

**Dossier zur Nutzenbewertung
gemäß § 35a SGB V**

*Aclidiniumbromid (Eklira[®] Genuair[®] 322 Mikrogramm
Pulver zur Inhalation /
Bretaris[®] Genuair[®] 322 Mikrogramm Pulver zur
Inhalation)*

Almirall Hermal GmbH

Modul 4 A

*Bronchodilatatorische Dauertherapie bei
Erwachsenen mit COPD*

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	4
Abbildungsverzeichnis	19
Abkürzungsverzeichnis	25
4 Modul 4 – allgemeine Informationen	28
4.1 Zusammenfassung der Inhalte von Modul 4	29
4.2 Methodik	33
4.2.1 Fragestellung	33
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	33
4.2.3 Informationsbeschaffung	39
4.2.3.1 Studien des pharmazeutischen Unternehmers	39
4.2.3.2 Bibliografische Literaturrecherche	40
4.2.3.3 Suche in Studienregistern	41
4.2.3.4 Selektion relevanter Studien	42
4.2.4 Bewertung der Aussagekraft der Nachweise.....	43
4.2.5 Informationssynthese und -analyse	45
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	45
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	45
4.2.5.3 Meta-Analysen.....	52
4.2.5.4 Sensitivitätsanalysen	55
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren	56
4.2.5.6 Indirekte Vergleiche	60
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	67
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	68
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	68
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	68
4.3.1.1.2 Studien aus der bibliographischen Literaturrecherche	74
4.3.1.1.3 Studien aus der Suche in Studienregistern	77
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	78
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	79
4.3.1.2.1 Studiendesign und Studienpopulationen	79
4.3.1.2.2 Verzerrungspotential auf Studienebene	85
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	85
4.3.1.3.1 Endpunkt Morbidität – RCT	87
4.3.1.3.1.1 CFB FEV ₁ AUC – RCT.....	87
4.3.1.3.1.2 CFB Trough-FEV ₁ – RCT.....	95
4.3.1.3.1.3 CFB Peak-FEV ₁ – RCT.....	99
4.3.1.3.1.4 CFB COPD-Symptomatik – RCT	103
4.3.1.3.2 Gesundheitsbezogene Lebensqualität - RCT	109

4.3.1.3.2.1	Behandlungszufriedenheit- RCT	109
4.3.1.3.3	Unerwünschte Arzneimittelwirkungen – RCT.....	114
4.3.1.3.4	Mortalität - RCT.....	120
4.3.1.3.4.1	COPD-assoziierte kardiovaskuläre Morbidität und Mortalität - RCT	120
4.3.1.3.4.2	COPD bedingte Letalität, Gesamtmortalität - RCT.....	123
4.3.1.4	Subgruppenanalysen – RCT	125
A priori geplante Subgruppenanalysen wurden in den direkten Vergleichsstudien nicht durchgeführt und können nicht dargestellt werden.		142
4.3.1.5	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien	143
4.3.2	Weitere Unterlagen.....	145
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	145
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	145
4.3.2.1.1.1	Studien des pharmazeutischen Unternehmers	145
4.3.2.1.1.2	Studien aus der bibliografischen Literaturrecherche	151
4.3.2.1.1.3	Studien aus der Suche in Studienregistern	154
4.3.2.1.1.4	Resultierender Studienpool: Studien für indirekte Vergleiche.....	157
4.3.2.1.2	Charakteristika der in die Bewertung eingeschlossenen Studien – Studien für indirekte Vergleiche	161
4.3.2.1.2.1	Studiendesign und Studienpopulationen	161
4.3.2.1.2.2	Verzerrungspotenzial auf Studienebene	184
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	188
4.3.2.1.3.1	CFB Trough-FEV ₁ – indirekte Vergleiche aus RCT.....	194
4.3.2.1.3.2	CFB Peak-FEV ₁ – indirekte Vergleiche aus RCT.....	214
4.3.2.1.3.3	TDI Focal Score – indirekte Vergleiche aus RCT.....	228
4.3.2.1.3.4	TDI -Responder – indirekte Vergleiche aus RCT	240
4.3.2.1.3.5	CFB EXACT-Respirationssymptome – indirekte Vergleiche aus RCT	252
4.3.2.1.3.6	CFB Nächtliche COPD-Symptome – indirekte Vergleiche aus RCT	256
4.3.2.1.3.7	CFB Morgendliche COPD-Symptome – indirekte Vergleiche aus RCT	260
4.3.2.1.3.8	Exazerbationsrate – indirekte Vergleiche aus RCT.....	265
4.3.2.1.3.9	Krankenhausaufenthalte aufgrund von Exazerbationen – indirekte Vergleiche aus RCT	280
4.3.2.1.3.10	SGRQ-Gesamtscore – indirekte Vergleiche aus RCT.....	290
4.3.2.1.3.11	SGRQ-Responder – indirekte Vergleiche aus RCT	304
4.3.2.1.3.12	Unerwünschte Ereignisse – indirekte Vergleiche aus RCT.....	318
4.3.2.1.3.13	Schwere unerwünschte Ereignisse – indirekte Vergleiche aus RCT	332
4.3.2.1.3.14	Studienabbrüche aufgrund unerwünschter Ereignisse – indirekte Vergleiche aus RCT.....	346
4.3.2.1.3.15	Kardiovaskuläre Morbidität und Mortalität – indirekte Vergleiche aus RCT.....	359
4.3.2.1.3.16	Letalität und Gesamtmortalität – indirekte Vergleiche aus RCT ..	365
4.3.2.1.3.17	Subgruppenanalysen – indirekte Vergleiche aus RCT	370

4.3.2.1.3.18 Zusammenfassung der Ergebnisse aus indirekten Vergleichen	423
4.3.2.2 Nicht randomisierte vergleichende Studien	433
4.3.2.3 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	433
4.3.2.3.1 Charakteristika der nicht randomisierten vergleichenden Studien.....	433
4.3.2.3.2 Ergebnisse aus nicht randomisierten vergleichenden Studien	434
4.3.2.3.2.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	434
4.3.2.3.2.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien	435
4.3.2.4 Weitere Untersuchungen.....	435
4.3.2.4.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen	435
4.3.2.4.2 Charakteristika der weiteren Untersuchungen	436
4.3.2.4.3 Ergebnisse aus weiteren Untersuchungen	436
4.3.2.4.3.1 <Endpunkt xxx> – weitere Untersuchungen	436
4.3.2.4.3.2 Subgruppenanalysen – weitere Untersuchungen	437
4.3.2.5 Zusammenfassung der Ergebnisse aus weiteren Unterlagen	437
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens	438
4.4.1 Beurteilung der Aussagekraft der Nachweise	438
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	440
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	443
4.4.4 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens....	444
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte.....	445
4.5.1 Begründung für die Vorlage indirekter Vergleiche.....	445
4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	446
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	446
4.5.4 Verwendung von Surrogatendpunkten	447
4.6 Liste der eingeschlossenen Studien	450
4.7 Referenzliste	454
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche	460
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....	467
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund	471
Anhang 4-D : Liste der abgebrochenen Studien	474
Anhang 4-E : Liste der laufenden Studien.....	475
Anhang 4-F : Methodik der eingeschlossenen Studien – RCT.....	476
Anhang 4-G : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten	650

Tabellenverzeichnis

	Seite
Tabelle 4.2-1: Patientenrelevante Endpunkte.....	36
Tabelle 4.2-2: Einschlusskriterien – Suche nach direkt vergleichenden RCTs	37
Tabelle 4.2-3: Ausschlusskriterien – Suche nach direkt vergleichenden RCTs	37
Tabelle 4.2-4: Einschlusskriterien – Suche nach RCTs für den indirekten Vergleich.....	38
Tabelle 4.2-5: Ausschlusskriterien – Suche nach RCTs für den indirekten Vergleich.....	39
Tabelle 4.2-6: Im Dossier berichtete patientenrelevante Endpunkte	47
Tabelle 4.2-7: A priori Subgruppenanalysen in der ATTAIN-Studie.....	57
Tabelle 4.2-8: Subgruppen nach COPD-Schwergrad-Einteilung der NVL	58
Tabelle 4.3-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	69
Tabelle 4.3-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel für den direkten Vergleich	73
Tabelle 4.3-3: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	78
Tabelle 4.3-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	79
Tabelle 4.3-5: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	80
Tabelle 4.3-6: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	82
Tabelle 4.3-7: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel	83
Tabelle 4.3-8: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	85
Tabelle 4.3-9: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	85
Tabelle 4.3-10: Im indirekten Vergleich herangezogene Endpunkte.....	86
Tabelle 4.3-11: Operationalisierung von CFB FEV ₁ AUC.....	88
Tabelle 4.3-12: Bewertung des Verzerrungspotenzials für CFB FEV ₁ AUC in RCT mit direktem Vergleich.....	89
Tabelle 4.3-13: Ergebnisse für CFB FEV ₁ AUC (I) aus RCT mit direktem Vergleich	90
Tabelle 4.3-14: Operationalisierung von CFB Trough-FEV ₁	96
Tabelle 4.3-15: Bewertung des Verzerrungspotenzials für CFB Trough-FEV ₁ in RCT mit direktem Vergleich.....	96
Tabelle 4.3-16: Ergebnisse für CFB Trough-FEV ₁ (I) aus RCT mit direktem Vergleich	97
Tabelle 4.3-17: Operationalisierung von CFB Peak-FEV ₁	99

Tabelle 4.3-18: Bewertung des Verzerrungspotenzials für CFB Peak-FEV ₁ in RCT mit direktem Vergleich.....	100
Tabelle 4.3-19: Ergebnisse für CFB Peak-FEV ₁ (I) aus RCT mit direktem Vergleich	101
Tabelle 4.3-20: Operationalisierung von COPD-Symptomatik	103
Tabelle 4.3-21: Bewertung des Verzerrungspotenzials für COPD-Symptomatik in RCT mit direktem Vergleich.....	105
Tabelle 4.3-22: Ergebnisse für COPD-Symptomatik aus RCT mit direktem Vergleich	106
Tabelle 4.3-23: Operationalisierung der Behandlungszufriedenheit.....	110
Tabelle 4.3-24: Bewertung des Verzerrungspotenzials für die Behandlungszufriedenheit in RCT mit direktem Vergleich.....	110
Tabelle 4.3-25: Ergebnisse für die Behandlungszufriedenheit aus RCT mit direktem Vergleich	111
Tabelle 4.3-26: Operationalisierung von unerwünschten Arzneimittelwirkungen	114
Tabelle 4.3-27: Bewertung des Verzerrungspotenzials für unerwünschte Arzneimittelwirkungen in RCT mit direktem Vergleich	117
Tabelle 4.3-28: Ergebnisse für unerwünschte Arzneimittelwirkungen aus RCT mit direktem Vergleich	118
Tabelle 4.3-29: Operationalisierung von COPD-assozierte kardiovaskuläre Morbidität und Mortalität.....	120
Tabelle 4.3-30: Bewertung des Verzerrungspotenzials für COPD-assozierte kardiovaskuläre Morbidität und Mortalität in RCT mit dem zu bewertenden Arzneimittel..	121
Tabelle 4.3-31: Ergebnisse für COPD-assozierte kardiovaskuläre Morbidität und Mortalität aus RCT mit einem direkten Vergleich.....	122
Tabelle 4.3-32: Operationalisierung von COPD bedingte Letalität, Gesamtmortalität	123
Tabelle 4.3-33: Bewertung des Verzerrungspotenzials für COPD bedingte Letalität, Gesamtmortalität in RCT mit dem zu bewertenden Arzneimittel.....	124
Tabelle 4.3-34: Darstellung der Patientenpopulation der M/34273/39 nach definierten Subgruppen.....	126
Tabelle 4.3-35: Subgruppenergebnisse für FEV ₁ AUC (I) aus RCT mit direktem Vergleich	127
Tabelle 4.3-36: Subgruppenergebnisse für Trough-FEV ₁ (I) aus RCT mit direktem Vergleich	131
Tabelle 4.3-37: Subgruppenergebnisse für Peak-FEV ₁ (I) aus RCT mit direktem Vergleich	132
Tabelle 4.3-38: Subgruppenergebnisse für COPD-Symptome aus RCT mit direktem Vergleich	134
Tabelle 4.3-39: Subgruppenergebnisse für die Behandlungszufriedenheit aus RCT mit einem direkten Vergleich	138
Tabelle 4.3-40: Subgruppenergebnisse für unerwünschte Arzneimittelwirkungen aus RCT mit einem direkten Vergleich	141

Tabelle 4.3-41: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	146
Tabelle 4.3-42: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel für indirekte Vergleiche	150
Tabelle 4.3-43: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel und der Vergleichstherapie für den indirekten Vergleich	155
Tabelle 4.3-44: Studienpool – RCT mit dem zu bewertenden Arzneimittel und der Vergleichstherapie für den indirekten Vergleich	158
Tabelle 4.3-45: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel für den indirekten Vergleich.....	162
Tabelle 4.3-46: Charakterisierung der Interventionen – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel für den indirekten Vergleich	173
Tabelle 4.3-47: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel für den indirekten Vergleich.....	178
Tabelle 4.3-48: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel für den indirekten Vergleich	184
Tabelle 4.3-49: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche nach 12 Wochen	190
Tabelle 4.3-50: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche nach 24 Wochen	192
Tabelle 4.3-51: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Trough-FEV ₁ – 12 Wochen – begleitende Behandlung mit LABAs nicht gestattet].....	194
Tabelle 4.3-52: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Trough-FEV ₁ – 12 Wochen – alle Studien]	195
Tabelle 4.3-53: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Trough-FEV ₁ - 24 Wochen- begleitende Behandlung mit LABAs nicht gestattet]	196
Tabelle 4.3-54: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Trough-FEV ₁ - 24 Wochen – alle Studien]	197
Tabelle 4.3-55: Operationalisierung von CFB Trough-FEV ₁	199
Tabelle 4.3-56: Bewertung des Verzerrungspotenzials für CFB Trough-FEV ₁ in RCT für indirekte Vergleiche	201
Tabelle 4.3-57: Ergebnisse für CFB Trough-FEV ₁ (ml) aus RCT für indirekte Vergleiche (12 Wochen).....	202
Tabelle 4.3-58: Ergebnisse für CFB Trough-FEV ₁ (ml) aus RCT für indirekte Vergleiche (24 Wochen).....	204
Tabelle 4.3-59: DIC für CFB Trough-FEV ₁ (ml) – indirekter Vergleich aus RCT	210

Tabelle 4.3-60: CFB Trough-FEV ₁ (ml) Indirekter Vergleich aus RCT (12 Wochen); Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit zufälligen Effekten ...	210
Tabelle 4.3-61: CFB Trough-FEV ₁ (ml) Indirekter Vergleich aus RCT (12 Wochen); Szenario 2: Alle Studien; Modell mit zufälligen Effekten	211
Tabelle 4.3-62: CFB Trough-FEV ₁ (ml) Indirekter Vergleich aus RCT (24 Wochen); Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	212
Tabelle 4.3-63: CFB Trough-FEV ₁ (ml) Indirekter Vergleich aus RCT (24 Wochen); Szenario 2: Alle Studien; Modell mit zufälligen Effekten	212
Tabelle 4.3-64: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Peak-FEV ₁ - 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]	214
Tabelle 4.3-65: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Peak-FEV ₁ - 12 Wochen - alle Studien]	215
Tabelle 4.3-66: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Peak-FEV ₁ - 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]	216
Tabelle 4.3-67: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Peak-FEV ₁ - 24 Wochen - alle Studien]	216
Tabelle 4.3-68: Operationalisierung des CFB Peak-FEV ₁	218
Tabelle 4.3-69: Bewertung des Verzerrungspotenzials für CFB Peak-FEV ₁ in RCT für indirekte Vergleiche	219
Tabelle 4.3-70: Ergebnisse für CFB Peak-FEV ₁ (ml) aus RCT für indirekte Vergleiche (12 Wochen).....	220
Tabelle 4.3-71: Ergebnisse für CFB Peak-FEV ₁ (ml) aus RCT für indirekte Vergleiche (24 Wochen).....	221
Tabelle 4.3-72: DIC für CFB (ml) Peak-FEV ₁ - Indirekter Vergleich aus RCT	225
Tabelle 4.3-73: CFB (ml) Peak-FEV ₁ (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit zufälligen Effekten ...	225
Tabelle 4.3-74: CFB (ml) Peak-FEV ₁ (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: alle Studien; Modell mit zufälligen Effekten	226
Tabelle 4.3-75: CFB (ml) Peak-FEV ₁ (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit zufälligen Effekten ...	226
Tabelle 4.3-76: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI Focal Score – 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]	228
Tabelle 4.3-77: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI Focal Score – 12 Wochen - alle Studien]	229
Tabelle 4.3-78: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI Focal Score – 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]	230

Tabelle 4.3-79: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI Focal Score – 24 Wochen - alle Studien]	231
Tabelle 4.3-80: Operationalisierung des TDI Focal Score.....	232
Tabelle 4.3-81: Bewertung des Verzerrungspotenzials für TDI Focal Score in RCT für indirekte Vergleiche	233
Tabelle 4.3-82: Ergebnisse für TDI Focal Score aus RCT für indirekte Vergleiche (12 Wochen)	234
Tabelle 4.3-83: Ergebnisse für den TDI Focal Score aus RCT für indirekte Vergleiche (24 Wochen)	235
Tabelle 4.3-84: DIC des indirekten Therapievergleichs zwischen Acridiniumbromid und Tiotropiumbromid 18µg; Differenz des TDI Focal Scores	238
Tabelle 4.3-85: TDI Focal Score (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	238
Tabelle 4.3-86: TDI Focal Score (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	239
Tabelle 4.3-87: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI-Responder (%) - 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet].....	240
Tabelle 4.3-88: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI-Responder (%) - 12 Wochen - alle Studien]	241
Tabelle 4.3-89: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI-Responder (%) – 24 Wochen – begleitende Behandlung mit LABAs nicht gestattet].....	242
Tabelle 4.3-90: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI-Responder (%) – 24 Wochen – alle Studien]	242
Tabelle 4.3-91: Operationalisierung von Prozentsatz der TDI-Responder	244
Tabelle 4.3-92: Bewertung des Verzerrungspotenzials für TDI-Responder in RCT für indirekte Vergleiche	245
Tabelle 4.3-93: Ergebnisse für TDI-Responder aus RCT für indirekte Vergleiche (12 Wochen)	246
Tabelle 4.3-94: Ergebnisse für TDI-Responder aus RCT für indirekte Vergleiche (24 Wochen)	247
Tabelle 4.3-95: DIC des indirekten Therapievergleichs zwischen Acridiniumbromid und Tiotropiumbromid 18µg; Anteil der TDI-Responder	249
Tabelle 4.3-96: TDI-Responder (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	250
Tabelle 4.3-97: TDI-Responder (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	250

Tabelle 4.3-98: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB EXACT-Respirationssymptome – 12 Wochen]	252
Tabelle 4.3-99: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB EXACT-Respirationssymptome – 24 Wochen]	252
Tabelle 4.3-100: Operationalisierung von CFB EXACT-Respirationssymptomen.....	253
Tabelle 4.3-101: Bewertung des Verzerrungspotenzials für CFB EXACT-Respirationssymptome in RCT für indirekte Vergleiche	253
Tabelle 4.3-102: Ergebnisse für CFB EXACT-Respirationssymptome aus RCT für indirekte Vergleiche (CFB nach 12 Wochen)	254
Tabelle 4.3-103: Ergebnisse für CFB EXACT-Respirationssymptome aus RCT für indirekte Vergleiche (CFB nach 24 Wochen)	254
Tabelle 4.3-104: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Nächtliche COPD-Symptome – 12 Wochen]	256
Tabelle 4.3-105: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Nächtliche COPD-Symptome – 24 Wochen]	256
Tabelle 4.3-106: Operationalisierung von CFB nächtliche COPD-Symptome	257
Tabelle 4.3-107: Bewertung des Verzerrungspotenzials für CFB nächtliche COPD-Symptome in RCT für indirekte Vergleiche	257
Tabelle 4.3-108: Ergebnisse für CFB nächtliche COPD-Symptome (12 Wochen) aus RCT für indirekte Vergleiche	258
Tabelle 4.3-109: Ergebnisse für CFB nächtliche COPD-Symptome (24 Wochen) aus RCT für indirekte Vergleiche	259
Tabelle 4.3-110: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Morgendliche COPD-Symptome – 12 Wochen]	260
Tabelle 4.3-111: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Morgendliche COPD-Symptome – 24 Wochen]	261
Tabelle 4.3-112: Operationalisierung von CFB morgendliche COPD-Symptome.....	261
Tabelle 4.3-113: Bewertung des Verzerrungspotenzials für CFB morgendliche COPD-Symptome in RCT für indirekte Vergleiche	262
Tabelle 4.3-114: Ergebnisse für CFB morgendliche COPD-Symptome (12 Wochen) aus RCT für indirekte Vergleiche.....	262
Tabelle 4.3-115: Ergebnisse für CFB morgendliche COPD-Symptome (24 Wochen) aus RCT für indirekte Vergleiche.....	263
Tabelle 4.3-116: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Exazerbationsrate; begleitende Behandlung mit LABAs nicht gestattet]	265

Tabelle 4.3-117: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Exazerbationsrate; alle Studien]	266
Tabelle 4.3-118: Operationalisierung der Exazerbationsraten	268
Tabelle 4.3-119: Bewertung des Verzerrungspotenzials für Exazerbationsraten in RCT für indirekte Vergleiche	270
Tabelle 4.3-120: Ergebnisse für die Exazerbationsrate aus RCT für indirekte Vergleiche ...	271
Tabelle 4.3-121: DIC für Exazerbationsraten – Indirekter Vergleich aus RCT.....	277
Tabelle 4.3-122: Exazerbationsraten - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne LABA-Begleitmedikation; Modell mit zufälligen Effekten.....	277
Tabelle 4.3-123: Exazerbationsraten - Indirekter Vergleich aus RCT; Szenario 2: alle Studien; Modell mit zufälligen Effekten	278
Tabelle 4.3-124: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Krankenhausaufenthalte aufgrund von Exazerbationen – 12 Wochen: alle Studien]	280
Tabelle 4.3-125: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Krankenhausaufenthalte aufgrund von Exazerbationen – 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]	280
Tabelle 4.3-126: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Krankenhausaufenthalte aufgrund von Exazerbationen – 24 Wochen: alle Studien]	281
Tabelle 4.3-127: Operationalisierung von Krankenhauseinweisungen aufgrund von Exazerbationen	283
Tabelle 4.3-128: Bewertung des Verzerrungspotenzials für Krankenhauseinweisungen aufgrund von Exazerbationen in RCT für indirekte Vergleiche	284
Tabelle 4.3-129: Ergebnisse für Patienten mit Krankenhausaufenthalten aufgrund von Exazerbation aus RCT für indirekte Vergleiche (12 Wochen)	285
Tabelle 4.3-130: Ergebnisse für Patienten mit Krankenhausaufenthalten aufgrund von Exazerbation aus RCT für indirekte Vergleiche (24 Wochen)	286
Tabelle 4.3-131: DIC des indirekten Vergleichs von Acclidiniumbromid und Tiotropiumbromid 18µg; Krankenhausaufenthalte aufgrund von Exazerbationen	288
Tabelle 4.3-132: Krankenhausaufenthalte aufgrund von Exazerbationen (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	288
Tabelle 4.3-133: Krankenhausaufenthalte aufgrund von Exazerbationen (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: alle Studien; Modell mit festen Effekten.....	289
Tabelle 4.3-134: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Gesamtwert– 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet].....	290
Tabelle 4.3-135: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Gesamtwert – 12 Wochen – alle Studien]	291

Tabelle 4.3-136: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Gesamtwert – 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet].....	292
Tabelle 4.3-137: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Gesamtwert – 24 Wochen – alle Studien]	293
Tabelle 4.3-138: Operationalisierung des SGRQ-Gesamtwerts	294
Tabelle 4.3-139: Bewertung des Verzerrungspotenzials für den SGRQ-Gesamtwert in RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel	296
Tabelle 4.3-140: Ergebnisse für SGRQ-Gesamtwert (12 Wochen) aus RCT für indirekte Vergleiche	297
Tabelle 4.3-141: Ergebnisse für SGRQ (24 Wochen) - Gesamtwert aus RCT für indirekte Vergleiche	298
Tabelle 4.3-142: DIC des SGRQ-Gesamtwerts – indirekte Vergleich aus RCT	301
Tabelle 4.3-143: SGRQ-Gesamtwert (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit zufälligen Effekten ...	302
Tabelle 4.3-144: SGRQ-Gesamtwert (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	302
Tabelle 4.3-145: SGRQ-Gesamtwert (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: Alle Studien; Modell mit festen Effekten	303
Tabelle 4.3-146: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Responder (%) – 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet].....	304
Tabelle 4.3-147: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Responder (%) – 12 Wochen – alle Studien].....	305
Tabelle 4.3-148: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Responder (%) – 24 Wochen – begleitende Behandlung mit LABAs nicht gestattet].....	306
Tabelle 4.3-149: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Responder (%) – 24 Wochen - alle Studien].....	307
Tabelle 4.3-150: Operationalisierung von SGRQ Respondern.....	308
Tabelle 4.3-151: Bewertung des Verzerrungspotenzials für den SGRQ-Wert in RCT für indirekte Vergleiche	310
Tabelle 4.3-152: Ergebnisse für SGRQ-Responder aus RCT für indirekte Vergleiche (12 Wochen)	311
Tabelle 4.3-153: Ergebnisse für SGRQ-Responder aus RCT für indirekte Vergleiche (24 Wochen)	312
Tabelle 4.3-154: DIC für SGRQ-Responder - indirekter Vergleich aus RCT	315

Tabelle 4.3-155: SGRQ-Responder (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	315
Tabelle 4.3-156: SGRQ-Responder (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	316
Tabelle 4.3-157: SGRQ-Responder (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: Alle Studien; Modell mit festen Effekten	316
Tabelle 4.3-158: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Unerwünschte Ereignisse - 12 Wochen- begleitende Behandlung mit LABAs nicht gestattet].....	318
Tabelle 4.3-159: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Unerwünschte Ereignisse - 12 Wochen - Alle Studien]	319
Tabelle 4.3-160: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Unerwünschte Ereignisse - 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet].....	320
Tabelle 4.3-161: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Unerwünschte Ereignisse - 24 Wochen – alle Studien, inkl. alle Studien]	320
Tabelle 4.3-162: Operationalisierung von unerwünschten Ereignissen.....	322
Tabelle 4.3-163: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse in RCT für indirekte Vergleiche.....	324
Tabelle 4.3-164: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (12 Wochen)	325
Tabelle 4.3-165: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (24 Wochen)	326
Tabelle 4.3-166: DIC des indirekten Vergleichs von Acridiniumbromid und Tiotropiumbromid 5µg und 18µg (gepoolt); unerwünschte Ereignisse.....	329
Tabelle 4.3-167: Unerwünschte Ereignisse nach 12 Wochen - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten	330
Tabelle 4.3-168: Unerwünschte Ereignisse nach 12 Wochen - Indirekter Vergleich aus RCT; Szenario 2: alle Studien: Modell mit festen Effekten.....	330
Tabelle 4.3-169: Unerwünschte Ereignisse nach 24 Wochen - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA: Modell mit festen Effekten	331
Tabelle 4.3-170: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Schwere unerwünschte Ereignisse - 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]	332
Tabelle 4.3-171: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Schwere unerwünschte Ereignisse - 12 Wochen - alle Studien]	333
Tabelle 4.3-172: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Schwere unerwünschte Ereignisse - 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]	334

Tabelle 4.3-173: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Schwere unerwünschte Ereignisse - 24 Wochen – alle Studien]	334
Tabelle 4.3-174: Operationalisierung von schweren unerwünschten Ereignissen.....	336
Tabelle 4.3-175: Bewertung des Verzerrungspotenzials für schwere unerwünschte Ereignisse in RCT für indirekte Vergleiche	338
Tabelle 4.3-176: Ergebnisse für schwere unerwünschte Ereignisse aus RCT für indirekte Vergleiche (12 Wochen)	339
Tabelle 4.3-177: Ergebnisse für schwere unerwünschte Ereignisse aus RCT für indirekte Vergleiche (24 Wochen)	340
Tabelle 4.3-178: DIC des indirekten Therapievergleichs von Aclidiniumbromid und Tiotropiumbromid 5µg und 18µg (gepoolt); Unterschied des mittleren CFB (ml) bei schwerwiegenden unerwünschten Ereignissen.	343
Tabelle 4.3-179: Schwere unerwünschte Ereignisse nach 12 Wochen - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	343
Tabelle 4.3-180: Schwere unerwünschte Ereignisse nach 12 Wochen - Indirekter Vergleich aus RCT; Szenario 2: alle Studien: Modell mit festen Effekten	344
Tabelle 4.3-181: Schwere unerwünschte Ereignisse nach 24 Wochen - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA: Modell mit festen Effekten.....	344
Tabelle 4.3-182: Schwere unerwünschte Ereignisse nach 24 Wochen - Indirekter Vergleich aus RCT; Szenario 2: alle Studien: Modell mit festen Effekten	345
Tabelle 4.3-183: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Ausschlüsse aufgrund von AEs – 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]	346
Tabelle 4.3-184: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Ausschlüsse aufgrund von AEs – 12 Wochen – alle Studien]	347
Tabelle 4.3-185: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Ausschlüsse aufgrund von AEs – 24 Wochen – begleitende Behandlung mit LABAs nicht gestattet]	348
Tabelle 4.3-186: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Ausschlüsse aufgrund von AEs – 24 Wochen – alle Studien]	348
Tabelle 4.3-187: Operationalisierung von Studienabbrüchen aufgrund unerwünschter Ereignisse	350
Tabelle 4.3-188: Bewertung des Verzerrungspotenzials für Ausschlüsse aufgrund von unerwünschten Ereignissen in RCT für indirekte Vergleiche.....	351
Tabelle 4.3-189: Ergebnisse für Studienabbrüche aufgrund von unerwünschten Ereignissen aus der RCT für indirekte Vergleiche (12 Wochen).....	352

Tabelle 4.3-190: Ergebnisse für Studienabbrüche aufgrund von unerwünschten Ereignissen aus der RCT für indirekte Vergleiche (24 Wochen).....	353
Tabelle 4.3-191: DIC des indirekten Vergleichs von Acridiniumbromid und Tiotropiumbromid 5µg bzw. 18µg (gepoolt); Studienabbruch aufgrund unerwünschter Ereignisse	356
Tabelle 4.3-192: Studienabbrüche aufgrund unerwünschter Ereignisse (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten.....	357
Tabelle 4.3-193: Studienabbrüche aufgrund unerwünschter Ereignisse (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: alle Studien: Modell mit festen Effekten.....	357
Tabelle 4.3-194: Studienabbrüche aufgrund unerwünschter Ereignisse (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA: Modell mit festen Effekten.....	358
Tabelle 4.3-195: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Kardiovaskuläre Morbidität und Mortalität – 12 Wochen]	359
Tabelle 4.3-196: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Kardiovaskuläre Morbidität und Mortalität – 24 Wochen]	359
Tabelle 4.3-197: Operationalisierung von kardiovaskulärer Morbidität und Mortalität.....	360
Tabelle 4.3-198: Bewertung des Verzerrungspotenzials für für kardiovaskuläre Morbidität und Mortalität in RCT für indirekte Vergleiche.....	361
Tabelle 4.3-199: Ergebnisse für kardiovaskuläre Morbidität und Mortalität aus RCT für indirekte Vergleiche (CFB nach 12 Wochen)	362
Tabelle 4.3-200: Ergebnisse für kardiovaskuläre Morbidität und Mortalität aus RCT für indirekte Vergleiche (CFB nach 24 Wochen)	362
Tabelle 4.3-201: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Letalität und Gesamtmortalität (12 Wochen)].....	365
Tabelle 4.3-202: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Letalität und Gesamtmortalität (24 Wochen)].....	366
Tabelle 4.3-203: Operationalisierung von Letalität und Gesamtmortalität.....	366
Tabelle 4.3-204: Bewertung des Verzerrungspotenzials für Letalität und Gesamtmortalität in RCT für indirekte Vergleiche.....	367
Tabelle 4.3-205: Ergebnisse für Letalität und Gesamt-Mortalität aus RCT für indirekte Vergleiche (CFB nach 12 Wochen)	368
Tabelle 4.3-206: Ergebnisse für Letalität und Gesamt-Mortalität aus RCT für indirekte Vergleiche (CFB nach 24 Wochen)	369
Tabelle 4.3-207: Darstellung der Patientenpopulation der Studien ACCORD I, ATTAIN und ACCORD II nach definierten Subgruppen	370

Tabelle 4.3-208: Ergebnisse der Subgruppenanalyse für CFB Trough-FEV ₁ aus RCT für indirekte Vergleiche	372
Tabelle 4.3-209: Ergebnisse der Subgruppenanalyse für CFB Peak-FEV ₁ aus RCT für indirekte Vergleiche	375
Tabelle 4.3-210: Ergebnisse der Subgruppenanalyse für den TDI-Score aus RCT für indirekte Vergleiche	378
Tabelle 4.3-211: Ergebnisse der Subgruppenanalyse für TDI-Responder aus RCT für indirekte Vergleiche	381
Tabelle 4.3-212: Ergebnisse der Subgruppenanalyse für EXACT-Respirationssymptome (12 Wochen) aus RCT für indirekte Vergleiche	384
Tabelle 4.3-213: Ergebnisse der Subgruppenanalyse für EXACT-Respirationssymptome (24 Wochen) aus RCT für indirekte Vergleiche	386
Tabelle 4.3-214: Ergebnisse der Subgruppenanalyse für nächtliche COPD-Symptome (12 Wochen) aus RCT für indirekte Vergleiche	387
Tabelle 4.3-215: Ergebnisse der Subgruppenanalyse für nächtliche COPD-Symptome (24 Wochen) aus RCT für indirekte Vergleiche	389
Tabelle 4.3-216: Ergebnisse der Subgruppenanalyse für morgendliche COPD-Symptome (12 Wochen) aus RCT für indirekte Vergleiche	390
Tabelle 4.3-217: Ergebnisse der Subgruppenanalyse für morgendliche COPD-Symptome (24 Wochen) aus RCT für indirekte Vergleiche	392
Tabelle 4.3-218: Ergebnisse der Subgruppenanalyse für die Exazerbationsrate aus RCT für indirekte Vergleiche	394
Tabelle 4.3-219: Ergebnis der Subgruppenanalyse für Krankenhausaufenthalte aufgrund von Exazerbationen (12 Wochen) aus RCT für indirekte Vergleiche	397
Tabelle 4.3-220: Ergebnis der Subgruppenanalyse für Krankenhausaufenthalte aufgrund von Exazerbationen (24 Wochen) aus RCT für indirekte Vergleiche	398
Tabelle 4.3-221: Ergebnisse der Subgruppenanalyse für den SGRQ-Score aus RCT für indirekte Vergleiche	399
Tabelle 4.3-222: Ergebnisse der Subgruppenanalyse für SGRQ-Responder aus RCT für indirekte Vergleiche	402
Tabelle 4.3-223: Ergebnisse der Subgruppenanalyse für Unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 24 Wochen)	404
Tabelle 4.3-224: Ergebnisse der Subgruppenanalyse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 24 Wochen)	406
Tabelle 4.3-225: Ergebnisse der Subgruppenanalyse für schwere unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 12 Wochen)	407
Tabelle 4.3-226: Ergebnisse der Subgruppenanalyse für schwere unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 24 Wochen)	409
Tabelle 4.3-227: Ergebnisse der Subgruppenanalyse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 12 Wochen)	410

Tabelle 4.3-228: Ergebnisse der Subgruppenanalyse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 24 Wochen)	412
Tabelle 4.3-229: Ergebnisse der a priori gelanten Subgruppenanalyse für CFB Trough-FEV1 (12 Wochen) aus RCT für indirekte Vergleiche	414
Tabelle 4.3-230: Ergebnisse der a priori gelanten Subgruppenanalyse für CFB Trough-FEV1 (24 Wochen) aus RCT für indirekte Vergleiche	415
Tabelle 4.3-231: Ergebnisse der a priori gelanten Subgruppenanalyse für CFB Peak-FEV1 (24 Wochen) aus RCT für indirekte Vergleiche	417
Tabelle 4.3-232: Ergebnisse der a priori gelanten Subgruppenanalyse für SGRQ-Responder (24 Wochen) aus RCT für indirekte Vergleiche	419
Tabelle 4.3-233: Ergebnisse der a priori gelanten Subgruppenanalyse für TDI-Responder (24 Wochen) aus RCT für indirekte Vergleiche nach 24 Wochen	420
Tabelle 4.3-234: Im direkten Vergleich herangezogene Endpunkte	423
Tabelle 4.3-235: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	434
Tabelle 4.3-236: Operationalisierung von <Endpunkt xxx>	434
Tabelle 4.3-237: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	435
Tabelle 4.3-238: Operationalisierung von <Endpunkt xxx>	436
Tabelle 4.4-1: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens	443
Tabelle 4.4-2: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens	444
Tabelle 4.5-1: Übersicht über die Aussagen zur Validierung/ Korrelation/ Relevanz von FEV ₁	448
Tabelle 4.7-1 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel	474
Tabelle 4.7-2 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel	475
Tabelle 4.7-3 (Anhang): Studiendesign und -methodik für Studie <Studienbezeichnung> ..	477
Tabelle 4.7-4 (Anhang): Studiendesign und -methodik für Studie M/34273/23	480
Tabelle 4.7-5 (Anhang): Studiendesign und -methodik für Studie M/34273/39	488
Tabelle 4.7-6 (Anhang): Studiendesign und -methodik für Studie ATTAIN	497
Tabelle 4.7-7 (Anhang): Studiendesign und -methodik für Studie ACCORD I	506
Tabelle 4.7-8 (Anhang): Studiendesign und -methodik für Studie ACCORD II	514
Tabelle 4.7-9 (Anhang): Studiendesign und -methodik für Studie Brusasco 2003	521
Tabelle 4.7-10 (Anhang): Studiendesign und -methodik für Studie Casaburi 2000	527
Tabelle 4.7-11 (Anhang): Studiendesign und -methodik für Studie Casaburie 2002	533
Tabelle 4.7-12 (Anhang): Studiendesign und -methodik für Studie SAFE (Chan 2007)	539

Tabelle 4.7-13 (Anhang): Studiendesign und -methodik für Studie Covelli 2005	545
Tabelle 4.7-14 (Anhang): Studiendesign und -methodik für Studie Donohue 2002	551
Tabelle 4.7-15 (Anhang): Studiendesign und -methodik für Studie INHANCE (Donohue 2010).....	557
Tabelle 4.7-16 (Anhang): Studiendesign und -methodik für Studie SPRUCE (Freeman 2007).....	563
Tabelle 4.7-17 (Anhang): Studiendesign und -methodik für Studie SAFE Portugal (Moita 2008).....	569
Tabelle 4.7-18 (Anhang): Studiendesign und -methodik für Studie Niewoehner 2005.....	575
Tabelle 4.7-19 (Anhang): Studiendesign und -methodik für Studie UPLIFT (Tashkin 2008).....	581
Tabelle 4.7-20 (Anhang): Studiendesign und -methodik für Studie TIPHON (Tonnel 2008).....	590
Tabelle 4.7-21 (Anhang): Studiendesign und -methodik für Studie Verkindre 2006.....	595
Tabelle 4.7-22 (Anhang): Studiendesign und -methodik für Studie Vogelmeier 2008	601
Tabelle 4.7-23 (Anhang): Studiendesign und -methodik für Studie Dusser 2006.....	608
Tabelle 4.7-24 (Anhang): Studiendesign und -methodik für Studie Powrie 2007.....	614
Tabelle 4.7-25 (Anhang): Studiendesign und -methodik für Studie Batemen 2010a	619
Tabelle 4.7-26 (Anhang): Studiendesign und -methodik für Studie Bateman 2010b	626
Tabelle 4.7-27 (Anhang): Studiendesign und -methodik für Studie Voshaar 2008.....	632
Tabelle 4.7-28 (Anhang): Studiendesign und -methodik für Studie Johansson 2008.....	638
Tabelle 4.7-29 (Anhang): Studiendesign und -methodik für Studie Beeh 2006.....	644
Tabelle 4.7-30 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie M/3473/23	653
Tabelle 4.7-31 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie M/3473/39	690
Tabelle 4.7-32 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie ACCORD I.....	726
Tabelle 4.7-33 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie ACCORD II.....	772
Tabelle 4.7-34 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie ATTAIN	817
Tabelle 4.7-35 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Bateman 2010a.....	866
Tabelle 4.7-36 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Bateman 2010b.....	887
Tabelle 4.7-37 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Brusasco 2003	908

Tabelle 4.7-38 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Beeh 2006.....	933
Tabelle 4.7-39 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Casaburi 2000.....	945
Tabelle 4.7-40 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Casaburi 2002.....	960
Tabelle 4.7-41 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SAFE (Chan 2007).....	975
Tabelle 4.7-42 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Covelli 2005.....	984
Tabelle 4.7-43 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie INHANCE (Donohue 2010).....	1001
Tabelle 4.7-44 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Donohue 2002.....	1035
Tabelle 4.7-45 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MISTRAL (Dusser 2006).....	1058
Tabelle 4.7-46 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SPRUCE (Freeman 2007).....	1067
Tabelle 4.7-47 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Johansson 2008.....	1082
Tabelle 4.7-48 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SAFE Portugal (Moita 2008).....	1103
Tabelle 4.7-49 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Niewoehner 2005.....	1126
Tabelle 4.7-50 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Powrie 2007.....	1144
Tabelle 4.7-51 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UPLIFT (Tashkin 2008).....	1153
Tabelle 4.7-52 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie TIPHON (Tonnel 2008).....	1168
Tabelle 4.7-53 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Verkindre 2006.....	1185
Tabelle 4.7-54 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Vogelmeier 2008.....	1208
Tabelle 4.7-55 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Voshaar 2008.....	1228

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	75
Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach RCTs für direkten Vergleich	76
Abbildung 3: CFB FEV ₁ (l) über 24 Stunden an Tag 15 (ITT-Population).....	92
Abbildung 4: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	152
Abbildung 5: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach RCTs für indirekte Vergleiche	153
Abbildung 6: Netzwerkdiagramm für die Analyse des CFB Trough-FEV ₁ nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	195
Abbildung 7: Netzwerkdiagramm für die Analyse des CFB Trough-FEV ₁ nach 12 Wochen (alle Studien).....	196
Abbildung 8: Netzwerkdiagramm für die Analyse des CFB Trough-FEV ₁ nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	197
Abbildung 9: Netzwerkdiagramm für die Analyse des CFB Trough-FEV ₁ nach 24 Wochen (alle Studien).....	198
Abbildung 10: Meta-Analyse für CFB Trough-FEV ₁ nach 12 Wochen (ml) aus RCT; Acclidiniumbromid im Vergleich zu Placebo.....	205
Abbildung 11: Meta-Analyse für CFB Trough-FEV ₁ nach 12 Wochen (ml) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.	205
Abbildung 12: Meta-Analyse für CFB Trough-FEV ₁ nach 24 Wochen (ml) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.	206
Abbildung 13: Sensitivitätsanalyse (Verzerrungspotential) für CFB Trough-FEV ₁ nach 12 Wochen (ml) aus RCT; Acclidiniumbromid im Vergleich zu Placebo.....	207
Abbildung 14: Sensitivitätsanalyse (COPD-Schweregrad) für CFB Trough-FEV ₁ nach 12 Wochen (ml) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo.	208
Abbildung 15: Sensitivitätsanalyse (Verzerrungspotential) für CFB Trough-FEV ₁ nach 12 Wochen (ml) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo.	208
Abbildung 16: Sensitivitätsanalyse (Verzerrungspotential) für CFB Trough-FEV ₁ nach 24 Wochen (ml) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo.	209
Abbildung 17: Netzwerk-Diagramm für die Analyse des CFB Peak-FEV ₁ nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	215
Abbildung 18: Netzwerkdiagramm für die CFB Peak-FEV ₁ -Analyse nach 12 Wochen	215
Abbildung 19: Netzwerk-Diagramm für die Analyse des CFB Peak-FEV ₁ nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	216
Abbildung 20: Netzwerkdiagramm für die CFB Peak-FEV ₁ -Analyse nach 24 Wochen	217

Abbildung 21: Meta-Analyse für CFB Peak-FEV ₁ nach 12 Wochen (ml) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	221
Abbildung 22: Meta-Analyse für CFB Peak-FEV ₁ nach 12 Wochen (ml) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo	222
Abbildung 23: Meta-Analyse für CFB Peak-FEV ₁ nach 24 Wochen (ml) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.....	222
Abbildung 24: Sensitivitätsanalyse (Verzerrungspotential) für CFB Peak-FEV ₁ (ml) nach 12 Wochen aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	223
Abbildung 25: Sensitivitätsanalyse (Dosierung) für CFB Peak-FEV ₁ (ml) nach 12 Wochen aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.	224
Abbildung 26: Netzwerkdiagramm für die Analyse des TDI Focal Score nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	229
Abbildung 27: Netzwerkdiagramm für die Analyse des TDI Focal Score nach 12 Wochen	229
Abbildung 28: Netzwerkdiagramm für die Analyse des TDI Focal Score nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	231
Abbildung 29: Netzwerkdiagramm für die Analyse des TDI Focal Score nach 24 Wochen	231
Abbildung 30: Meta-Analyse für TDI Focal Score (12 Wochen) aus RCT; Aclidiniumbromid 400µg im Vergleich zu Placebo.....	235
Abbildung 31: Meta-Analyse für TDI Focal Score (12 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.....	236
Abbildung 32: Meta-Analyse für TDI Focal Score (24 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.....	236
Abbildung 33: Sensitivitätsanalyse (Verzerrungspotential) für den TDI Focal Score nach 12 Wochen (ml) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.	237
Abbildung 34: Netzwerkdiagramm für die Analyse der TDI-Responder nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	241
Abbildung 35: Netzwerkdiagramm für die Analyse der TDI-Responder nach 12 Wochen ..	241
Abbildung 36: Netzwerkdiagramm für die Analyse der TDI-Responder nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	242
Abbildung 37: Netzwerkdiagramm für die Analyse der TDI-Responder nach 24 Wochen ..	243
Abbildung 38: Meta-Analyse für den Prozentsatz der TDI-Responder (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	247
Abbildung 39: Meta-Analyse für den Prozentsatz der TDI-Responder (24 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.....	248
Abbildung 40: Sensitivitätsanalyse (Verzerrungspotential) für TDI-Responder nach 12 Wochen aus RCT; Aclidiniumbromid im Vergleich zu Placebo.	248
Abbildung 41: Netzwerkdiagramm für die Analyse der Exazerbationsrate (begleitende Behandlung mit LABAs nicht gestattet)	266
Abbildung 42: Netzwerkdiagramm für die Analyse der Exazerbationsrate	267

Abbildung 43: Meta-Analyse für Exazerbationsraten aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	272
Abbildung 44: Meta-Analyse für Exazerbationsraten aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.	273
Abbildung 45: Sensitivitätsanalyse (Verzerrungspotential) für Exazerbationsraten aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	274
Abbildung 46: Sensitivitätsanalyse (Verzerrungspotential) für Exazerbationsraten aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.	275
Abbildung 47: Sensitivitätsanalyse (Dosierung) für Exazerbationsraten aus RCT; Tiotropiumbromid 5µg im Vergleich zu Placebo.	275
Abbildung 48: Sensitivitätsanalyse (Dosierung) für Exazerbationsraten aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.....	276
Abbildung 49: Netzwerk-Diagramm für die Analyse von Krankenhauseinweisungen aufgrund von Exazerbationen nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	281
Abbildung 50: Netzwerk-Diagramm für die Analyse von Krankenhauseinweisungen aufgrund von Exazerbationen nach 24 Wochen (alle Studien)	282
Abbildung 51: Meta-Analyse für die Anzahl von Patienten, die aufgrund von unerwünschten Ereignissen nach 12 Wochen in ein Krankenhaus eingeliefert wurden; Aclidiniumbromid im Vergleich zu Placebo.....	286
Abbildung 52: Meta-Analyse für die Anzahl von Patienten, die aufgrund von Exazerbationen in ein Krankenhaus eingeliefert wurden (24 Wochen); Tiotropiumbromid 18µg im Vergleich zu Placebo	287
Abbildung 53: Netzwerkdiagramm für die Analyse des SGRQ-Gesamtwerts nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	291
Abbildung 54: Netzwerkdiagramm für die Analyse des SGRQ-Gesamtwerts nach 12 Wochen (alle Studien).....	292
Abbildung 55: Netzwerkdiagramm für die Analyse des SGRQ-Gesamtwerts nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	293
Abbildung 56: Netzwerkdiagramm für die Analyse des SGRQ-Gesamtwerts nach 24 Wochen (alle Studien).....	294
Abbildung 57: Meta-Analyse für den SGRQ-Score (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	299
Abbildung 58: Meta-Analyse für SGRQ-Score (12 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo	299
Abbildung 59: Meta-Analyse für den SGRQ-Score (24 Wochen) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.	300
Abbildung 60: Sensitivitätsanalyse (Verzerrungspotential) für SGRQ-Gesamtwert nach 12 Wochen (ml) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.	300
Abbildung 61: Netzwerkdiagramm für die Analyse der SGRQ-Responder (%) nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	305

Abbildung 62: Netzwerkdiagramm für die Analyse der SGRQ-Responder (%) nach 12 Wochen (alle Studien).....	305
Abbildung 63: Netzwerkdiagramm für die Analyse der SGRQ-Responder (%) nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet).....	306
Abbildung 64: Netzwerkdiagramm für die Analyse der SGRQ-Responder (%) nach 24 Wochen (alle Studien).....	307
Abbildung 65: Meta-Analyse für den Prozentsatz der SGRQ-Responder (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	312
Abbildung 66: Meta-Analyse für den Prozentsatz der SGRQ-Responder (12 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.....	313
Abbildung 67: Meta-Analyse für den Prozentsatz der SGRQ-Responder (24 Wochen) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo.....	313
Abbildung 68: Sensitivitätsanalyse (Verzerrungspotential) für SGRQ-Responder nach 12 Wochen (ml) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	314
Abbildung 69: Netzwerk-Diagramm für die Analyse unerwünschter Ereignisse nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet).....	319
Abbildung 70: Netzwerk-Diagramm für die Analyse unerwünschter Ereignisse nach 12 Wochen.....	319
Abbildung 71: Netzwerk-Diagramm für die Analyse unerwünschter Ereignisse nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet).....	320
Abbildung 72: Netzwerk-Diagramm für die Analyse unerwünschter Ereignisse nach 24 Wochen (alle Studien).....	321
Abbildung 73: Meta-Analyse für unerwünschte Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	326
Abbildung 74: Meta-Analyse für unerwünschte Ereignisse (12 Wochen) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.....	327
Abbildung 75: Meta-Analyse für unerwünschte Ereignisse (24 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.....	327
Abbildung 76: Sensitivitätsanalyse (COPD-Schweregrad) für unerwünschte Ereignisse (12 Wochen) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.....	328
Abbildung 77: Netzwerk-Diagramm für die Analyse schwerwiegender unerwünschter Ereignisse nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet).....	332
Abbildung 78: Netzwerk-Diagramm für die Analyse schwerer unerwünschter Ereignisse nach 12 Wochen.....	333
Abbildung 79: Netzwerk-Diagramm für die Analyse schwerer unerwünschter Ereignisse nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet).....	334
Abbildung 80: Netzwerk-Diagramm für die Analyse schwerer unerwünschter Ereignisse nach 24 Wochen.....	335
Abbildung 81: Meta-Analyse für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	340

Abbildung 82: Meta-Analyse für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo	341
Abbildung 83: Meta-Analyse für schwere unerwünschte Ereignisse (24 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.....	341
Abbildung 84 : Sensitivitätsanalyse (COPD-Schweregrad) für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Tiotropiumbromid 5µg und 18µg OD im Vergleich zu Placebo.	342
Abbildung 85: Netzwerk-Diagramm für die Analyse von Ausschlüssen wegen unerwünschter Ereignisse nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	347
Abbildung 86: Netzwerk-Diagramm für die Analyse von Ausschlüssen wegen unerwünschter Ereignisse nach 12 Wochen.....	347
Abbildung 87: Netzwerk-Diagramm für die Analyse von Ausschlüssen wegen unerwünschter Ereignisse nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)	348
Abbildung 88: Netzwerk-Diagramm für die Analyse von Ausschlüssen wegen unerwünschter Ereignisse nach 24 Wochen.....	349
Abbildung 89: Meta-Analyse für Studienabbrüche aufgrund von unerwünschten Ereignissen (12 Wochen); Acclidiniumbromid im Vergleich zu Placebo.	353
Abbildung 90: Meta-Analyse für Studienabbrüche aufgrund von unerwünschten Ereignissen (12 Wochen); Tiotropiumbromid 5µg im Vergleich zu Placebo.....	354
Abbildung 91: Meta-Analyse für Studienabbrüche aufgrund von unerwünschten Ereignissen (24 Wochen); Tiotropiumbromid 18µg im Vergleich zu Placebo.....	354
Abbildung 92: Sensitivitätsanalyse (COPD-Schweregrad) für Studienabbrüche aufgrund von unerwünschten Ereignissen (12 Wochen); Tiotropiumbromid 5µg im Vergleich zu Placebo	355
Abbildung 93: Netzwerk-Diagramm für die Analyse kardiovaskulärer Morbidität und Mortalität nach 12 Wochen	360
Abbildung 94: Meta-Analyse für die Anzahl von Patienten mit kardiovaskulärer Morbidität und Mortalität (12 Wochen); Acclidiniumbromid im Vergleich zu Placebo	363
Abbildung 95: Netzwerk-Diagramm für die Analyse der Letalität und der Gesamt-Mortalität nach 12 Wochen	365
Abbildung 96: Netzwerk-Diagramm für die Analyse der Letalität und der Gesamt-Mortalität nach 24 Wochen	366
Abbildung 97: Meta-Analyse für mittelschwere COPD für CFB Trough-FEV ₁ (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo.....	373
Abbildung 98: Meta-Analyse für schwere COPD für CFB Trough-FEV ₁ (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo.....	373
Abbildung 99: Meta-Analyse für mittelschwere COPD für FEV ₁ -Peak (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo.....	376
Abbildung 100: Meta-Analyse für schwere COPD für FEV ₁ -Peak (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo.....	376

Abbildung 101: Meta-Analyse für mittelschwere COPD für den TDI Score (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo	379
Abbildung 102: Meta-Analyse für schwere COPD für den TDI Score (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	379
Abbildung 103: Meta-Analyse für mittelschwere COPD für TDI-Responder (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo	382
Abbildung 104: Meta-Analyse für schwere COPD für TDI-Responder (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	382
Abbildung 105: Meta-Analyse für mittelschwere COPD für Exazerbationsrate aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	395
Abbildung 106: Meta-Analyse für mittelschwere COPD für Exazerbationsrate aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	396
Abbildung 107: Meta-Analyse für mittelschwere COPD für den SGRQ-Wert (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	400
Abbildung 108: Meta-Analyse für schwere COPD für den SGRQ-Wert (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	400
Abbildung 109: Meta-Analyse für mittelschwere COPD für den SGRQ-Wert (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	403
Abbildung 110: Meta-Analyse für schwere COPD für den SGRQ-Wert (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	403
Abbildung 111: Meta-Analyse für mittelschwere COPD für unerwünschte Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	405
Abbildung 112: Meta-Analyse für schwere COPD für unerwünschte Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	405
Abbildung 113: Meta-Analyse für mittelschwere COPD für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.....	408
Abbildung 114: Meta-Analyse für schwere COPD für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo	408
Abbildung 115: Meta-Analyse für mittelschwere COPD für Ausschlüsse aufgrund unerwünschter Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo	411
Abbildung 116: Meta-Analyse für schwere COPD für Ausschlüsse aufgrund unerwünschter Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo	411

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AB	Acridiniumbromid
AE	Adverse Event (unerwünschtes Ereignis)
ANCOVA	Analysis of covariance (Kovarianzanalyse)
ATS	American Thoracic Society
AUC	Area under the curve (Fläche unter der Kurve)
BID, bid	bis in die (zweimal täglich)
BDI	Baseline Dyspnea Index
BMI	Body Mass Index
CFB	Change from baseline
CONSORT	Consolidated Standards of Reporting Trials
COPD	Chronic Obstructive Pulmonary Disease (Chronisch obstruktive Lungenerkrankung)
CPMP	Committee for Proprietary Medicinal Products
CRF	Case Report Form (Prüfbogen)
CSR	Clinical Study Report
DIC	Deviance Information Criterion
DIMDI	Deutsches Institut für Medizinische Dokumentation
DMP	Disease Management Programm
EG	Europäische Gemeinschaft
EMA	European Medicines Agency
ERS	European Respiratory Society
E-RS Score	EXACT- Respiratory Symptoms (E-RS) Fragebogen - Score
EXACT	Exacerbations of chronic pulmonary disease tool
EXACT-PRO	Exacerbations of chronic pulmonary disease tool – patient reported outcome
EXACT-RS	Exacerbations of chronic pulmonary disease tool for respiratory symptoms
FDA	Food and Drug Administration
FEV ₁	Forced Expiratory Volume (Forciertes expiratorisches Ein-Sekunden-Volumen)
FVC	forced vital capacity (Forcierte Vitalkapazität)
G-BA	Gemeinsamer Bundesausschuss

GOLD	Global Initiative for Chronic Obstructive Lung Disease
ICS	Inhalative Corticosteroide
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
IV	intravenös
KI	Konfidenzintervall
l	Liter
LABA	Long-acting beta2 agonist (langwirksame Beta-2-Sympathomimetika)
LAMA	Long-acting muscarinic antagonists (langwirksame Anticholinergika)
LS	Least squares (Quadrat-Mittelwerte)
MDI	Metered-dose inhaler (Treibgasdosieraerosole)
µg	Mikrogramm
ml	Milliliter
MTC	Mixed Treatment Comparison
n	Anzahl
n.d.	nicht durchgeführt
NMA	Netzwerk-Metaanalyse
NR	Not reported
NVL	Nationale VersorgungsLeitlinie
OR	Odds Ratio
p	probability (Wahrscheinlichkeit)
Pbo	Placebo
QAM, qad	quaque ante meridiem (einmal morgens)
QD, qd	quaque die (einmal täglich)
QPM, qpm	quaque post meridiem (einmal abends)
RCT	Randomized Controlled Trial (randomisierte kontrollierte Studien)
SABA	Short-acting beta2 agonists (kurzwirksame Beta-2-Sympathomimetika)
SAE	Serious Adverse Event (schwerwiegendes unerwünschtes Ereignis)
SAMA	Short-acting muscarinic antagonists (kurzwirksame Anticholinergika)
SD	standard deviation (Standardabweichung)
SE	standard error (Standardfehler)

SGB	Sozialgesetzbuch
SGRQ	St. George's Respiratory Questionnaire
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TB	Tiotropiumbromid
TDI	Transition Dyspnea Index
Tio	Tiotropiumbromid
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
VK	Vitalkapazität
WHO	World Health Organization

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Für Arzneimittel, die zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt. Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, müssen jedoch vorgelegt werden. Zu diesem Zweck enthält die Dokumentvorlage den Abschnitt 4.4.4, der ausschließlich in Dossiers für die oben genannten Arzneimittel zur Behandlung eines seltenen Leidens bearbeitet werden soll. In diesem Abschnitt sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich der begründenden Nachweise vorzulegen. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

4.1 Zusammenfassung der Inhalte von Modul 4

Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.

Fragestellung

Welchen medizinischen Nutzen und Zusatznutzen weist die Behandlung von erwachsenen Patienten mit chronisch obstruktiver Lungenerkrankung (Stufe II-IV nach Nationaler VersorgungsLeitlinie) mit Acclidiniumbromid im Vergleich zu Tiotropiumbromid auf?

Die Bewertung erfolgt anhand randomisierter kontrollierter Studien (RCTs). Es werden patientenrelevante Endpunkte zur Beurteilung herangezogen, diese sind in Abschnitt 4.2.2 detailliert dargestellt.

Datenquellen

Für die Bewertung des Nutzen und Zusatznutzens von Acclidiniumbromid wurden alle in Bezug auf patientenrelevante Endpunkte aussagekräftigen RCTs des pharmazeutischen Herstellers sowie placebokontrollierte Studien der zweckmäßigen Vergleichstherapie berücksichtigt, die in den Studienregistern und in der systematischen Literaturrecherche identifizierbar waren.

Ein-/Ausschlusskriterien für Studien

Eingeschlossen wurden RCTs, die die Wirksamkeit von Acclidiniumbromid und/ oder Tiotropiumbromid bei Erwachsenen mit chronisch obstruktiver Lungenerkrankung (COPD) untersuchten. Diese RCTs mussten Ergebnisse zu patientenrelevanten Endpunkten beschreiben und außerdem als Volltextpublikation zur Verfügung stehen.

Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen

Die Verzerrungsaspekte der Ergebnisse wurden für jede einzelne der eingeschlossenen Studien beschrieben und bewertet. Die Bewertung der einzelnen Aspekte erfolgte getrennt auf Ebene der einzelnen RCT sowie für jeden Endpunkt. Die Ergebnisse wurden in den im Dossier vorgegebenen Bewertungsbögen dokumentiert.

Im vorliegenden Dossier wird der medizinische Nutzen und Zusatznutzen sowohl anhand direkter als auch anhand indirekter Evidenz nachgewiesen.

Die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus den identifizierten RCTs für einen direkten Vergleich wurden für Acclidiniumbromid im Vergleich zur zweckmäßigen Vergleichstherapie (Tiotropiumbromid) für jeden der patientenrelevanten Endpunkte tabellarisch gegenübergestellt, analysiert und zusammenfassend beschrieben. Meta-Analysen konnten aufgrund der Heterogenität des Studiendesigns nicht durchgeführt werden. Abschließend wurde zur Beschreibung des Zusatznutzens sowie für dessen Ausmaß und Wahrscheinlichkeit eine Zusammenfassung und Bewertung der vorliegenden Ergebnisse als solche vorgenommen, unter Berücksichtigung der Validität der herangezogenen Endpunkte, der Studienqualität sowie der Evidenzstufe der vorliegenden Informationen.

Mit den für den indirekten Vergleich identifizierten RCTs (Aclidiniumbromid bzw. Tiotropiumbromid vs. Placebo) wurde bei der Gegenüberstellung der Ergebnisse genauso verfahren. Die Zusammenfassung der Ergebnisse zu einem Gesamteffektschätzer erfolgte für die jeweilige Therapie mit Hilfe einer Meta-Analyse. Anschließend wurden die Ergebnisse der placebokontrollierten RCTs in einem indirekten Vergleich mit Hilfe eines bayesianischen Netzwerk-Meta-Analysemodells (NMA) ausgewertet.

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Direkter Vergleich

Der Nutzen von Aclidiniumbromid und der Zusatznutzen gegenüber Tiotropiumbromid bzgl. des patientenrelevanten Endpunkts Morbidität lässt sich mit Hilfe der als primäres Therapieziel definierten Linderung der COPD-Symptome zeigen. Dazu gehören die drei COPD-Hauptsymptome Husten, Auswurf und Atemnot, die nächtlichen- und morgendlichen Symptome, sowie die Veränderung der forcierten expiratorischen Volumina in einer Sekunde (FEV₁).

In Bezug auf die Linderung der Hauptsymptome und der nächtlichen- und morgendlichen Symptome sind die Werte numerisch zugunsten von Aclidiniumbromid. Die Einschränkung der körperlichen Aktivität durch COPD-Symptome wird von Aclidiniumbromid-Patienten signifikant (-0,10, KI [-0,19; -0,01] p=0,04) geringer als von Tiotropiumbromid-Patienten empfunden.

Aclidiniumbromid ermöglicht durch die zweimal tägliche Gabe eine konstante Bronchodilatation über 24 Stunden (FEV₁ AUC) und wirkt effektiv auf die besonders belastenden COPD-Symptome in der Nacht und am Morgen. Das Bronchodilatationsprofil ist dem Profil von Tiotropiumbromid mindestens gleichwertig.

Die nachhaltige Verbesserung der Lungenfunktion wird durch überwiegend klinisch relevante Veränderungen zur Baseline und die im Vergleich zu Tiotropiumbromid mindestens gleichwertige Trough-FEV₁-Werte gezeigt.

Die Effektivität von Aclidiniumbromid wird durch das Erreichen hoher Peak-FEV₁-Werte bereits am ersten Behandlungstag deutlich.

Der mit der Behandlungszufriedenheit beschriebene Lebensqualitätsaspekt zeigt einen signifikanten Vorteil für den Genuair[®] Inhalator gegenüber dem HandiHaler[®] (Tiotropiumbromid): Bis zu 80% der Patienten bevorzugen die Handhabung des Genuair[®] - Inhalationssystems. Dieser bedeutende Vorteil wird ergänzend zur Nutzenbewertung herangezogen.

Bei den Nebenwirkungen, Studienabbrüchen aufgrund von Nebenwirkungen sowie für die Aspekte der Mortalität unterscheiden sich die Therapien nicht voneinander. Bei Aclidiniumbromid sind im Gegensatz zu Tiotropiumbromid keine Vorsichtsmaßnahmen und Dosisanpassungen bei Patienten mit mittlerer bis schwerer Nierenfunktionsstörung nötig. Dies stellt einen Vorteil bezüglich der Therapiebrotweite dar.

Indirekter Vergleich

Der indirekte Vergleich von Acclidiniumbromid mit der zweckmäßigen Vergleichstherapie mit Placebo als Brückenkomparator wurde für die Endpunkte Trough-FEV₁, SGRQ (Lebensqualität), SGRQ-Responder, TDI focal Score (Dyspnoe), TDI-Responder, Peak-FEV₁, (schwere) unerwünschte Ereignisse und Studienabbrüche wegen unerwünschter Ereignisse nach 12 und 24 Wochen durchgeführt. Außerdem wurden die Exazerbationsraten und die aufgrund von Exazerbationen beobachteten Hospitalisierungen nach 24 Wochen indirekt verglichen.

Die Ergebnisse sprechen für eine weitgehende Vergleichbarkeit der beiden Therapien. In allen indirekten Vergleichen umfassten die 95 %-Konfidenzintervalle die Null (Endpunkte mit binärem oder kontinuierlichem Messniveau) bzw. die Eins (Raten). Beim Endpunkt Hospitalisierungen aufgrund von Exazerbationen nach 24 Wochen betrug der obere Grenzwert des 95 %-Konfidenzintervalls eins, was auf eine starke Tendenz zugunsten von Acclidiniumbromid hinweist. Die Ergebnisse des indirekten Vergleichs belegen den zumindest vergleichbaren patientenrelevanten Nutzen von Acclidiniumbromid und Tiotropiumbromid bei COPD.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Anhand der im Dossier analysierten patientenrelevanten Endpunkte lassen sich die Linderung von COPD-Symptomen, die Verringerung der Anzahl von Exazerbationen und Hospitalisierungen, eine bedeutende Verbesserung der Lebensqualität sowie, auch als notwendige Grundbedingung der übrigen Wirkungen, eine große Behandlungssicherheit und -zufriedenheit mit dem Genuair[®]-Inhalator nachweisen.

Nach §5 (7) Nr. 2 des 5. Kapitels der Verfahrensordnung wurde gegenüber der zweckmäßigen Vergleichstherapie eine moderate und nicht nur eine geringfügige Verbesserung des therapielevanten Nutzens erreicht. Dies betrifft vor allen Dingen die Verringerung nicht schwerwiegender Symptome bei einem gleichwertigen Nebenwirkungs- und Sicherheitsprofil. Die nicht erforderlichen Dosisanpassungen oder das Ergreifen von Vorsichtsmaßnahmen bei Patienten mit mittlerer bis schwerer Nierenfunktionsstörung bergen gegenüber Tiotropiumbromid einen klaren Vorteil bezüglich der Therapiebreite.

In dem vorliegenden Dossier wird der Zusatznutzen belegt und aufgrund der gegenüber der zweckmäßigen Vergleichstherapie erzielten moderaten und nicht nur geringfügigen Verbesserung des therapielevanten Nutzens als „gering“ im Sinne der Verfahrensordnung (VerfO) eingestuft. Die im Dossier als Resultat der systematischen Übersicht dargelegten Nachweise zum Zusatznutzen von Acclidiniumbromid berücksichtigen Ergebnisse von zwei RCTs mit direktem Vergleich sowie 24 RCTs für den indirekten Vergleich und entsprechen so den Anforderungen des Evidenzlevels der Stufe 1.

Basierend auf der Ergebnisvollständigkeit der dargestellten Endpunkte kann von einer hohen Sicherheit des dargestellten Zusatznutzens von Acclidiniumbromid ausgegangen werden. Im direkten Vergleich kann der Zusatznutzen von Acclidiniumbromid gegenüber

Tiotropiumbromid anhand der signifikant besseren Werte an Tag 1 für CFB Peak-FEV1, CFB FEV1 AUC₁₂₋₂₄, und durch eine deutliche Patientenpräferenz für das Inhalationssystem in zwei RCTs gezeigt werden. In jeweils einem RCT wird signifikant belegt, dass die Einschränkung der körperlichen Aktivität durch die COPD-Symptome nach Behandlung mit Acclidiniumbromid geringer ist, die Lungenfunktion gemessen als CFB FEV1 AUC₁₂₋₂₄, an Tag 15 und der Wert für CFB Trough FEV1 an Tag 1 signifikant besser als für Tiotropiumbromid ist.

Die weiteren gleichgerichteten, konsistenten Ergebnisse für die COPD-Symptome und die FEV1-Werte in beiden RCTs untermauern die bessere Wirksamkeit von Acclidiniumbromid gegenüber Tiotropiumbromid. Dies entspricht der Aussagewahrscheinlichkeit eines Beleges.

Der indirekte Vergleich mit insgesamt 24 Studien zeigt nicht nur eine Nichtunterlegenheit für Acclidiniumbromid, sondern deutet mindestens auf eine Gleichwertigkeit der beiden Wirkstoffe hin. In einigen Endpunkten - insbesondere bei längerer Studiendauer von 24 Wochen - ist eine numerisch positive Tendenz für Acclidiniumbromid erkennbar.

Der Zusatznutzen von Acclidiniumbromid gegenüber der zweckmäßigen Vergleichstherapie kann im Dossier für die Behandlung aller Patienten der Zielpopulation (COPD-Stufen II–IV) nachgewiesen werden. Die Ergebnisse der durchgeführten Subgruppenanalyse haben keinen Einfluss auf die Nutzenbewertung. Durch die moderate und nicht nur geringfügige Verbesserung des therapielevanten Nutzens wird ein Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie erreicht, der im Sinne der VerfO als „gering“ eingestuft wird.

4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

In vorliegendem Dossier soll folgende Frage beantwortet werden:

Welchen medizinischen Nutzen und Zusatznutzen weist die Behandlung von erwachsenen Patienten mit chronisch obstruktiver Lungenerkrankung (Stufe II-IV nach Nationaler VersorgungsLeitlinie) mit Acclidiniumbromid im Vergleich zu Tiotropiumbromid auf?

Die Bewertung des medizinischen Nutzen und Zusatznutzen erfolgt anhand randomisierter kontrollierter Studien (RCTs). Es werden patientenrelevante Endpunkte zur Beurteilung herangezogen, diese sind in Abschnitt 4.2.2 detailliert dargestellt.

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.

Im vorliegenden Dossier soll der medizinische Nutzen und Zusatznutzen sowohl anhand direkter als auch anhand indirekter Evidenz nachgewiesen werden. Die Suche nach direkt vergleichenden Studien wird getrennt von der Suche nach Studien für einen indirekten Vergleich durchgeführt. Die Ein- und Ausschlusskriterien der beiden Herangehensweisen der systematischen Auswahl von Studien sind im Folgenden erläutert.

Population

Die Studienpopulation ist in beiden Fragestellungen identisch und ist wie folgt definiert: Erwachsene Patienten mit einer nach nationalen und/oder internationalen Leitlinien diagnostizierten COPD der Stufe II bis IV¹ [1]. Die Definition und Diagnosekriterien der Erkrankung sind in Modul 3 dargestellt.

Intervention

Für die Suche nach direkt vergleichenden Studien stellt Acridiniumbromid in der zugelassenen Dosierung (Eklira[®] Genuair[®]/ Bretaris[®] Genuair[®] 322 µg² 2 mal täglich) die Intervention dar.

In einem indirekten Vergleich sollen Studien des zu bewertenden Arzneimittels und der Vergleichstherapie gegenüber gestellt werden. Dem zu Folge werden bei der Suche nach Studien für einen indirekten Vergleich RCTs eingeschlossen, die entweder Acridiniumbromid oder Tiotropiumbromid als Intervention untersuchen. Tiotropiumbromid wird als Spiriva[®] 18 Mikrogramm und als Spiriva[®] Respimat[®] vertrieben. Beide Präparate werden in die Nutzenbewertung einbezogen.

Vergleichstherapie

Für die Suche nach direkt vergleichenden Studien stellt Tiotropiumbromid (Spiriva[®] 18 Mikrogramm und Spiriva[®] Respimat[®]) in der zugelassenen Dosierung die Vergleichstherapie dar.

Die in der Recherche nach Studien für einen indirekten Vergleich zu berücksichtigende Vergleichstherapie wird maßgeblich von dem Brückenkomparator beeinflusst. Der Brückenkomparator ermöglicht den indirekten Vergleich des zu bewertenden und des zu vergleichenden Arzneimittels.

Bei den zur Verfügung stehenden RCTs des pharmazeutischen Unternehmers mit der zulassungskonformen Anwendung von Acridiniumbromid (Monotherapie mit 400µg 2 mal

¹ Diese Einschränkung beruht auf der Nationalen VersorgungsLeitlinie, die erst ab der COPD-Stufe II langwirksame Bronchodilatoren empfiehlt.

² Die in der gesamten außerklinischen und klinischen Entwicklung von Acridiniumbromid verwendete Bezeichnung lautet Eklira[®] Genuair[®]/ Bretaris[®] Genuair[®] 400 µg. Dies entspricht der gemessenen Dosis (Dosiervolumen) des Wirkstoffs (Acridiniumbromid) und nicht dem aktiven Teil (Acridinium). Da die Leitlinien der EMA vorgeben, dass sich die Stärkeangabe in der SmPC auf die verabreichte Dosis des aktiven Teils (Acridinium) beziehen sollte, wurde diese Bezeichnung im Laufe des Zulassungsverfahrens auf Eklira[®] Genuair[®]/ Bretaris[®] Genuair[®] 322 µg geändert. Es handelt sich bei beiden Bezeichnungen um das gleiche Arzneimittel – nur zu unterschiedlichen Zeitpunkten des Entwicklungs- und Zulassungsprozesses.

täglich) und einer hinreichenden Studiendauer (vgl. Einschlusskriterium „Studiendauer“) stehen lediglich placebokontrollierte Studien zur Verfügung (vgl. Abschnitt 4.3.1.1.1). Als Brückenkompator wird somit Placebo gewählt. Der indirekte Vergleich erfolgt dementsprechend auf Grundlage von placebokontrollierten Studien; Placebo wird als Vergleichstherapie für die Suche nach Studien für den indirekten Vergleich definiert.

Um den Studienpool für den indirekten Vergleich unabhängig vom Studienpool des direkten Vergleichs zu vervollständigen, werden auch Studien mit einem direkten Vergleich von Acclidiniumbromid und Tiotropiumbromid gesucht.

Studiendesign

Die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention liefern randomisierte kontrollierte Studien (RCTs) [2]. Das vorliegende Dossier verwendet ausschließlich methodisch adäquate und der jeweiligen Fragestellung angemessene RCTs. Dieses Vorgehen ermöglicht Aussagen zum medizinischen Nutzen und Zusatznutzen mit hoher Ergebnissicherheit.

Patientenrelevante Endpunkte

In der Niederschrift vom 21.06.2012 zum Beratungsgespräch [3] am 24.05.2012 werden folgende patientenrelevante Endpunkte genannt:

- „Mortalität, insbesondere COPD-bedingte Letalität
- Morbidität, insbesondere COPD-Symptomatik, wie Anzahl und Schwere der COPD-Symptome, Dyspnoe und Exazerbationen
- Gesundheitsbezogene Lebensqualität
- Unerwünschte Ereignisse“

Lungenfunktionsparameter gelten in der Versorgungspraxis als allgemein anerkannter patientenrelevanter Endpunkte und werden ergänzend zur Nutzenbewertung herangezogen.

Ergänzend kann laut Methodenpapier des IQWiG [2] (S. 31) „die Zufriedenheit der Patientinnen und Patienten mit der Behandlung berücksichtigt werden.“ In der COPD-Therapie handelt es sich dabei um die Zufriedenheit mit dem verwendeten Inhalationssystem und um die Handhabbarkeit desselben. Dieser Endpunkt wird der Endpunktkategorie Lebensqualität zugeordnet.

Die Ausprägungen der vom G-BA genannten und im vorliegenden Dossier eingeschlossenen Endpunkte sind in Tabelle 4.2-1 dargestellt. Studien müssen mindestens einen dieser Endpunkte analysiert haben um dieses Einschlusskriterium zu erfüllen.

Tabelle 4.2-1: Patientenrelevante Endpunkte

Endpunktkategorie	Ausprägungen der Endpunktkategorie
Mortalität	<ul style="list-style-type: none"> • COPD-assoziierte kardiovaskuläre Mortalität • COPD-bedingte Letalität • Gesamtmortalität
Morbidität	<ul style="list-style-type: none"> • Husten, Auswurf, Atemnot • Nächtliche und morgendliche Symptome • Dyspnoe • Exazerbationen • Veränderung der forcierten expiratorischen Volumens in einer Sekunde (FEV₁)
Gesundheitsbezogene Lebensqualität	<ul style="list-style-type: none"> • St. George's Respiratory Questionnaire (SGRQ) • Behandlungszufriedenheit
Unerwünschte Arzneimittelwirkungen	<ul style="list-style-type: none"> • unerwünschte Ereignisse (AEs) • schwerwiegende unerwünschte Ereignisse (SAEs) • Studienabbruch wegen unerwünschter Ereignisse

Aufgrund der Erhebungszeitpunkte der Studien des pU nach 12 und 24 Wochen gelten die Erhebungszeitpunkte 12 bzw. 24 Wochen³ nach Studienbeginn als allein gültige Erhebungszeitpunkte.⁴

Studiendauer

Für Effektivitätsnachweise in der Indikation COPD empfiehlt die EMA eine Studiendauer von 12-24 Wochen [4]. Werden in der Suche nach Studien für den direkten oder indirekten Vergleich keine Studien mit der empfohlenen Studiendauer gefunden, wird aufgrund der auch kurzfristig möglichen Behandlungseffekte von der empfohlenen Mindestdauer abgesehen. Bei diesen bereits sehr kurzfristig einsetzenden Behandlungseffekten ist insbesondere die mit der Bronchodilatation erreichte Verbesserung der Lungenfunktion relevant. Die Verbesserung der Lungenfunktion hat außerdem einen positiven Einfluss auf die Atemnot (Dyspnoe) und weitere kurzfristig relevante COPD-Symptome (Husten und Auswurf).

Für die Bewertung des Zusatznutzens werden in diesem Fall insbesondere solche Endpunkte herangezogen, die den schnell eintretenden Erfolg der Therapie sichtbar machen. Dies sind: COPD-Symptome, nächtliche und morgendliche Symptome und Lungenfunktionsparameter.

³ Eine Varianz von 2 Wochen, d. h. 10-14 bzw. 22-26 Wochen, sind laut den Allgemeinen Methoden des IQWiG zulässig.

⁴ Unter der Voraussetzung, dass keine Studien Dritter mit Acclidiniumbromid mit abweichenden Erhebungszeitpunkten existieren.

Alle übrigen in den Studien berichteten patientenrelevanten Endpunkte werden zum Zwecke der Vollständigkeit ebenfalls dargestellt.

Werden in der Suche nach Studien für den direkten oder indirekten Vergleich hingegen Studien mit der empfohlenen Studiendauer gefunden, werden Studien mit einer kürzeren als der empfohlenen ausgeschlossen. So wird ein möglichst homogener Studienpool mit möglichst langer Beobachtungsdauer für weiterführende Analysen (Meta-Analyse, indirekter Vergleich) gewährleistet.

Aus den oben beschriebenen Ein- und Ausschlusskriterien ergibt sich folgende Übersicht:

Tabelle 4.2-2: Einschlusskriterien – Suche nach direkt vergleichenden RCTs

Kurzbezeichnung	Einschlusskriterium
E1	Patienten mit COPD (Stufe II-IV)
E2	Intervention: Aclidiniumbromid, Dosierung laut Zulassung von Eklira [®] / Bretaris [®] Genuair [®] (Aclidiniumbromid 322 µg 2 mal täglich)
E3	Vergleichstherapie: Tiotropiumbromid, Dosierung laut Zulassung von Spiriva [®] 18 Mikrogramm (18 µg 1 mal täglich) und Spiriva [®] Respimat [®] (5 µg 1 mal täglich)
E4	Studien im RCT-Design
E5	Erhebung patientenrelevanter Endpunkte
E6	Studiendauer: wenn vorhanden \geq 12 Wochen , sonst auch < 12 Wochen

Tabelle 4.2-3: Ausschlusskriterien – Suche nach direkt vergleichenden RCTs

Kurzbezeichnung	Ausschlusskriterium
A1	Patienten, die nicht der definierten Patientenpopulation entsprechen
A2	Intervention nicht wie definiert
A3	Vergleichstherapie nicht wie definiert
A4	Studie nicht im RCT-Design
A5	Keine patientenrelevanten Endpunkte
A6	Studien < 12 Wochen, wenn Studien \geq 12 Wochen vorhanden
A7	Keine Volltextpublikation oder Studienberichte verfügbar
A8	Mehrfachpublikationen ohne (für die Bewertung des Zusatznutzens) relevante Zusatzinformation

A9	Publikation in einer anderen als der englischen oder deutschen Sprache
A10	Studien mit einem anderen Studienziel als dem des Wirksamkeitsvergleiches der Intervention und Vergleichstherapie

Tabelle 4.2-4: Einschlusskriterien – Suche nach RCTs für den indirekten Vergleich

Kurzbezeichnung	Einschlusskriterium
E1	Patienten mit COPD (Stufe II-IV)
E2	Intervention: Aclidiniumbromid, Dosierung laut Zulassung von Eklira [®] / Bretaris [®] Genuair [®] (Aclidiniumbromid 322 µg 2 mal täglich) oder Tiotropiumbromid, Dosierung laut Zulassung von Spiriva [®] 18 Mikrogramm (18 µg 1 mal täglich) und Spiriva [®] Respimat [®] (5 µg 1 mal täglich)
E3	Vergleichstherapie: Placebo oder Aclidiniumbromid, Dosierung laut Zulassung von Eklira [®] / Bretaris [®] Genuair [®] (Aclidiniumbromid 322 µg 2 mal täglich) oder Tiotropiumbromid, Dosierung laut Zulassung von Spiriva [®] 18 Mikrogramm (18 µg 1 mal täglich) und Spiriva [®] Respimat [®] (5 µg 1 mal täglich)
E4	Studien im RCT-Design
E5	Erhebung patientenrelevanter Endpunkte zu einem vergleichbaren Zeitpunkt
E6	Studiendauer: wenn vorhanden ≥ 12 Wochen, sonst auch < 12 Wochen

Tabelle 4.2-5: Ausschlusskriterien – Suche nach RCTs für den indirekten Vergleich

Kurzbezeichnung	Ausschlusskriterium
A1	Patienten, die nicht der definierten Patientenpopulation entsprechen
A2	Intervention nicht wie definiert
A3	Vergleichstherapie nicht wie definiert
A4	Studie nicht im RCT-Design
A5	Keine patientenrelevanten Endpunkte oder keine patientenrelevanten Endpunkte zu einem vergleichbaren Zeitpunkt
A6	Studien < 12 Wochen, wenn Studien \geq 12 Wochen vorhanden
A7	Keine Volltextpublikation oder Studienberichte verfügbar
A8	Mehrfachpublikationen ohne (für die Bewertung des Zusatznutzens) relevante Zusatzinformation
A9	Publikation in einer anderen als der englischen oder deutschen Sprache
A10	Studien mit einem anderen Studienziel als dem des Wirksamkeitsvergleiches der Intervention und Vergleichstherapie

Es werden Studien eingeschlossen, die alle Einschlusskriterien aber keines der Ausschlusskriterien erfüllen. Die Einschlusskriterien gelten als erfüllt, wenn in der zu bewertenden Studie mindestens 80 % der Patienten die Einschlusskriterien erfüllen [2].

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2.1, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“.

Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist in folgenden Fällen erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt:

- Für das zu bewertende Arzneimittel wurde eine Zulassung in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wurde, mindestens 12 Monate vor dem Zeitpunkt der Dossiereinreichung erteilt. Dabei sind auch Zulassungen außerhalb Deutschlands bzw. Europas relevant.
- Es werden indirekte Vergleiche zur Bewertung des medizinischen Zusatznutzens herangezogen.

Sofern eine bibliografische Literaturrecherche erforderlich ist, soll diese mindestens in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Geben Sie dabei an, ob Sie eine gemeinsame Suche für mehrere Fragestellungen durchgeführt haben (z. B. gemeinsame Suche nach Studien mit dem zu bewertenden Arzneimittel und Studien zur zweckmäßigen Vergleichstherapie für indirekte Vergleiche). Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, ob bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Die Zulassung des zu bewertenden Arzneimittels im Anwendungsgebiet wurde weniger als 12 Monate vor der Einreichung des Dossiers erteilt. Zur Sicherstellung eines vollständigen Studienpools und aufgrund der Durchführung eines indirekten Vergleichs wurde eine vollständige bibliografische Literaturrecherche durchgeführt.

Die bibliografische Literaturrecherche wurde in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt.

Nach RCTs mit einem direkten Vergleich des zu bewerteten Arzneimittels mit der zweckmäßigen Vergleichstherapie wurden die Datenbanken MEDLINE und EMBASE sowie der Cochrane-Datenbanken systematisch durchsucht. Die Suche nach Studien für den indirekten Vergleich wurde in MEDLINE, MEDLINE In Process und EMBASE mit dem OVID-System durchgeführt. Außerdem wurde in den Cochrane-Datenbanken recherchiert.

Die Suchstrategien beider Recherchen in jeweils drei Datenbanken sind mitsamt der verwendeten Suchbegriffe im Anhang 4-A detailliert beschrieben. Die Suche nach Studien für den indirekten Vergleich wurde am 5. Juli 2012 durchgeführt. Die Suche nach Studien für den direkten Vergleich wurde am 28. August 2012 letztmalig aktualisiert. Die Suchstrategien wurden unter Verwendung eines Pools von verwandten Suchbegriffen in Bezug auf die Einschlusskriterien der jeweiligen Recherche für jede Datenbank angepasst. Die Suchstrategien beinhalten jeweils eine Kombination von Freitext und Begriffen der Medical Subject Headings (MeSH).

Bei der Recherche nach RCTs mit einem direkten Vergleich wurden aufgrund der engen Suchkriterien keine generellen Einschränkungen vorgenommen. Bei der Recherche nach RCTs für einen indirekten Vergleich war die Suche auf deutsch- und englischsprachige Publikationen beschränkt. Studien, die nur Tiere in ihrer Nomenklatur hatten, wurden ausgeschlossen.

4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien von Dritten vollständig identifiziert werden.

Die Suche soll mindestens in den Studienregistern clinicaltrials.gov, clinicalstudyresults.org sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden.

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Eine gemeinsame Suche nach Studien zu mehreren Fragestellungen (z. B. direkt vergleichende Studien sowie Studien für einen indirekten Vergleich) ist möglich. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Beschreiben Sie nachfolgend, in welchen Studienregistern die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, ob bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.

Zusätzlich zur systematischen Suche in medizinischen Datenbanken wurde auch eine Suche in öffentlich zugänglichen Registern durchgeführt. Das Ziel dieser Suche war die Identifikation aller klinischen Studien zu Tiotropiumbromid und Aclidiniumbromid bei COPD. Die Suchen in den Registern galten der Identifizierung aller abgeschlossenen, abgebrochenen und laufenden Studien.

Die Suche in Studienregistern wurde in dem Studienregister clinicaltrials.gov und über das International Clinical Trials Registry Plattform Search Portal (ICTRP Search Portal, Suchportal der WHO) durchgeführt. Die Suche im Studienregister clinicalstudyresults.org war aufgrund fehlenden Zugriffs nicht möglich.

Im ersten Schritt wurden Studien für den direkten Vergleich des zu bewerteten Arzneimittels mit der zweckmäßigen Vergleichstherapie gesucht (letzte Aktualisierung der Suche am: 29. August 2012). Im zweiten Schritt wurden am 6. Juli 2012 Studien mit dem zu bewertendem Arzneimittel und/ oder der Vergleichstherapie für einen indirekten Vergleich gesucht. Bei beiden Recherchen wurden keine generellen Einschränkungen vorgenommen.

Suchstrategien wurden für alle Register und beide Rechercheziele separat entwickelt. Die Suchbegriffe waren u. a. Bezeichnungen für die Interventionen (Tiotropiumbromid und Aclidiniumbromid) und COPD. Die in den Studien verwendete Sprache wurde nicht eingegrenzt. Anhang 4-B gibt einen Überblick über die verwendeten Suchstrategien sowie die Anzahl der Suchergebnisse.

4.2.3.4 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2 und 4.2.3.3 beschriebenen Rechenschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Auswahl der geeigneten Studien aus den in medizinischen Datenbanken und Studienregistern identifizierten Studien fand nach der Entfernung von Dubletten anhand der festgelegten Ein- und Ausschlusskriterien in zwei Schritten statt und wurde von zwei voneinander unabhängigen Personen vollzogen. Dazu wurde im ersten Schritt ein Titel-/Abstractsreening und im zweiten Schritt ein Volltextscreening durchgeführt. Die Ausschlussgründe wurden in beiden Schritten dokumentiert. Abweichende Beurteilungen wurden im abschließenden Konsensverfahren durch beide Personen geklärt und entsprechend dokumentiert.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-G):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie des Behandlers
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen.

Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Validität der in diesem Dossier präsentierten Evidenz wurde in Übereinstimmung mit den oben beschriebenen Anforderungen und den Kriterien in Anhang 4-G bewertet. Dazu wurden die Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben. Die Bewertung der einzelnen Aspekte wurde für jeden Endpunkt im Bewertungsbogen entsprechend Anhang 4-G dokumentiert.

Die Studien wurden einzeln evaluiert, um potentielle Ursachen von Verzerrungen auf Studienebene und der Ebene der einzelnen bewerteten Endpunkte zu finden. Die auf Studienebene berücksichtigten Faktoren waren u. a.: Erstellung einer Randomisierungsreihenfolge, Geheimhaltung der Gruppenzuweisung, Verblindung von Patient und Mediziner, Vergleichbarkeit der Eigenschaften der Studienarme bei Studienbeginn, ergebnisorientierte oder selektive Berichterstattung und Berücksichtigung von weiteren für die Prognose relevanten Faktoren. Die auf Endpunktebene berücksichtigten Faktoren waren u. a.: Verblindung der Endpunkterheber, Umsetzung des ITT-Prinzips, ergebnisorientierte oder selektive Berichterstattung und andere Faktoren, die zur Verzerrung der Ergebnisse beitragen könnten.

Das Verzerrungspotential wurde entweder als niedrig oder hoch eingestuft. Wenn eine Verzerrung der relevanten Ergebnisse mit hoher Wahrscheinlichkeit ausgeschlossen werden konnte, wurde dies als niedriges Verzerrungspotential angesehen. Wenn von der Verzerrungsquelle eine relevante Änderung der Ergebnisse zu erwarten war, wurde dies als hohes Verzerrungspotential angesehen. Dieser Fall musste auch angenommen werden, wenn keine ausreichenden Informationen zur Bewertung des Verzerrungspotentials vorlagen (z. B. bei der ausschließlichen Verfügbarkeit von Publikationen). Lagen hingegen zusätzlich Unterlagen – z.B. Informationen von Zulassungsbehörden – vor, wurden diese in die Bewertung des Verzerrungspotentials einbezogen.

Die Verzerrungspotentialbewertung wurde zur Bewertung der Aussagekraft der Nachweise verwendet und führte nicht zum Ausschluss von Daten.

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2 jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)⁵. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-⁶ bzw. STROBE-Statements⁷ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

In Modul 4 werden zur Nutzenbewertung ausschließlich randomisierte klinische Studien berücksichtigt. Die Darstellung der Studien entsprechen dabei den in diesem Dossier zitierten CONSORT Anforderungen Items 2b bis 14 sowie dem CONSORT-Flow-Chart.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 4.3.2 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

Benennen Sie die Patientencharakteristika und Endpunkte, zu denen Ergebnisse dargestellt werden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

⁵ Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.

⁶ Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.

⁷ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

Im vorliegenden Dossier wird der Zusatznutzen von Acridiniumbromid gegenüber Tiotropiumbromid untersucht. Die in den Studien betrachtete Patientenpopulation wird mit folgenden Patientenmerkmalen beschrieben:

Alter, Geschlecht, Schweregrad (mittel/schwer), Raucherstatus (Packungsjahre), Zeit seit der COPD-Diagnose, FEV₁ (prozentual vom Sollwert und prozentual vom FVC) zu Studienbeginn, Studienabbruchrate.

Die Ergebnisse der Studien mit einem direkten Vergleich werden in Abschnitt 4.3.1 dargestellt. Zur Gegenüberstellung der Ergebnisse der Einzelstudien werden die berichteten patientenrelevanten Endpunkte tabellarisch dargestellt und zusammenfassend beschrieben.

Die Ergebnisse der Studien für den indirekten Vergleich werden in Abschnitt 4.3.2.1 dargestellt. Auch hier werden zur Gegenüberstellung der Ergebnisse der Einzelstudien die berichteten patientenrelevanten Endpunkten tabellarisch dargestellt und zusammenfassend beschrieben.

Nach nationalen [1] und internationalen [5] Leitlinien sollte sich die Behandlung der COPD auf die Linderung der Beschwerden (durch Linderung der COPD-Symptome, Steigerung der körperlichen Belastbarkeit, Verbesserung des Gesundheitsstatus und der Lebensqualität), der Vermeidung zukünftiger Risiken (durch Vorbeugen und Behandeln von Exazerbationen und Komplikationen, sowie Verringerung der Mortalität) konzentrieren. Um auf die Deckung des in Modul 3 ausführlich beschriebenen Therapiebedarfs bei der Beurteilung des Nutzens und Zusatznutzens zu fokussieren, werden bevorzugt solche Endpunkte ausgewertet, die sich auf die oben beschriebenen Therapieziele beziehen.

Die im Dossier berichteten patientenrelevanten Endpunkte werden in der folgenden Tabelle 4.2-6 dargestellt. Es ist zu beachten, dass nicht alle Endpunkte von allen Studien berichtet werden.

Tabelle 4.2-6: Im Dossier berichtete patientenrelevante Endpunkte

Endpunktkategorie und Ausprägung	Erhebung und ggf. Validität	Details
Mortalität		
COPD-assoziierte kardiovaskuläre Morbidität und Mortalität	Auswertung der Dokumentation der schweren unerwünschten Ereignisse	Häufigkeit im Verhältnis zur Studienpopulation
COPD-bedingte Letalität	Auswertung der Dokumentation der schweren unerwünschten Ereignisse	Häufigkeit im Verhältnis zur Studienpopulation
Gesamtmortalität	Auswertung der dokumentierten Todesfälle	Häufigkeit im Verhältnis zur Studienpopulation
Morbidität (COPD-Symptome)		
Husten, Auswurf, Atemnot	Dokumentation in einem Patiententagebuch In der Validierung befindlicher EXACT-RS Score	CFB
Nächtliche und morgendliche Symptome	Dokumentation in einem Patiententagebuch	CFB
Dyspnoe	Validierter „Transition Dyspnea Index“ (TDI)	CFB Anteil der Patienten mit einer Verbesserung um ≥ 1 Einheit
Exazerbationen	Akute Verschlechterung der COPD, Operationalisierungen in den einzelnen Studien vgl. Abschnitt 4.3.1 und 4.3.2.1	Rate = Anzahl der Exazerbationen je Personenjahr
Morgendlicher Trough- FEV_1 -Wert	Messung mittels Spirometer Dabei werden die Anforderungen der American Thoracic Society (ATS) und der European Respiratory Society (ERS) bzgl. der Genauigkeit der Messung von Lungenfunktionsparametern [6] beachtet (z. B. wiederholte Messung zur Erzeugung von größtmöglicher Reliabilität)	CFB der Messung vor der morgendlichen Inhalation
Peak- FEV_1 -Wert		CFB der Messung nach der morgendlichen (und abendlichen) Inhalation
FEV_1 AUC		CFB der Ermittlung mehrerer Messwerte über den Tag verteilt
Gesundheitsbezogene Lebensqualität		

Lebensqualität	Validierter St. George's Respiratory Questionnaire (SGRQ)	CFB Anteil der Patienten mit einer Verbesserung um ≥ 4 Einheiten
Behandlungszufriedenheit	Patientenpräferenz erhoben mit: Anwendbarkeit Erlernbarkeit Handhabbarkeit	Anteil der Patienten, die einen der beiden Inhalatoren bevorzugten
Unerwünschte Arzneimittelwirkungen		
unerwünschte Ereignisse (AEs)	Auswertung der Dokumentation der unerwünschten Ereignisse	Anzahl und Anteil an Studienpopulation
schwerwiegende unerwünschte Ereignisse (SAEs)	Auswertung der Dokumentation der schwerwiegenden unerwünschten Ereignisse	Anzahl und Anteil an Studienpopulation
Studienabbruch wegen unerwünschter Ereignisse	Auswertung der Dokumentation der unerwünschten Ereignisse, die zum Studienabbruch führten	Anzahl und Anteil an Studienpopulation
CFB = Veränderung zur Baseline		

COPD-Symptome

Lungenfunktion

Im Versorgungsalltag basiert die Klassifizierung der Schwere der COPD auf der gemessenen Atemwegsobstruktion. Das forcierte expiratorische Volumen in einer Sekunde (FEV₁-Wert) gilt wie die inspiratorische Vitalkapazität (VK) und deren Verhältnis zueinander (FEV₁/FVC) als eine der Kenngrößen der Atemwegsobstruktion, besitzt jedoch die größte Evidenz [1, 5]. Der FEV₁-Wert stellt außerdem den Lungenfunktionsparameter mit der höchsten Reliabilität dar [4, 7]. Patienten mit COPD haben üblicherweise einen kleineren FEV₁-Wert, als man entsprechend ihres Alters und Geschlechts erwarten würde (Details vgl. Modul 3).

Die Veränderung des FEV₁ wird als erster Endpunkt zur Evaluation des Disease Management Programms (DMP) COPD durch das Bundesversicherungsamt herangezogen [8]. Er soll die Optimierung der Lungenfunktion unter DMP-Behandlung beschreiben. Nationale und internationale Leitlinien bestätigen, dass die Veränderung des FEV₁ der Beurteilung des Therapieansprechens auf Bronchodilatoren dient [1, 5].

Die zur Beurteilung des Therapieansprechens und zur Evaluation strukturierter Behandlungsprogramme herangezogene Veränderung des FEV₁ wird zu folgenden Messzeitpunkten als patientenrelevanter Endpunkt zur Nutzenbewertung herangezogen:

- Vor der Einnahme der Studienmedikation (FEV₁-Trough). Der FEV₁-Trough ist ein Maß für die Bronchodilatation am Ende des Dosierintervalls (nach 24 Stunden), an dem normalerweise ein geringerer Grad an Bronchodilatation beobachtet wird. Daher ist der FEV₁-Trough ein Maß der Wirkungsdauer der untersuchten Behandlung. Er wird in der Mehrheit der klinischen Studien zur Bewertung der nachhaltigen Verbesserung der Lungenfunktion bei COPD verwendet.
- Der FEV₁-Spitzenwert (FEV₁-Peak). Der FEV₁-Peak ist der höchste innerhalb einer Dauer von drei Stunden nach der morgendlichen Dosis beobachtete FEV₁-Wert. Der FEV₁-Peak wird als Maß der sofortigen Wirkung einer COPD-Behandlung auf die Atemwegsobstruktion des Patienten verwendet.
- Zur Bewertung der Verbesserung der Lungenfunktion über den gesamten Tagesverlauf wird ein weiterer Wert verwendet: der FEV₁ AUC. Die Fläche unter der Kurve (area under the curve, AUC) bildet die Lungenfunktion über 24 Stunden (FEV₁ AUC₀₋₂₄), in der ersten Tageshälfte (FEV₁ AUC₀₋₁₂) sowie in der zweiten Tageshälfte (FEV₁ AUC₁₂₋₂₄) ab. Dieser Endpunkt wird bei der Beurteilung der konstanten bronchodilatatorischen Wirkung auch im Zusammenhang mit den durch die zirkadiane Rhythmik verursachten tageszeitlichen Schwankungen herangezogen.

Zur Beurteilung des Therapieansprechens ist eine klinisch relevante Veränderung des FEV₁-Wertes von mindestens 100 ml definiert [9]. Dieser wurde auch im Festbetragsbeschluss zu Indacaterol herangezogen [10].

Dyspnoe

Dyspnoe als Folge einer Atemwegsobstruktion ist eine Hauptursache von Behinderungen und Angstgefühlen von Patienten mit COPD. Der Schweregrad einer Dyspnoe kann mit dem Transitional Dyspnea Index (TDI) gemessen werden. Dieser Index misst den Schweregrad der Atemlosigkeit bei symptomatischen COPD-Patienten. Der TDI ist ein validiertes Maß für symptomatische Patienten und enthält eine Bewertung für 3 Kategorien von Faktoren, die den Schweregrad der vom Patienten wahrgenommenen Dyspnoe beeinflussen. Diese definierten Kategorien sind:

- Funktionelle Beeinträchtigung, die die Auswirkungen der Atemlosigkeit auf die Fähigkeit zur Ausführung von Aktivitäten bestimmt
- Ausmaß der Aufgabe, welche die Atemlosigkeit verursacht
- Ausmaß der Anstrengung, welche die Atemlosigkeit verursacht

Damit man einen Focal Score im Bereich zwischen -9 (Verschlechterung der Dyspnoe) und +9 (Verbesserung der Dyspnoe) erhält, werden die von den Patienten angegebenen Werte für

die drei Kategorien aufsummiert. Eine positive Veränderung von 1 Punkt im TDI wird als die kleinste bedeutsame Verbesserung angesehen und Patienten mit einer Verbesserung von ≥ 1 Punkt im TDI werden als TDI-Responder angesehen [11].

COPD-Symptomatik

Die drei Hauptsymptome der COPD (Husten, Auswurf, Atemnot) werden mit den Ergebnissen des EXACT-RS-Fragebogen ausgewertet. Der EXACT-RS-Fragebogen stellt einen Teil des in der Validierung befindlichen EXACT-Fragebogens dar, der die Schwere der charakteristischen Symptome der COPD (Atemnot, Husten & Auswurf, Brustenge) untersucht [12, 13]. Das **EX**acerbations of **Ch**ronic Pulmonary Disease **T**ool (EXACT) ist ein Messinstrument für Patient-Reported-Outcomes, das entwickelt wurde, um standardisierte internationale Methoden zur Messung der Evaluation von Frequenz, Schwere und Dauer von akuten COPD-Exazerbationen und chronischer Bronchitis für internationale und klinische Studien zu erschaffen. Der Fragebogen ist zur Validierung bei der EMA und der FDA eingereicht⁸.

Nächtliche und morgendliche COPD-Symptome sind weitere wichtige Maße zum Beschreiben des wahrgenommenen Gesundheitszustands von COPD-Patienten. Sie geben Hinweise auf die durch die zirkadiane Rhythmik verursachten tageszeitlichen Schwankungen der COPD [14, 15]. Die Auswirkungen von COPD-Symptomen auf Patienten werden für den Schlaf, morgendliche Aktivitäten und für die erste Stunde nach dem Aufstehen bewertet.

Exazerbationen

Eine Exazerbation einer COPD ist als akutes Ereignis definiert, das durch eine Verschlechterung der respiratorischen Symptome charakterisiert ist, und das in einem Ausmaß, das die normalen Tagesschwankungen übertrifft und zu einer Änderung der regulären Medikation führt. Da Exazerbationen die Abnahme der Lungenfunktion, die Verschlechterung des Gesundheitszustandes und die Verstärkung der Todesgefahr zur Folge haben, wird die Bewertung des Exazerbationsrisikos als patientenrelevanter Endpunkt angesehen [5]. Exazerbationen wurden auf Grundlage des EXACT-Fragebogens zu COPD-Symptomen und der Anzahl der Exazerbationen während des Krankenhausaufenthalts in die Analysen mit aufgenommen.

Gesundheitsbezogene Lebensqualität

Zur Bewertung der Lebensqualität von COPD-Patienten wird der „St. George’s Respiratory Questionnaire“ (SGRQ) verwendet. Dies ist ein objektiver Fragebogen zur Messung der Lebensqualität von Patienten mit einer Atemwegserkrankung. Der SGRQ ist in der GOLD-Leitlinie [5] als einer der erkrankungsspezifischen Fragebögen zur Bewertung des Gesundheitszustands anerkannt. Der Fragebogen besteht aus 50 Punkten, die in zwei Teile unterteilt sind und drei Abschnitte enthalten:

- Symptome: betrifft COPD-Symptome und deren Häufigkeit sowie Schwere

⁸ <http://www.exactproinitiative.com/default.php>

- **Aktivität:** betrifft Aktivitäten, die die Atemnot verstärken oder die aufgrund von Atemnot weniger gut ausgeführt werden können
- **Auswirkungen:** deckt einen Bereich von Aspekten mit Bezug zu sozialen, funktionalen und psychologischen Störungen ab, die aufgrund der Erkrankung auftreten

Der erste Teil des Fragebogens (Fragen 1 bis 8) ermittelt die Punktzahl für den Abschnitt „Symptome“ und fragt nach Erinnerungen der Patienten an Symptome in einem vorangegangenen Zeitraum. Der zweite Teil (Fragen 9 bis 17) ermittelt die Punktzahlen für die Abschnitte „Aktivität“ und „Auswirkungen“ und fragt nach dem gegenwärtigen Zustand der Patienten. Für die Analyse werden das Gesamtergebnis des SGRQ und die Anzahl von SGRQ-Respondern verwendet. Patienten gelten als SGRQ-Responder, wenn die über alle Messungen angegebenen Veränderungen einen Unterschied im SGRQ-Ergebnis von mindestens 4 Punkten ergeben, da dies als kleinste klinisch relevante Änderung angesehen wird [16].

Die Zufriedenheit der Patientinnen und Patienten mit der Behandlung kann ebenfalls berücksichtigt werden [2] und wird als ein weiterer Lebensqualitätsaspekt ausgewertet. Dazu wird die Zufriedenheit und Handhabbarkeit des verwendeten Inhalationssystems ergänzend zur Nutzen- und Zusatznutzenbewertung herangezogen.

Unerwünschte Arzneimittelwirkungen

Bei der Bewertung einer medikamentösen Behandlung im Rahmen einer klinischen Studie ist die Erfassung unerwünschter Ereignisse von größter Bedeutung für die Sicherheitsbewertung des Arzneimittels. Daher werden in diesem Bericht alle unerwünschten Ereignisse als Ergebnis angegeben. Schwere unerwünschte Ereignisse sind Ereignisse, die zum Tod führten, das Leben des Patienten bedrohten oder Einweisungen in ein Krankenhaus, Behinderungen oder Geburtsfehler nach sich zogen. Zusätzlich wurden Studienabbrüche aufgrund von unerwünschten Ereignissen zur Bewertung der Verträglichkeit der Studienmedikation angegeben.

Mortalität

Beim Endpunkt Mortalität handelt es sich um einen in der VerFO definierten patientenrelevanten Endpunkt. Die Analysen im vorliegenden Dossier schließen die Letalität, die Gesamtmortalität als auch die kardiovaskuläre Morbidität und Mortalität ein.

4.2.5.3 Meta-Analysen

Sofern die vorliegenden Studien dazu geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst werden. Dabei sollen verschiedene Studientypen nicht in einer Meta-Analyse zusammengefasst werden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten⁹ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär anhand des Odds Ratio durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet¹⁰ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes I^2 und des statistischen Tests auf Vorliegen von Heterogenität¹¹ erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

⁹ DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-188.

¹⁰ Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

¹¹ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.

Allgemein

Für die beiden für den direkten Vergleich herangezogenen Studien kann aufgrund der unterschiedlichen Charakteristika der Studien keine Meta-Analyse durchgeführt werden. Dies betrifft vor allen Dingen die unterschiedliche Studiendauer (15 Tage vs. 6 Wochen) und das Design (3-Perioden-cross-over vs. parallel).

Für den indirekten Vergleich werden drei placebokontrollierte Studien des zu bewertenden Arzneimittels und 21 placebokontrollierte Studien der Vergleichstherapie herangezogen. Für diese werden jeweils Meta-Analysen durchgeführt.

Dazu werden Effektschätzer und Konfidenzintervalle aus den Studien mittels Forest Plots zusammengefasst. Die für die sinnvolle Zusammenfassung von Studienergebnissen erforderliche Homogenität der Studien wird mit Hilfe des statistischen Maßes I^2 und des p-Wertes zur Beschreibung der Signifikanz des Ergebnisses geprüft.

Für die Zusammenfassung der Acclidiniumbromid-Studien werden aufgrund der in Abschnitt 4.3.2.1.2 dargestellten vergleichbaren Charakteristika des Studiendesigns und Studienpopulation gute Homogenitätswerte erwartet. Bei der Zusammenfassung der teils vom Studiendesign und der Patientenpopulation her unterschiedlichen Tiotropiumbromid-Studien kann es bei manchen Endpunkten zu Heterogenität kommen. Ist dies der Fall, werden Sensitivitätsanalysen durchgeführt.

Methodik

Meta-Analysen wurden für alle Endpunkte und Behandlungen (Acclidiniumbromid, Tiotropiumbromid) durchgeführt, zu denen Daten verfügbar waren. Die Meta-Analyse wurde mit dem Softwarepaket R¹² unter Verwendung der meta-group und metafor Funktionen durchgeführt. Meta-group wurde verwendet, wenn detaillierte Ergebnisse (Baseline, Follow-up, Veränderung) pro Behandlungsarm zur Verfügung standen und es sich bei den Ergebnissen nicht um Raten handelte. Metafor wurde zur Analyse von Raten und von Publikationen verwendet, die nur den Unterschied (Veränderung) zwischen den Behandlungsarmen berichteten. Die angegebenen Ergebnisse beziehen sich auf ITT-Analysen. Alle Meta-Analysen wurden auf der Basis des Modells mit zufälligen Effekten zusammengefasst. Die Heterogenität wurde mittels Q-Statistik und I^2 -Statistik überprüft.

Binäre Variablen werden mit dem OR (Odds Ratio) zusammengefasst. Je nach Art der Publikation der Ergebnisse werden kontinuierliche Variablen mit der Differenz der Mittelwerte und des Standardfehlers der Differenz der Mittelwerte (Verwendung der Funktionen der meta-group) oder mit der Mittelwertdifferenz und des Standardfehlers der Mittelwertdifferenz (Verwendung der metafor Funktionen) zusammengefasst. Die Exazerbationsraten wurden von auf der Poisson-Likelihood-Verteilung basierenden Raten mit Hilfe von metafor zusammengefasst.

¹² <http://www.r-statistik.de/>

Die Heterogenität der Studien wurde unter Verwendung von Forest Plots visualisiert, wobei die Behandlungseffekte der einzelnen Studien (und auch eine Zusammenfassung der erwarteten, studienübergreifenden Effekte) mit entsprechenden Konfidenzintervallen zusammen mit den zusammengefassten Behandlungseffekten auf der Basis des Modells mit zufälligen Effekten dargestellt wurden. Die quantitative Einschätzung der statistischen Heterogenität zwischen den Studien wurden mit einem Heterogenitätstest (Q-Statistik, p-Wert) und der Berechnung der I^2 -Statistik durchgeführt [17]. Für den Grenzwert einer bestehenden oder nicht bestehenden Heterogenität existieren lediglich grobe Richtwerte [2]. Entscheidender ist das Vorliegen einer statistisch signifikanten Heterogenität. Bei einem p-Wert unter 0,2 sollte und wird von einer Zusammenfassung der Ergebnisse abgesehen werden [2].

Der Heterogenität wurde mit der Verwendung von Modellen mit zufälligen Effekten und mit Sensitivitätsanalysen Rechnung getragen. Ein weiterer Ansatz zum Umgang mit Heterogenität wäre der Einsatz von erklärenden Kovariaten in die Analysen gewesen. Aufgrund der kleinen Zahl von Studien (weniger als 10) in jeder Meta-Analyse wurde dies als nicht informativ bzw. zielführend angesehen.

Im Base-Case der Meta-Analysen wurden alle Studien aufgenommen. Wenn Daten für beide Tiotropiumbromiddosierungen (5 oder 18 µg) verfügbar waren, wurden diese Daten gepoolt. Die beiden Präparate wurden von einer Studie von Boehringer Ingelheim Pharmaceuticals [18] als vergleichbar erachtet; deshalb wurde die Tiotropiumbromiddosierung nicht als primäre Heterogenitätsquelle betrachtet. Wenn Daten entweder nur für 18 µg Tiotropiumbromid oder nur für 5 µg Tiotropiumbromid angegeben wurden, wurde die Meta-Analyse nur für die entsprechende Dosierung durchgeführt.

Einige Faktoren standen vor der Durchführung der Meta-Analyse im Verdacht, eine Heterogenitätsquelle zu sein (z. B. das Ungleichgewicht der Ausgangswerte der Patienteneigenschaften bei ACCORD II); andere wurden im Verlauf der Analyse gefunden (z. B. Tiotropiumbromid-Studien, die aufgrund des nicht verwendeten ITT-Prinzips ein hohes Verzerrungspotential aufwiesen). Zur Verifizierung der Heterogenität wurden Sensitivitätsanalysen durchgeführt. Dies wurde für jeden potenziellen Heterogenitätsfaktor vorgenommen (vgl. Abschnitt 4.2.5.4), sofern der Datensatz der aufgenommenen Studie den Anforderungen dieser Analysen gerecht wurde.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden, wenn sich die Studien bezüglich dieser Faktoren unterscheiden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Robustheit einer Metaanalyse oder einer Netzwerk-Meta-Analyse kann mit Bootstrapmethoden untersucht werden. Hierzu werden eine oder mehrere Studien ausgeschlossen und anschließend die Analyse wiederholt. Mit diesem Vorgehen kann gezeigt werden, wie stabil die Ergebnisse der Analyse sind und welche Studien die Ergebnisse am stärksten beeinflussen.

Bei der vorgenommenen Metaanalyse der Aclidiniumbromid-Studien wird aufgrund des Verzerrungspotentials der ACCORD II Studie [19] je Endpunkt ein Szenario ohne diese Studie dargestellt. So wird die Aussagesicherheit der Aclidiniumbromid-Studien geprüft. Bei der vorgenommenen Metaanalyse der Tiotropiumbromid-Studien wird mit verzerrten Studien ebenso verfahren – sie werden in einem Szenario aus der Analyse ausgeschlossen. Sollten weniger als zwei nicht verzerrte Studien im zu analysierenden Studienpool verbleiben, kann die beschriebene Sensitivitätsanalyse nicht durchgeführt werden.

Um den Einfluss der Unterschiede zwischen den Patientenpopulationen auf die Ergebnisse bewerten zu können, wird mit Hilfe eines weiteren Szenarios eine Sensitivitätsanalyse durchgeführt, die die Studien Johansson 2008 [20] und Freeman 2007 (SPRUCE) [21] aus der Metaanalyse ausschließt. Die Patientenpopulation dieser Studie bestand fast nur aus Patienten mit mittelschwerer COPD (Stufe II), während die anderen in der Analyse berücksichtigten Studien auch einen größeren Anteil von Patienten mit schwererer Ausprägung von COPD enthielten.

Der trotz der Device-Studie [18] theoretisch mögliche Fall einer durch die unterschiedliche Dosierung von Tiotropiumbromid begründeten Heterogenität wird mit einer separaten Analyse der 5 und 18 µg Studien begegnet.

Kann die Heterogenität der Tiotropiumbromid-Studien nicht mit den beschriebenen Sensitivitätsanalysen begründet werden, kann diese nicht statistisch erklärt werden. Ist dies der Fall, werden mögliche Antworten aus dem klinischen Zusammenhang gesucht. Dabei

kann es sich beispielsweise um die unterschiedliche Operationalisierung der Endpunkte in den Studien handeln.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressions- oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollten, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. -stadium

Sollten sich aus den verfügbaren Informationen Anhaltspunkte für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen.

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.

Die Durchführung von Subgruppenanalysen dient der Identifizierung von Faktoren (z. B. Patientencharakteristika), die zu unterschiedlichen Effekten in Teilpopulationen führen können. „Analog zum Vorgehen, Studien mit zu großer Heterogenität nicht mithilfe von Meta-Analysen zusammenzufassen, sollten auch Ergebnisse von Subgruppen nicht zu einem

gemeinsamen Effektschätzer zusammengefasst werden, wenn sich die Subgruppen zu stark voneinander unterscheiden.“ [2] (S. 139)

Die Ähnlichkeit der Subgruppen wird durch den Vergleich der jeweiligen Effekte (signifikant, gleichgerichtet) in den Subgruppen und mit Hilfe eines Interaktionstest überprüft. Ist die Interaktion auf einem Niveau von $p=0,05$ signifikant liegt eine Beleg für einen Unterschied zwischen den Subgruppen vor. Aussagen und Ergebnisschätzer sind so nur getrennt für die Subgruppen möglich. Liegt der p -Wert zwischen 0,05 und 0,2 liegt ein Hinweis auf Effektmodifikation vor – die Schätzer der Gesamtpopulation und der Subgruppen werden getrennt dargestellt und verglichen. Liegen bei den Subgruppen und der Gesamtpopulation gleichgerichtete Ergebnisse vor, können die Ergebnisse als Schätzer der Gesamtpopulation dargestellt werden. Bei Inkonsistenz zwischen den Subgruppen und der Gesamtpopulation kann die Aussagesicherheit für die Subgruppe ggf. herabgestuft werden. Bei einem Interaktionstestergebnis von $p > 0,2$ gibt es keinen Hinweis auf einen Unterschied zwischen den Subgruppen – sie werden als vergleichbar angesehen. [2]

Als relevante Effektmodifikatoren werden im Allgemeinen [2] das Alter, das Geschlecht und die Schwere der Erkrankung in Betracht gezogen. Die Nutzenbewertung des IQWiG zu Tiotropiumbromid (A05-18) [22] bezieht außerdem den Raucherstatus in die Subgruppenanalysen ein. Das IQWiG resümiert, dass die bezüglich der vier potentiellen Effektmodifikatoren (Alter, Geschlecht, Raucherstatus, Erkrankungsschwere) durchgeführten Subgruppenanalysen keine Auswirkungen auf die Nutzenbewertung hatten. Lediglich die Subgruppenergebnisse einer Tiotropiumbromid-Studie (UPLIFT) ergänzten das Fazit der Nutzenbewertung. Die zu Grunde liegenden Daten stammen aus Studienberichten und lagen zur Erstellung dieses Dossiers nicht vor. Eine aktuelle Netzwerk-Meta-Analyse (NMA) zur Behandlung mit LABAs [23] zeigt außerdem, dass Alter und Geschlecht in der Indikation COPD nicht als Effektmodifikatoren relevant sind.

Subgruppenergebnisse nach Alter, Geschlecht und Raucherstatus sind bis auf eine Ausnahme nicht in den eingeschlossenen Studien a priori geplant und werden nicht im Dossier dargestellt. Die a priori geplanten Subgruppenanalysen einer Studie des Studienpools (placebokontrollierte Acclidiniumbromid-Studie ATTAIN) werden in Abschnitt 4.3.2.1.3.17 für folgende (vgl. Tabelle 4.2-7) Subgruppen und Endpunkte präsentiert.

Tabelle 4.2-7: A priori Subgruppenanalysen in der ATTAIN-Studie

Subgruppen	Endpunkte
<ul style="list-style-type: none"> · Alter · Geschlecht · COPD-Schweregrad · Raucherstatus · Reversibility · BMI · Behandlung mit ICS 	<ul style="list-style-type: none"> · Peak-FEV₁ · SGRQ · TDI

Für die placebokontrollierten Tiotropiumbromid-Studien können aufgrund fehlender Subgruppendaten (a priori oder post-hoc) in den zur Verfügung stehenden Publikationen keine Subgruppenanalysen durchgeführt werden.

Dementsprechend werden im vorliegenden Dossier bzgl. Alter, Geschlecht und Raucherstatus keine post-hoc Subgruppenanalysen durchgeführt.

Der im Beratungsgespräch [3] empfohlenen Berücksichtigung der COPD-Schweregrade (nach dem Stufenschema der Nationalen VersorgungsLeitlinie (NVL) [1]) soll mit einer entsprechenden Subgruppenanalyse Rechnung getragen werden. Dazu werden Subgruppen untersucht, die sich aus dem Schweregrad (A) und der Exazerbationshistorie (B) zusammensetzen (vgl. Tabelle 4.2-8).

Tabelle 4.2-8: Subgruppen nach COPD-Schweregrad-Einteilung der NVL

Bestandteil Subgruppe	A	B
	Schweregrad nach NVL	Exazerbationen
1	Stufe II	-
2	Stufe III	$n \leq 2$ in den letzten 12 Monaten vor Studienbeginn
3	Stufe III	$n > 2$ in den letzten 12 Monaten vor Studienbeginn
4	Stufe IV	$n \leq 2$ in den letzten 12 Monaten vor Studienbeginn
5	Stufe IV	$n > 2$ in den letzten 12 Monaten vor Studienbeginn

Liegen keine Exazerbationsdaten vor oder kann die Analyse aufgrund zu kleiner Patientenzahlen nicht durchgeführt werden, wird die Subgruppenanalyse lediglich nach Schweregrad (Schema A) vorgenommen.

Im Rahmen der Subgruppenanalyse werden folgende Schritte vorgenommen:

1. Darstellung der einzelnen Endpunkte nach Subgruppen
2. Meta-Analyse je Subgruppe (Methodik analog zu 4.2.5.3)
3. Kontrolle der Interaktion mit Hilfe eines Interaktionstests

Für die Durchführung des Interaktionstestes bei kontinuierlichen und kategorialen Variablen werden folgende zwei Modelle verwendet.

Für kontinuierliche Variablen

Der Behandlungseffekt in Subpopulationen wurde bei kontinuierlichen Variablen mittels eines Modells der Kovarianzanalyse (ANCOVA) analysiert. Dabei stellten die Baselinewerte und das Alter Kovariate und Intervention, Geschlecht, Schweregrad und die Interaktion

zwischen Intervention und Subgruppe Faktoren im Fixed-Effects-Modell dar. Zur Ermittlung der Interaktion zwischen den Subgruppen wurden die Mittelwerte der kleinsten Quadrat-Mittelwerte (least squares (LS) means), die Standardabweichungen (standard error, SE), die 95%-igen Konfidenzintervalle (KI) und die p-Werte der Differenzen zwischen den beiden relevanten Vergleichsgruppen (Aclidiniumbromid 400 µg und Placebo) verwendet.

Für kategoriale Variablen

Der Behandlungseffekt in den Subgruppen wurde bei kategorialen Variablen mit Hilfe der logistischen Regression analysiert. Dabei stellten die Intervention, das Geschlecht, das Alter, die Baselinewerte und die Interaktion zwischen Intervention und Subgruppe die erklärenden Variablen dar. Die statistische Signifikanz der Koeffizienten des Modells wurde unter Verwendung des Wald-Tests überprüft. Zur Ermittlung der Interaktion zwischen den Subgruppen wurden die Odds Ratios (OR), die KIs von 95% und die Standardabweichungen zwischen den beiden relevanten Vergleichsgruppen (Aclidiniumbromid 400 µg und Placebo) verwendet.

In beiden Fällen ist die Berechnung der Interaktionen nur dann möglich, wenn die Anzahl der Beobachtungen die vom jeweiligen Modell benötigten Freiheitsgrade (Kovariate und Faktoren bzw. erklärende Variablen) übersteigt. Ist diese Voraussetzung nicht erfüllt, kann der Interaktionswert für diesen Endpunkt nicht berechnet werden. In diesem Fall werden die Subgruppenergebnisse hinsichtlich ihrer Gleichgerichtetheit miteinander verglichen und bewertet.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich¹³. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“¹⁴, „Multiple-Treatment-Meta-Analysen“¹⁵ oder auch „Netzwerk-Meta-Analysen“¹⁶, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹⁷. Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen¹⁸.

Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an,*

¹³ Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-148.

¹⁴ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

¹⁵ Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

¹⁶ Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17(3): 279-301.

¹⁷ B. Schöttker, D. Lühmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

¹⁸ Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).

- *Art und Umfang von Sensitivitätsanalysen.*

Statistisches Modell

Bayesianische Netzwerk-Meta-Analysemodelle (NMA) nach Caldwell et al. 2005 [24], Lu und Ades 2004 [25] und Jansen et al. 2008 [26] wurden zur Analyse der erhobenen patientenrelevanten Endpunkte verwendet, um gleichzeitig eine Synthese der Ergebnisse der aufgenommenen Studien und der relativen Behandlungseffekte zu erhalten.

Bayesianische Netzwerk-Meta-Analysemodelle beinhalten Daten, eine Wahrscheinlichkeitsverteilung, ein Modell mit Parametern und A-priori-Verteilungen [26]. Das Modell bezieht die Daten aus den einzelnen Studien auf grundlegende Parameter und reflektiert somit die (zusammengefassten) relativen Behandlungseffekte jeder Intervention im Vergleich zu einer allgemeinen Referenzbehandlung, z. B. Placebo. Auf der Grundlage dieser Parameter erhält man die relative Wirksamkeit jeder der verglichenen Interventionen.

Für die kontinuierlichen Variablen wurde ein Regressionsmodell mit normaler Wahrscheinlichkeitsverteilung (z. B. Trough-FEV₁) und für die binären Variablen ein Regressionsmodell mit einer binominalen Wahrscheinlichkeitsverteilung verwendet (z. B. SGRQ- und TDI-Responder) [25, 26]. Für die Exazerbationsraten wurde ein Regressionsmodell mit einer Poisson-Likelihood-Verteilung verwendet. Die Modellschlüssel sind im Unterpunkt „Sensitivitätsanalysen“ für jeden analysierten Ergebnistyp aufgeführt.

Um den Einfluss vorhandener Vermutungen bzgl. der Ergebnisse des Modells zu verhindern, wurden nicht-informative A-priori-Verteilungen verwendet. Für kontinuierliche Variablen waren A-priori-Verteilungen der relativen Behandlungseffekte Normalverteilungen mit einem Mittelwert von 0 und einer Varianz von 10¹⁰. Für Modelle mit binären Variablen und Raten waren A-priori-Verteilungen des Referenz-Behandlungseffekts ebenfalls Normalverteilungen mit einem Mittelwert von 0 und einer Varianz von 10⁴.

Initialwerte für den relativen Behandlungseffekt wurden automatisch von der Software generiert. Wenn die von der Software generierten Werte zu Berechnungsproblemen führten, wurde der Initialwert gleich 0 gesetzt. Es wurde eine Markov-Kette mit einem Startwert von 20.000 Iterationen und einer Länge von 80.000 Iterationen ausgeführt.

Homogenität

Für jeden Endpunkt wurden ein Modell mit festen und eines mit zufälligen Effekten evaluiert. Das Modell mit festen Effekten basiert auf der Annahme, dass der wahre Behandlungseffekt von jedem paarweisen Behandlungsvergleich nicht zwischen den Studien variiert.

Das Modell mit zufälligen Effekten basiert hingegen auf der Annahme, dass der wahre Behandlungseffekt eines jeden Behandlungsvergleiches z. B. aufgrund von Patienten- und Studieneigenschaften variiert.

Es wurde davon ausgegangen, dass das Ausmaß der Heterogenität bei allen Behandlungsvergleichen gleich war.

Das Modell mit zufälligen Effekten kann für den Umgang mit der Heterogenität verwendet werden, wenn diese Heterogenität zufälliger und nicht strukturierter Natur ist. Deswegen wurde zusätzlich zur Verwendung des Modells mit zufälligen Effekten eine Sensitivitätsanalyse durchgeführt, die Studien mit einer zugelassenen Begleitbehandlung mit einem LABA ausschlossen.

Um das am besten geeignete Modell (d. h. mit festen oder zufälligen Effekten, mit oder ohne Studien mit begleitender LABA-Behandlung) zu identifizieren, kann die Eignung der Modellvorhersagen für die entsprechenden Daten durch Berechnung der verbleibenden Abweichung des posteriori-Mittelwerts \bar{D} gemessen werden [27]. Die Entscheidung, ob Modelle mit zufälligen oder festen Effekten verwendet werden, kann das Ergebnis der Analyse beeinflussen. Daher wurde der indirekte Vergleich jeweils mit zufälligen und festen Effekten durchgeführt. Das Abweichungsinformationskriterium DIC (Deviance Information Criterion) wurde für den Vergleich der Modelle mit festen und zufälligen Effekten verwendet und ist ein Maß der Modellanpassung, das Modellkomplexität nach $DIC = \bar{D} + pD$, $pD = \bar{D} - \hat{D}$ [28] benachteiligt. pD ist die „effektive Parameterzahl“ und \hat{D} ist die Abweichung, die aus dem posteriori-Mittelwert der Modellparameter evaluiert wurde. Das Modell mit dem niedrigsten DIC und somit der „besten“ Anpassung wurde als Basismodell für den verwendeten Datensatz angesehen.

Für die A-priori-Verteilung der Heterogenität in den Modellen mit zufälligen Effekten wurde eine gleichmäßige Verteilung in einem Bereich von 0 bis 5 verwendet. Die Initialwerte wurden auf 1 gesetzt.

Konsistenz

Eine Prüfung auf Konsistenz der Ergebnisse im Netzwerk wurde nicht durchgeführt, da keine direkten Vergleiche zwischen Acridiniumbromid und Tiotropiumbromid mit vergleichbarer Studiendauer vorlagen, die für eine Konsistenzprüfung herangezogen werden könnten.

Verwendete Software und Codes

Die Statistiksoftware WinBUGS 1.4.1 wurde für die Analysen verwendet [29].

Die Modelle mit zufälligen Effekten für die unterschiedlichen Endpunktarten sind unten aufgeführt (kontinuierliche Variablen, binäre Variablen und Raten). Bei Modellen mit festen Effekten, wird σ^2 gleich Null gesetzt. Alle WinBUGS-Codes basieren auf den vom NICE [30] publizierten Codes.

Bayesianisches Modell zur Netzwerk-Meta-Analyse kontinuierlicher Variablen**a. Random Effects (zufällige Effekte)**

$$diff_{j bk} \sim Normal(\delta_{j bk}, \sigma_{j bk}^2) \quad (\text{Wahrscheinlichkeit})$$

$$\delta_{j bk} \sim Normal(d_{bk}, \sigma^2) = Normal(d_{Ak} - d_{Ab}, \sigma^2) \quad (\text{Modell mit zufälligen Effekten})$$

$$d_{AA} = 0$$

$$d_{Ak} \sim Normal(0, 0.0001) \quad (\text{A-priori-Verteilungen})$$

$$\sigma \sim Uniform(0, 10)$$

Wo

j = Studie

b = Kontrollgruppe, kann Behandlung A, B, C, \dots sein

k = Behandlungsgruppe, kann B, C, D, \dots sein

$diff_{j bk}$ = Beobachteter Unterschied der Änderung gegenüber Ausgangswert eines kontinuierlichen Ergebnisses von Intervention k im Vergleich zur Referenzbehandlung b in Studie j

$\delta_{j bk}$ = Unterschied der Änderung gegenüber Ausgangswert von Behandlung k in Relation zu Komparator b in Studie j

$\sigma_{j bk}$ = Unsicherheit („sampling error“) im Unterschied der Änderung gegenüber Ausgangswert von Intervention k im Vergleich zur Referenzbehandlung b in Studie j

d_{Ak} = Gepoolter Unterschied der Änderung gegenüber Ausgangswert von Intervention k in Relation zur Referenzbehandlung A

σ^2 = Konstante Varianz der relativen Effekte jeder Behandlung im Vergleich zur Referenzbehandlung A über alle Studien als Maß der Heterogenität

b. Fixed Effects (feste Effekte)

Bei Modellen mit festen Effekten, wird σ gleich Null gesetzt: $\delta_{j bk} = d_{bk}$

Bayesianisches Modell zur Netzwerk-Meta-Analyse binärer Variablen**a. Random Effects (zufällige Effekte)**

$$r_{jk} \sim \text{Binomial}(p_{jk}, n_{jk}) \quad (\text{Wahrscheinlichkeit})$$

$$\text{logit}(p_{jk}) = \begin{cases} \mu_{jb} & b = A, B, C, \dots \quad \text{if } k = b \\ \mu_{jb} + \delta_{jbk} & \text{if } k \text{ alphabetically after } b \end{cases}$$

$$\delta_{jbk} \sim \text{Normal}(d_{bk}, \sigma^2) = \text{Normal}(d_{Ak} - d_{Ab}, \sigma^2) \quad (\text{Modell mit zufälligen Effekten})$$

$$d_{AA} = 0$$

$$\mu_{jb} \sim \text{Normal}(0, 0.0001)$$

$$d_{Ak} \sim \text{Normal}(0, 0.0001) \quad (\text{A-priori-Verteilungen})$$

$$\sigma \sim \text{Uniform}(0, 10)$$

Wo

j = Studie

b = Kontrollgruppe, kann Behandlung A, B, C, ... sein

k = Behandlungsgruppe, kann B, C, D, ... sein

r_{jk} = Anzahl der bei Behandlung k in Studie j beobachteten Ereignisse

p_{jk} = Wahrscheinlichkeit eines Ereignisses bei Behandlung k in Studie j

n_{jk} = Anzahl der Patienten für Behandlung k in Studie j

μ_{jb} = Logarithmische Quote (Odd) eines Ereignisses bei „Grundbehandlung“ b in Studie j

δ_{jbk} = Logarithmisches Odds Ratio (OR) von Behandlung k in Relation zu Behandlung b in Studie j

d_{bk} = Gepooltes logarithmisches OR eines Ereignisses bei Behandlung k in Relation zu Behandlung b

σ^2 = Varianz zwischen den Studien und „Heterogenitätsparameter“

d_{Ak} = Gepooltes logarithmisches OR eines Ereignisses bei Behandlung k im Vergleich zu A

d_{Ab} = Gepooltes logarithmisches OR eines Ereignisses bei Behandlung b im Vergleich zu A

b. Fixed Effects (feste Effekte)

Bei Modellen mit festen Effekten, wird σ^2 gleich Null gesetzt: $\delta_{jbk} = d_{bk}$

Bayesianisches Modell zur Netzwerk-Meta-Analyse von Raten

a. Random Effects (zufällige Effekte)

$$y_{jk} \sim \text{Poisson}(\lambda_{jk} * T_{jk}) \quad (\text{Wahrscheinlichkeit})$$

$$\log(\lambda_{jk}) = \begin{cases} \mu_{jb} & b = A, B, C, \dots \quad \text{if } k = b \\ \mu_{jb} + \delta_{jkb} & \text{if } k \text{ alphabetically after } b \end{cases}$$

$$\delta_{jkb} \sim \text{Normal}(d_{bk}, \sigma^2) = \text{Normal}(d_{Ak} - d_{Ab}, \sigma^2) \quad (\text{Modell mit zufälligen Effekten})$$

$$d_{AA} = 0$$

$$\mu_{jb} \sim \text{Normal}(0, 0.00001)$$

$$d_{Ak} \sim \text{Normal}(0, 0.00001) \quad (\text{A-priori-Verteilungen})$$

$$\sigma \sim \text{Uniform}(0, 10)$$

Wo

j = Studie

b = Kontrollgruppe, kann Behandlung A, B, C, ... sein

k = Behandlungsgruppe, kann B, C, D, ... sein

y_{jk} = Anzahl der bei Behandlung k in Studie j beobachteten Ereignisse

λ_{jk} = Ereignisrate („event rate“) pro Person und Jahr für Behandlung k in Studie j

T_{jk} = Anzahl der Patientenjahren für Behandlung k in Studie j

μ_{jb} = Logarithmus der Ereignisrate pro Patientjahr für die Baseline-Behandlung b in Studie j

δ_{jkb} = Logarithmus des Ratenverhältnisses von Behandlung k in Relation zu Behandlung b in Studie j

d_{bk} = Gepoolte Logarithmen des Ratenverhältnisses für Behandlung k in Relation zu Behandlung b

σ^2 = Varianz zwischen den Studien und „Heterogenitätsparameter“

d_{Ak} = Gepoolte Logarithmen des Ratenverhältnisses für Behandlung k im Vergleich zu A

d_{Ab} = Gepoolte Logarithmen des Ratenverhältnisses für Behandlung b im Vergleich zu A

b. Fixed Effects (feste Effekte)

Bei Modellen mit festen Effekten, wird σ gleich Null gesetzt: $\delta_{jkb} = d_{bk}$

Sensitivitätsanalysen

Modelle mit festen und zufälligen Effekten wurden evaluiert, um die Heterogenität der Evidenz einzuschätzen. Einige Studien verwendeten zusätzlich zur Behandlung mit Acclidiniumbromid oder Tiotropiumbromid eine begleitende LABA-Behandlung. Da diese die Wirkung von Acclidiniumbromid bzw. Tiotropiumbromid verstärkt, wird die begleitende LABA-Behandlung als potenziell relativer Behandlungseffektmodifikator angesehen. Deshalb wurden Szenarioanalysen ohne Studien mit LABA-Begleitbehandlung durchgeführt.

Für jeden Endpunkt wurden die folgenden Modelle berechnet:

- ein Modell mit festen Effekten ohne die Studien mit begleitender LABA-Behandlung
- ein Modell mit zufälligen Effekten ohne die Studien mit begleitender LABA-Behandlung
- ein Modell mit festen Effekten mit allen Studien
- ein Modell mit zufälligen Effekten mit allen Studien

Aufgrund der begrenzten Anzahl von Studien und der Unvollständigkeit der aufgeführten Kovariatswerte in den aufgenommenen Publikationen wurden keine Kovariate in diese Modelle eingebracht. Die in Erwägung gezogenen Kovariate waren Rauchen, Schweregrad der Erkrankung und Verwendung von ICS.

Subgruppenanalysen

Aufgrund der fehlenden Subgruppendaten (bzgl. der Schwere der COPD oder weiterer Effektmodifikatoren) in den Publikationen der Tiotropiumbromid-Studien war die Durchführung von Subgruppenanalysen im Rahmen des indirekten Vergleichs nicht möglich. Sollten innerhalb der Subgruppenanalyse (bzgl. der Schwere der COPD) der Acclidiniumbromid-Studien (vgl. 4.3.2.1.3.18) Unterschiede festgestellt werden, können die Ergebnisse des indirekten Vergleiches lediglich dahingehend diskutiert werden.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.1.3.18)
- Weitere Untersuchungen (Abschnitt 4.3.2.4)

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Benennen Sie in der nachfolgenden Tabelle alle Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war.

Von den randomisierten kontrollierten Studien, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden, sollen die Studien, deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle randomisierten kontrollierten Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.

Benennen Sie nur randomisierte, kontrollierte Studien, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

In der folgenden Tabelle 4.3-1 sind alle Studien des pharmazeutischen Unternehmers im Anwendungsgebiet COPD aufgelistet. Im ersten Teil der Tabelle sind die im Studienprogramm zuerst durchgeführten Studien mit der einmal täglichen Gabe aufgelistet. Im zweiten Abschnitt der Tabelle sind die Studien mit einer kombinierten Gabe von Acridiniumbromid und Formoterol dargestellt. In den darauf folgenden Abschnitten sind die im Rahmen der Weiterentwicklung des Studienprogramms durchgeführten Studien mit der (zumindest in einem Therapiearm) zulassungskonformen zweimal täglichen Gabe gelistet. Die Tabelle schließt mit der Auflistung der Studien zur Evaluation des Inhalators).

Tabelle 4.3-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studien- bezeichnung	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
<Studie 1>	ja	abgeschlossen	12 Monate	Medikament A, Medikament B, Placebo
Studien mit einmal täglicher Einnahme von Acridiniumbromid				
M/34273/00	nein	abgeschlossen	1 Tag	Acridiniumbromid 50, 300, 600 µg QD vs. Placebo (via Cycobaler®)
M/34273/01	nein	abgeschlossen	5 Tage	Acridiniumbromid 600-6000 µg QD vs. Placebo (via Cycobaler®)
M/34273/04	nein	abgeschlossen	1 Woche	Acridiniumbromid 400 µg QD [phenyl-U- ¹⁴ C] (IV) vs. Acridiniumbromid 400 µg QD [glycolyl-U- ¹⁴ C] (IV)
M/34273/05	nein	abgeschlossen	1 Woche	Teil I: Acridiniumbromid 25- 400 µg (IV) QD vs. Placebo (IV) Teil II: Acridiniumbromid 200 µg (IV), 200 µg QD (via Genuair®)
M/34273/06	nein	abgeschlossen	5 Tage	Acridiniumbromid 200, 400, 800 µg QD vs. Placebo (via Genuair®)
M/34273/08	nein	abgeschlossen	1 Tag	Acridiniumbromid 400 µg QD (via Genuair®) bei 4 Gruppen mit normaler Nierenfunktion, sowie leichter, mittlerer und schwerer Niereninsuffizienz
M/34273/09	nein	abgeschlossen	3 Tage	Acridiniumbromid 200, 400 µg QD (via Genuair®)
M/34273/11	nein	abgeschlossen	3 Tage	Acridiniumbromid 200, 800 µg QD (via Genuair®)
M/34273/21	nein	abgeschlossen	4 Tage	Acridiniumbromid 100, 300, 900 µg QD vs. Placebo (via Genuair®)
M/34273/22	nein	abgeschlossen	4 Wochen	Acridiniumbromid 25, 50, 100, 200, 400 µg QD vs. Placebo (via Genuair®)
M/34273/24	nein	abgeschlossen	1 Tag	Acridiniumbromid 200 µg QD (via Genuair®) vs. Placebo (via Genuair®) vs. Tiotropiumbromid 18 µg QD

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studien- bezeichnung	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
				(via HandiHaler®)
M/34273/25	nein	abgeschlossen	1 Woche	Acidiniumbromid 200 µg QAM/QPM vs. Placebo (via Genuair®)
LAS-MD-26	nein	abgeschlossen	6 Wochen	Acidiniumbromid 200 µg QD vs. Placebo (via Genuair®)
M/34273/30 (ACCLAIM I)	nein	abgeschlossen	52 Wochen	Acidiniumbromid 200 µg QD vs. Placebo (via Genuair®)
M/34273/31 (ACCLAIM II)	nein	abgeschlossen	52 Wochen	Acidiniumbromid 200 µg QD vs. Placebo (via Genuair®)
Studien mit der Kombinationstherapie aus Acidiniumbromid und Formoterol				
M/273FO/23	nein	abgeschlossen	4 Wochen	Acidiniumbromid 200 µg und Formoterol 6 µg QD vs. Acidiniumbromid 200 µg und Formoterol 12 µg QD vs. Acidiniumbromid 200 µg und Formoterol 18 µg QD vs. Acidiniumbromid 200 µg vs. Formoterol 12 µg QD vs. Placebo
M/40464/26	nein	abgeschlossen	4 Wochen	Acidiniumbromid 200 µg und Formoterol 6 µg BID vs. Acidiniumbromid 200 µg und Formoterol 12 µg BID vs. Acidiniumbromid 200 µg vs. Formoterol 12 µg BID vs. Placebo
LAC-MD-27	nein	abgeschlossen	2 Wochen	Acidiniumbromid 400 µg BID und Formoterol 6 µg BID vs. Acidiniumbromid 400 µg BID und Formoterol 12 µg BID vs. Acidiniumbromid 400 µg BID vs. Formoterol 12 µg BID vs. Placebo
Studien mit zweimal täglicher Einnahme				
M/34273/23	nein	abgeschlossen	15 Tage	Acidiniumbromid 400 µg BID (via Genuair®) vs. Placebo (via Genuair®) vs. Tiotropiumbromid 18 µg QD (via HandiHaler®)
M/34273/29	nein	abgeschlossen	1 Woche	Acidiniumbromid 100, 200,

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studien- bezeichnung	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
				400 µg BID (via Genuair [®]) vs. Placebo (via Genuair [®]) vs. Formoterol 12 µg BID (via Aerolizer [®])
LAS-MD-33 (ACCORD I)	ja	abgeschlossen	12 Wochen	Acclidiniumbromid 200, 400 µg BID vs. Placebo (via Genuair [®])
M/34273/34 (ATTAIN)	ja	abgeschlossen	24 Wochen	Acclidiniumbromid 200, 400 µg BID vs. Placebo (via Genuair [®])
LAS-MD-38 A (ACCORD II)	ja	abgeschlossen	12 Wochen	Acclidiniumbromid 200, 400 µg BID vs. Placebo (via Genuair [®])
M/34273/39	nein	abgeschlossen	6 Wochen	Acclidiniumbromid 400 µg BID (via Genuair [®]) vs. Placebo (via Genuair [®]) vs. Tiotropiumbromid 18 µg QD (via HandiHaler [®])
Langzeitstudien				
LAS-MD-35	nein	abgeschlossen	52 Wochen	Acclidiniumbromid 200, 400 µg BID (via Genuair [®])
LAS-MD-36	nein	abgeschlossen	52 Wochen	Acclidiniumbromid 200, 400 µg BID (via Genuair [®])
LAS-MD-38 B	nein	abgeschlossen	40 Wochen	Acclidiniumbromid 400 µg BID (via Genuair [®])
Laufende Studien				
M/34273/40	nein	laufend	3 Wochen	Acclidiniumbromid 400 µg BID vs. Placebo (via Genuair [®])
Device Studien				
M/34273/03	nein	abgeschlossen	1 Tag	Acclidiniumbromid 200 µg (via Genuair [®])
M/34273/07	nein	abgeschlossen	1 Tag	Placebo (via Genuair [®]) vs. Placebo (via HandiHaler [®])
M/34273/32	nein	abgeschlossen	1 Tag	Placebo (via Genuair [®]) vs. Placebo (via HandiHaler [®])
QD = einmal täglich, BID = zweimal täglich, QAM = einmal morgens, QPM = einmal abends, IV = intravenös, µg = Mikrogramm				

Geben Sie an, welchen Stand die Information in Tabelle 4.3-1 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus sollte nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Informationen in Tabelle 4-9 stellen den Stand vom 09.07.2012 dar.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4.3-1 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

In der Studienübersicht des pU werden keine den sonstigen Einschlusskriterien entsprechenden Studien mit einer empfohlenen Studiendauer von mindestens 12 Wochen gefunden. Deshalb werden, wie in Abschnitt 4.2.2 geplant, auch kürzere Studien in die Nutzenbewertung mit Hilfe des direkten Vergleiches eingeschlossen.

Die folgende Tabelle 4.3-2 zeigt die Studien des pharmazeutischen Unternehmers, die nicht für den direkten Vergleich berücksichtigt wurden. Bei der Begründung der Nichtberücksichtigung der Studie ist bei Zutreffen mehrerer Ausschlussgründe lediglich der primäre Grund aufgeführt.

Tabelle 4.3-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel für den direkten Vergleich

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
M/34273/00	Population (es handelt sich um männliche gesunde Freiwillige)
M/34273/01	Population (es handelt sich um männliche gesunde Freiwillige)
M/34273/03	Prüf- und Vergleichstherapie (es handelt sich um eine Device Studie)
M/34273/04	Population (es handelt sich um männliche gesunde Freiwillige)
M/34273/05	Population (es handelt sich um männliche gesunde Freiwillige)
M/34273/06	Population (es handelt sich um gesunde Freiwillige)
M/34273/07	Prüf- und Vergleichstherapie (es handelt sich um eine Device Studie)
M/34273/08	Vergleichstherapie (Ermittlung der Dosis toleranz bei Niereninsuffizienz)
M/34273/09	Vergleichstherapie (es handelt sich um eine Dosisfindungsstudie)
M/34273/11	Population (es handelt sich um gesunde Freiwillige)
LAS-PK-12	Population (es handelt sich um gesunde Freiwillige)
M/34273/21	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/34273/22	Vergleichstherapie (Placebo statt Tiotropiumbromid)
M/273FO/23	Vergleichstherapie (Placebo oder Formoterol statt Tiotropiumbromid)
M/34273/24	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/34273/25	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/40464/26	Vergleichstherapie (Placebo oder Formoterol statt Tiotropiumbromid)
LAS-MD-26	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
LAC-MD-27	Vergleichstherapie (Placebo oder Formoterol statt Tiotropiumbromid)
M/34273/29	Vergleichstherapie (Placebo und Formoterol statt Tiotropiumbromid)
M/34273/30	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/34273/31	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/34273/32	Prüf- und Vergleichstherapie (es handelt sich um eine Device Studie)
LAS-MD-33 (ACCORD I)	Vergleichstherapie (Placebo statt Tiotropiumbromid)
M/34273/34 (ATTAIN)	Vergleichstherapie (Placebo statt Tiotropiumbromid)
LAS-MD-35	Vergleichstherapie (Studie ist nicht kontrolliert)
LAS-MD-36	Vergleichstherapie (Studie ist nicht kontrolliert)
LAS-MD-38 A (ACCORD II A)	Vergleichstherapie (Placebo statt Tiotropiumbromid)
LAS-MD-38 B (ACCORD II B)	Vergleichstherapie (Studie ist nicht kontrolliert)

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
M/34273/40	Laufende Studie - es liegen noch keine Ergebnisse vor
QD = einmal täglich, µg = Mikrogramm	

Nach Ausschluss der oben gelisteten Studien verbleiben 2 Studien im Studienpool für den direkten Vergleich. Dabei handelt es sich um die Studien M/34273/23 und M/34273/39.

Stellen Sie die abgebrochenen und laufenden Studien, die nicht in die Bewertung eingehen, in Anhang 4-D und Anhang 4-E dar.

4.3.1.1.2 Studien aus der bibliographischen Literaturrecherche

Sofern eine bibliografische Literaturrecherche durchgeführt wurde, beschreiben Sie nachfolgend das Ergebnis dieser Recherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-C.

Falls die Recherche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel gemeinsam mit der Recherche zu einer anderen Fragestellung (z. B. zu indirekten Vergleichen) durchgeführt wurde, unterteilen Sie die Angaben zu relevanten Treffern und Studien entsprechend diesen Fragestellungen.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.

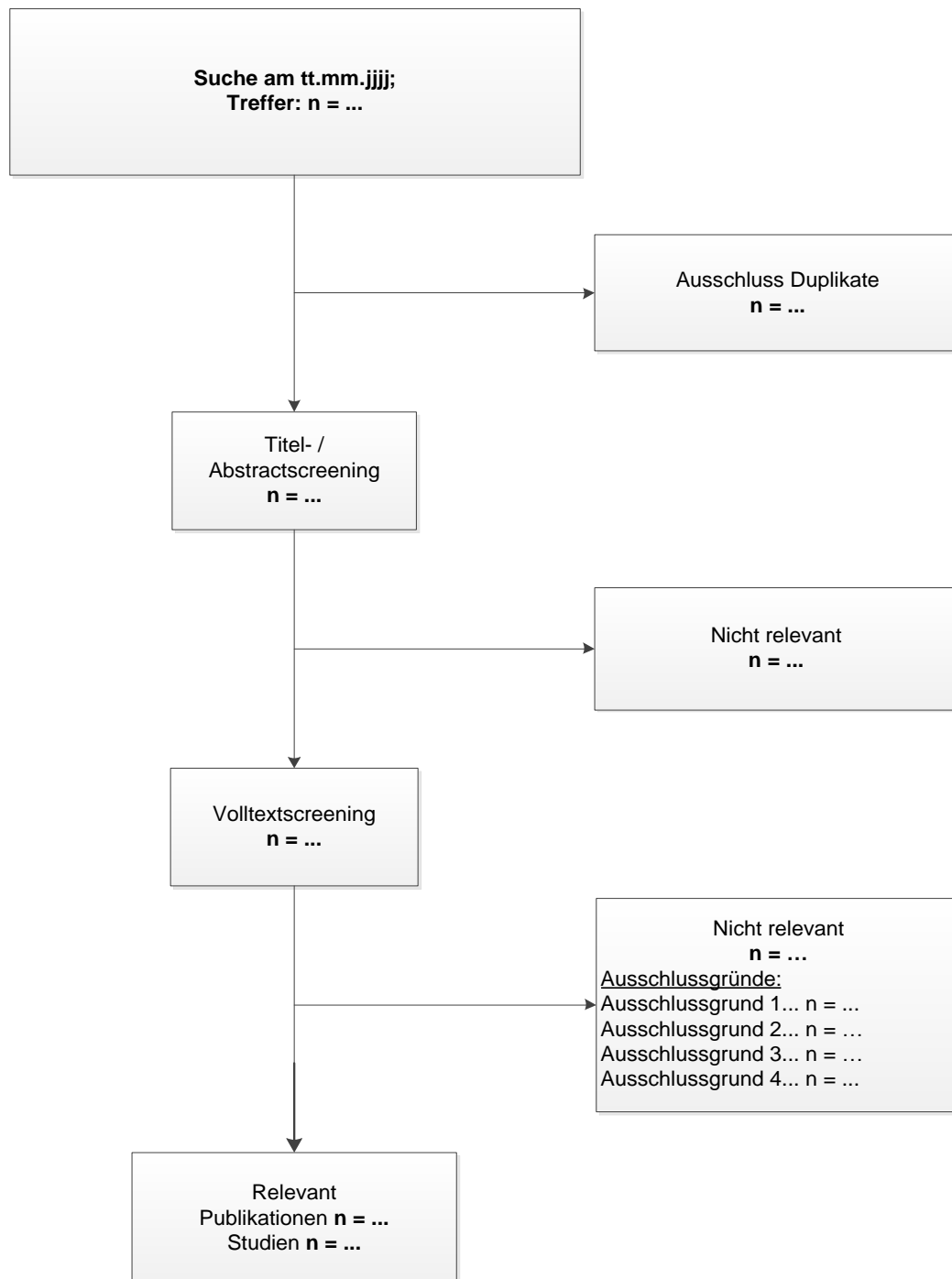


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die bibliografische Literaturrecherche wurde in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt. Es wurden RCTs mit einem direkten Vergleich des zu bewerteten Arzneimittels mit der Vergleichstherapie gesucht. Die Suchstrategie ist in Anhang 4-B dargestellt.

In MEDLINE wurden 3, in EMBASE 3 und in den Cochrane-Datenbanken 6 Treffer erzielt (vgl. Abbildung 2). Infolge des Dubletten-Checks wurden 4 Ergebnisse entfernt.

Nach dem Titel-/Abstractscreening wurden entsprechend der Ausschlusskriterien (Abschnitt 4.2.2) 5 Treffer ausgeschlossen:

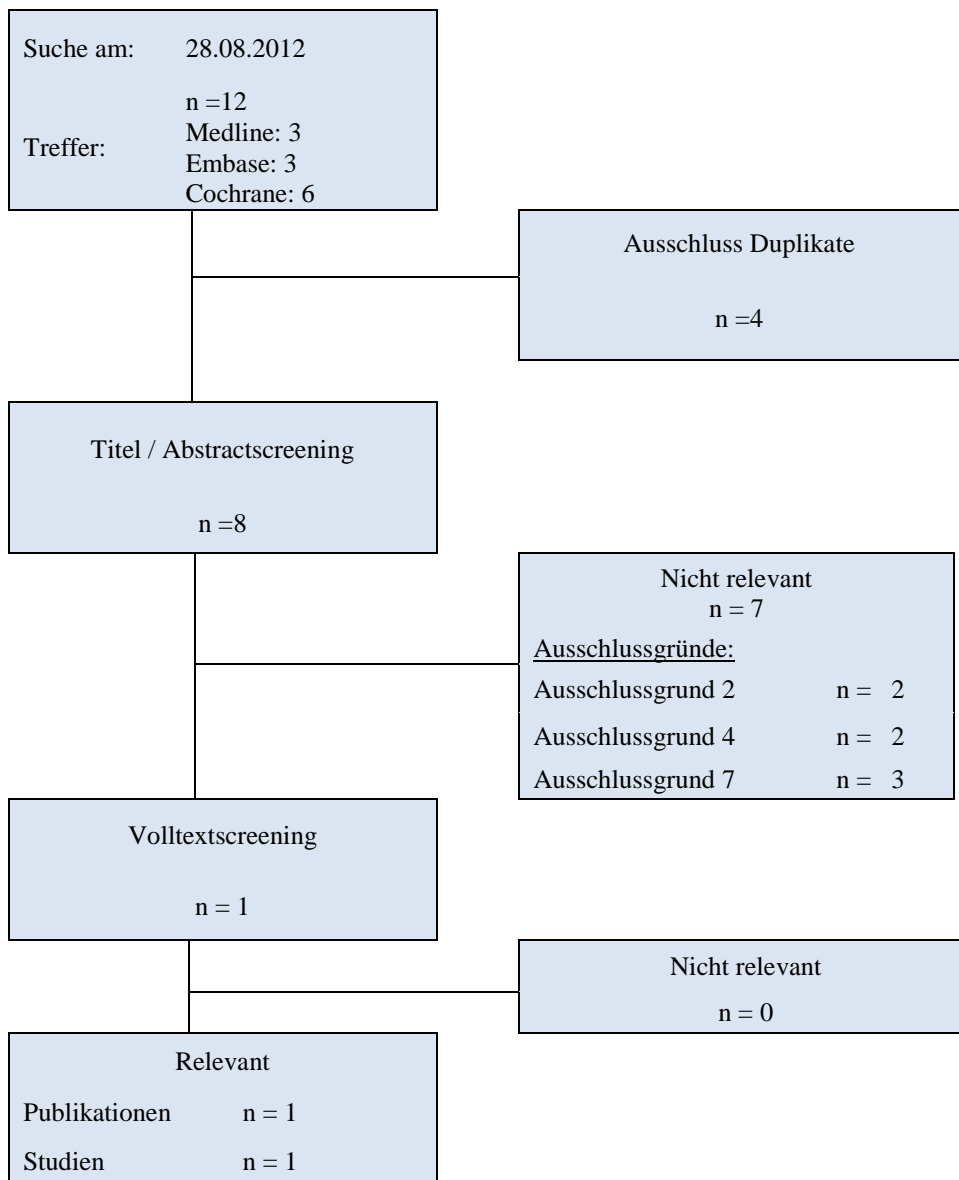


Abbildung 2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach RCTs für direkten Vergleich

In der bibliografischen Literaturrecherche nach RCTs mit einem direkten Vergleich des zu bewerteten Arzneimittels mit der Vergleichstherapie wurde eine Studie identifiziert. Dabei

handelt es sich um die Studie des pU M/34273/23, die von Fuhr et al. [31] im März 2012 publiziert wurde.

4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4.3-1) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde (sofern eine solche durchgeführt wurde). Fügen Sie für jede Studie eine neue Zeile ein.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Die folgende Tabelle 4.3-3 zeigt die durch die Suche in Studienregistern identifizierten Studien des zu bewertenden Arzneimittels.

Von den vier als relevant eingestuften Studien sind zwei (M/34273/40 und DW_EKL001) noch nicht abgeschlossen. Die Studien M/34273/23 und M/34273/39 werden in den Studienpool aufgenommen.

Tabelle 4.3-3: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d. ^b)
<Studie 1>	clinicaltrials.gov [Zitate], clinicalstudyresults.org [Zitate]	ja	nein
M/34273/23	clinicaltrials.gov und ICTPR [NCT00868231]	ja	ja [31]
M/34273/39	clinicaltrials.gov und ICTPR [NCT01462929]	ja	nein
M/34273/40	ICTPR [EUCTR2011-002665-38-GB, EUCTR2011-002665-38-ES]	ja	nein
DW_EKL001	ICTPR [NCT01636401]	nein	nein
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.			

4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle für die Nutzenbewertung berücksichtigten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

In der Informationsbeschaffung wurden keine Studien mit der empfohlenen Dauer von 12 Wochen gefunden. Deshalb werden, wie in Abschnitt 4.2.2 geplant, auch kürzere Studien in die Nutzenbewertung mit Hilfe des direkten Vergleiches einbezogen. Für die Bewertung des Zusatznutzens werden insbesondere solche Endpunkte herangezogen, die den schnell eintretenden Erfolg der Therapie sichtbar machen. Dies sind: COPD-Symptome, wie Husten, Auswurf, Atemnot, nächtliche und morgendliche Symptome sowie

Lungenfunktionsparameter. Alle übrigen in den Studien berichteten patientenrelevanten Endpunkte werden zum Zwecke der Vollständigkeit ebenfalls dargestellt.

Tabelle 4.3-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
Placebokontrolliert						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
M/34273 /23	Nein	ja	nein	ja [32]	ja [NCT00868231]	ja [31]
M/34273 /39	Nein	ja	nein	ja [33]	ja [NCT01462929]	nein
a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.						
b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.						
c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-F zu hinterlegen.

Tabelle 4.3-5: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
<Studie 1>	<i>RCT, doppelblind, parallel</i>	<i>Jugendliche und Erwachsene, leichtes bis mittelschweres Asthma</i>	<Gruppe 1> (n=354) <Gruppe 2> (n=347)	<i>Run-in: 2 Wochen Behandlung: 6 Monate</i>	<i>Europa (Deutschland, Frankreich, Polen) 9/2003 – 12/2004</i>	<i>FEV₁; Asthma-Symptome, Exazerbationen, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse</i>
M/34273/23	multizentrische Mehrfachdosis- Studie, doppelblind, double-dummy, randomisiert, 3- Perioden-cross-over, placebo- und aktivkontrolliert	Patienten ≥40 Jahren, mittelschwere bis schwere COPD, eine Raucherhistorie ≥ 10 Packungsjahren. Nach Gabe eines Bronchodilatators FEV ₁ /FVC <70% FEV ₁ ≥30% bis <80% des Sollwerts	Acidiniumbromid 400 µg (n=30); davon relevante Patienten ^a : n=29 Tiotropiumbromid 18 µg (n=30); davon relevante Patienten ^a : n=28 Placebo (n=30); davon relevante Patienten ^a : n=30	Run-in: 5-9 Tage Behandlung: 15 Tage* Follow-up: 4-6 Tage	Deutschland (2 Zentren) 03/2009 – 07/2009	Primärer: FEV ₁ AUC ₀₋₁₂ Änderung gegenüber dem Baselinewert (CFB) zum Studienende; Sekundäre: CFB FEV ₁ AUC ₀₋₁₂ am 1. Tag der Behandlung; CFB FVC AUC ₀₋₁₂ am 1. und 15. Tag der Behandlung; CFB FEV ₁ und FVC AUC ₀₋₂₄ am 1. und 15. Tag der Behandlung • CFB FEV ₁ und FVC AUC ₁₂₋₂₄ am 1. und 15. Tag der Behandlung • CFB Trough-FEV ₁ - und FVC Wert am

						<p>Morgen und am Abend am 1. und 15. Tag der Behandlung</p> <ul style="list-style-type: none"> • CFB Peak-FEV₁- und FVC Wert am Morgen und am Abend am 1. und 15. Tag der Behandlung • CFB FEV₁- und FVC Wert zu jeder spezifischen Tageszeit • CFB der begleitenden Medikation nach ein und zwei Wochen der Behandlung <p>CFB COPD-Symptome nach ein und zwei Wochen der Behandlung</p>
M/34273/39	multizentrische Mehrfachdosis-Studie, doppelblind, double-dummy, randomisiert, parallel, placebo- und aktivkontrolliert	<p>Patienten ≥ 40 Jahren, mittelschwere bis schwere COPD, eine Raucherhistorie ≥ 10 Packungsjahren.</p> <p>Nach Gabe eines Bronchodilatators FEV₁/FVC <70% FEV₁ $\geq 30\%$ bis <80% des Sollwerts</p>	<p>Acclidiniumbromid 400 μg (n=171); davon relevante Patienten^a: n=171</p> <p>Tiotropiumbromid 18 μg (n=158); davon relevante Patienten^a: n=158</p> <p>Placebo (n=85); davon relevante Patienten^a: n=85</p>	<p>Run-in: 2-3 Wochen Behandlung: 6 Wochen Follow-up: 2 Wochen</p>	<p>Europa (60 Zentren) 11/2011 - 06/2012</p>	<p>Primärer: FEV₁ AUC₀₋₂₄ Änderung gegenüber dem Baselinewert (CFB) zum Studienende;</p> <p>Sekundäre: CFB FEV₁ AUC₀₋₂₄ zum am ersten Tag</p>

a: Intention to Treat (ITT) population, *30 Patienten werden nacheinander in allen drei Therapiearmen á 15 Tage behandelt

Tabelle 4.3-6: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	<Gruppe 1>	<Gruppe 2>	<Gruppe 3>	Begleitende Medikation
<Studie 1>	xxx 250 µg, 1 Inhalation bid + Placebo 2 Inhalationen bid	yyy 200 µg, 2 Inhalationen bid + Placebo 1 Inhalation bid		Vorbehandlung: zzz 1000 µg pro Tag, 4 Wochen vor Studienbeginn Bedarfsmedikation: aaa
M/34273/23	Aclidinium- bromid 400 µg, 2 Inhalationen bid (morgens/abends, via Genuair®) + Placebo zu Tiotropium- bromid Inhalation od (morgens via HandiHaler®)	Tiotropiumbromid 18 µg, Inhalation od (morgens, via HandiHaler®) + Placebo zu Aclidiniumbromid 2 Inhalationen bid (morgens/abends via Genuair®)	Placebo zu Aclidiniumbromid 2 Inhalationen bid (morgens/abends via Genuair®) + Placebo zu Tiotropiumbromid Inhalation od (morgens via HandiHaler®)	<u>Bedarfsmedikation:</u> Salbutamol als zugelassenes atemwegserweiterndes Medikament, unterbrochen ≥6 h vor der Untersuchung <u>Begleitende Medikation:</u> Inhalative Corticosteroide (ICS) oder orale oder parenterale Corticosteroide, dauerhafte Nutzung ≥ 4 Wochen vor Run- In; Medikationen zu anderen Erkrankungen (z.B. Diuretika, Antidepressiva), dauerhafte Nutzung ≥ 4 Wochen vor Run- In; Sauerstofftherapie, dauerhafte Nutzung ≥ 4 Wochen vor Run- In (≤15 h/ täglich)
M/34273/39	Aclidinium- bromid 400 µg, 2 Inhalationen bid (morgens/abends, via Genuair®) + Placebo zu Tiotropium- bromid Inhalation od (morgens via HandiHaler®)	Tiotropium- bromid 18 µg, Inhalation od (morgens, via HandiHaler®) + Placebo zu Aclidiniumbromid 2 Inhalationen bid (morgens/abends via Genuair®)	Placebo zu Aclidiniumbromid 2 Inhalationen bid (morgens/abends via Genuair®) + Placebo zu Tiotropiumbromid Inhalation od (morgens via HandiHaler®)	<u>Bedarfsmedikation:</u> Salbutamol als zugelassenes atemwegserweiterndes Medikament, unterbrochen ≥6 h vor der Untersuchung <u>Begleitende Medikation:</u> Inhalative Corticosteroide (ICS), orale oder parenterale Corticosteroide oder orale Theophylline, dauerhafte Nutzung ≥ 4 Wochen vor Run- In;

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4.3-7: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel

Studie Gruppe	N ^a	Alter (Jahre) ^b	Geschlecht w /m (%)	Schweregrad ^c (mittel / schwer)%	Rauchen [Packungs- jahre] ^b	Zeit seit COPD- Diagnose [Jahre] ^b	FEV ₁ ^d zu Studienbeginn [%Sollwert] ^b	FEV ₁ ^d zu Studienbeginn [% FVC] ^b	Studienabbruch n (%)
<Studie 1>									
<Gruppe 1>									
<Gruppe 2>									
placebokontrollierte Studien mit dem zu bewertenden Arzneimittel									
M/34273/23	30	58,37 ± 7,9	36,7 / 63,3	63,3 / 33,3	41,1 (15,9)	9,2 (6,9)	55,7 (13,7)	46,2 (10,3)	3 (10,0) ^e
M/34273/39	414	62,3 ± 8,1	32,9 / 67,2	65,4 / 34,6	42,4 (20,9)	8,7 (6,1)	55,8 (12,9)	48,1 (11,2)	14 (3,4)
AB 400 µg	171	61,8 ± 8,2	33,3 / 66,7	63,2 / 36,8	41,0 (22,4)	8,8 (5,9)	55,8 (13,3)	47,6 (11,5)	5 (2,9)
TIO 18 µg	158	62,8 ± 7,9	26,6 / 73,4	66,2 / 33,8	45,0 (21,8)	8,2 (6,0)	56,0 (13,2)	48,6 (11,1)	4 (2,5)
Placebo	85	62,2 ± 8,2	43,5 / 56,5	68,2 / 31,8	39,6 (15,4)	9,6 (6,7)	55,5 (11,8)	48,4 (10,9)	5 (5,9)

a: Zahl der Patienten, für die in den Studien Charakteristika zu Studienbeginn dargestellt wurden (Sicherheits-Population; alle randomisierten und behandelten Patienten).

b: Mittelwert und Standardabweichung

c: Nach GOLD Leitlinie 2006 [34]; M/34273/23 enthält einen Patienten mit sehr schwerer COPD; die Summe der Anteile ergibt deshalb nicht 100%

d: Postbronchodilatatorisch

e: Während der Placebobehandlung aufgrund von AEs

Beschreiben Sie die Studien zusammenfassend. Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

Die in Tabelle 4.3-5 dargestellten multizentrischen, doppelblinden, double-dummy, randomisierten placebo- und aktivkontrollierten Mehrfachdosis-Studien unterscheiden sich im Wesentlichen durch die Studiendauer (15 Tage vs. 6 Wochen) und das Studiendesign. Die Studie M/34273/39 ist parallel, die Studie M/34273/23 im cross-over-Design angelegt. Die primären und sekundären Endpunkte fokussieren bei beiden Studien die COPD-Symptome (Lungenfunktionsparameter).

Tabelle 4.3-6 beschreibt die durch das double-dummy-Design der beiden Studien charakterisierte Medikation der drei Therapiearme. Die zugelassene Bedarfsmedikation und die begleitende Medikation sollen die Minimierung von Exazerbationen unterstützen.

Tabelle 4.3-7 zeigt die Charakteristika der in den 2 RCTs mit direktem Vergleich untersuchten Patienten. Aufgrund des cross-over Designs der Studie M/34273/23 werden die Patientencharakteristika nur für die gesamte Studienpopulation dargestellt. Bei der LAS-MD-39 werden sowohl die Charakteristika der Gesamtpopulation als auch die der drei Therapiearme dargestellt. Die Patientenpopulationen unterscheiden sich in den dargestellten Kriterien kaum voneinander. Lediglich die Studienabbruchrate ist in der M/34273/23 höher als in der LAS-MD-39. Bei beiden Patientenpopulationen handelt sich um vortherapierte Patienten mit einer Auswaschungsphase von 9-15 Tagen vor jeder Behandlungsperiode (M/34273/23) bzw. von 2-3 Wochen vor Studienbeginn (M/34273/39).

4.3.1.2.2 Verzerrungspotential auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-8: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	Ja	ja	ja	ja	nein	nein	niedrig
M/34273/39	Ja	ja	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial der beiden RCTs mit direktem Vergleich (M/34273/23 und M/34273/39) wird als niedrig eingestuft. Die Randomisierung, Verblindung und Berichterstattung geben keinen Hinweis auf Verzerrung.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-9: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität	Morbidität	Lebensqualität	Unerwünschte Arzneimittelwirkungen
M/34273/23	Ja	ja	ja	ja
M/34273/39	Ja	ja	ja	ja

Die in Tabelle 4.3-9 dargestellten und zur Nutzenbewertung herangezogenen Endpunktkategorien beruhen auf folgenden Ausprägungen (Tabelle 4.3-10):

Tabelle 4.3-10: Im indirekten Vergleich herangezogene Endpunkte

Endpunktkategorie und Ausprägung	M/34273/23	M/34273/39
Mortalität		
COPD-assoziierte kardiovaskuläre Morbidität und Mortalität	ja	ja
COPD-bedingte Letalität	ja	ja
Gesamtmortalität	ja	ja
Morbidität (COPD-Symptome)		
Husten, Auswurf, Atemnot	ja	ja
Nächtliche und morgendliche Symptome	ja	ja
Dyspnoe	nein	nein
Exazerbationen	nein	nein
Morgendlicher Trough-FEV ₁ -Wert	ja	ja
Peak-FEV ₁ -Wert	ja	ja
FEV ₁ AUC	ja	ja
Gesundheitsbezogene Lebensqualität		
Lebensqualität	nein	nein
Behandlungszufriedenheit	ja	ja
Unerwünschte Arzneimittelwirkungen		
unerwünschte Ereignisse (AEs)	ja	ja
schwerwiegende unerwünschte Ereignisse (SAEs)	ja	ja
Studienabbruch wegen unerwünschter Ereignisse	ja	ja

4.3.1.3.1 Endpunkt Morbidität – RCT

4.3.1.3.1.1 CFB FEV₁ AUC – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-11: Operationalisierung von CFB FEV₁ AUC

Studie	Operationalisierung
M/34273/23	<p>Das forcierte Ein-Sekunden-Volumen (FEV₁) wird mit Hilfe eines Spirometers ermittelt. Dazu müssen die Patienten in ein Mundstück aus voller Lunge schnell ausatmen. Gemessen wird die Luftmenge (in Litern), die in der ersten Sekunde ausgeatmet wird.</p> <p>Zur Berechnung der Fläche unter der Kurve (Area under the curve, AUC) werden 19 FEV₁ Werte über den Tag verteilt ermittelt und als Datenpunkte in ein Zeit-FEV₁-Diagramm eingetragen.</p> <p>FEV₁ AUC₀₋₂₄ beschreibt die Entwicklung des FEV₁ Wertes über 24 Stunden, FEV₁ AUC₀₋₁₂ in der ersten Tageshälfte (9:00 bis 21:00 Uhr, jeweils ± 1 Stunde) und FEV₁ AUC₁₂₋₂₄ in der zweiten Tageshälfte (21:00 Uhr bis 9:00 Uhr, jeweils ± 1 Stunde).</p> <p>Zur Berechnung der Veränderung gegenüber des Baseline-Wertes (CFB) werden die jeweiligen FEV₁ AUC Werte an Tag 1 und Tag 15 den FEV₁ AUC Werten vor Therapiebeginn gegenübergestellt.</p> <p>Die Baseline Werte werden als Durchschnitt zweier Messungen (eine und null Stunden vor Therapiebeginn) vor jeder der drei Behandlungsperioden der drei-armigen cross-over-Studie ermittelt.</p>
M/34273/39	<p>Das forcierte Ein-Sekunden-Volumen (FEV₁) wird mit Hilfe eines Spirometers ermittelt. Dazu müssen die Patienten in ein Mundstück aus voller Lunge schnell ausatmen. Gemessen wird die Luftmenge (in Litern), die in der ersten Sekunde ausgeatmet wird.</p> <p>Zur Berechnung der Fläche unter der Kurve (Area under the curve, AUC) werden 19 FEV₁ Werte über den Tag verteilt ermittelt und als Datenpunkte in ein Zeit-FEV₁-Diagramm eingetragen.</p> <p>FEV₁ AUC₀₋₂₄ beschreibt die Entwicklung des FEV₁ Wertes über 24 Stunden, FEV₁ AUC₀₋₁₂ in der ersten Tageshälfte (9:00 bis 21:00 Uhr, jeweils ± 1 Stunde) und FEV₁ AUC₁₂₋₂₄ in der zweiten Tageshälfte 21:00 Uhr bis 9:00 Uhr, jeweils ± 1 Stunde).</p> <p>Zur Berechnung der Veränderung gegenüber des Baseline-Wertes (CFB) werden die jeweiligen FEV₁ AUC Werte an Tag 1 und nach 6 Wochen den FEV₁ AUC Werten vor Therapiebeginn gegenübergestellt.</p> <p>Die Baseline Werte werden als Durchschnitt zweier Messungen (eine und null Stunden vor Therapiebeginn) vor Behandlungsbeginn ermittelt.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-12: Bewertung des Verzerrungspotenzials für CFB FEV₁ AUC in RCT mit direktem Vergleich

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	niedrig	ja	ja	nein	nein	niedrig
M/34273/39	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential der beiden RCTs mit direktem Vergleich (M/34273/23 und M/34273/39) ist niedrig. Das Verzerrungspotential des erhobenen Endpunktes wird in beiden Studien ebenfalls als niedrig eingestuft; die Endpunkterheber sind verblindet, die Auswertung der Ergebnisse erfolgt nach dem ITT-Prinzip, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-13: Ergebnisse für CFB FEV₁ AUC (l) aus RCT mit direktem Vergleich

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	Gruppendifferenz ¹ (SE) KI [] p-Wert
		Mean (SD)	Mean (SD)	CFB (SD)	
M/34273/23					
FEV₁ AUC₀₋₁₂					
Tag 1*	AB (n=29)	1,46 (0,50)	1,70 (0,54)	0,24 (0,20)	0,05 (0,02) KI [0,02; 0,10] p=0,04
	TB (n=28)	1,49 (0,47)	1,67 (0,53)	0,18 (0,20)	
Tag 15	AB (n=29)	1,46 (0,50)	1,70 (0,56)	0,24 (0,35)	-0,02 (0,04) KI [-0,11; 0,06] p=0,57
	TB (n=28)	1,49 (0,47)	1,74 (0,59)	0,25 (0,25)	
FEV₁ AUC₁₂₋₂₄					
Tag 1*	AB (n=28)	1,46 (0,50)	1,63 (0,53)	0,18 (0,27)	0,10 (0,03) KI [0,04; 0,16] p=0,002
	TB (n=27)	1,49 (0,47)	1,60 (0,54)	0,09 (0,25)	
Tag 15*	AB (n=28)	1,46 (0,50)	1,63 (0,55)	0,20 (0,28)	0,07 (0,03) KI [0,01; 0,14] p=0,02
	TB (n=28)	1,49 (0,47)	1,57 (0,54)	0,08 (0,25)	
FEV₁ AUC₀₋₂₄					
Tag 1*	AB (n=29)	1,46 (0,50)	1,68 (0,54)	0,21 (0,23)	0,07 (0,03) KI [0,02; 0,13] p=0,008
	TB (n=28)	1,49 (0,47)	1,63 (0,53)	0,13 (0,22)	
Tag 15	AB (n=28)	1,46 (0,50)	1,67 (0,56)	0,24 (0,27)	0,04 (0,03) KI [-0,1; 0,11] p=0,10
	TB (n=28)	1,49 (0,47)	1,66 (0,56)	0,16 (0,24)	
M/34273/39					
FEV₁ AUC₀₋₁₂					
Tag 1	AB (n=171)	1,46 (0,48)	1,60 (0,53)	0,14 (0,18)	0,01 (0,02) KI [-0,02; 0,05] p=0,50
	TB (n=158)	1,54 (0,54)	1,67 (0,57)	0,13 (0,17)	
6 Wochen	AB (n=171)	1,46 (0,48)	1,58 (0,53)	0,12 (0,22)	-0,02 (0,03) KI [-0,07; 0,03] p=0,48
	TB (n=158)	1,54 (0,54)	1,67 (0,57)	0,13 (0,24)	

FEV₁ AUC₁₂₋₂₄					
Tag 1*	AB (n=171)	1,46 (0,48)	1,56 (0,52)	0,09 (0,21)	0,07 (0,02)
	TB (n=158)	1,54 (0,54)	1,56 (0,55)	0,02 (0,20)	KI [0,03; 0,11] p=0,002
6 Wochen	AB (n=171)	1,46 (0,48)	1,51 (0,50)	0,05 (0,20)	0,04 (0,02)
	TB (n=158)	1,54 (0,54)	1,54 (0,55)	0,001 (0,24)	KI [-0,01; 0,08] p=0,12
FEV₁ AUC₀₋₂₄					
Tag 1*	AB (n=171)	1,46 (0,48)	1,58 (0,52)	0,12 (0,19)	0,04 (0,02)
	TB (n=158)	1,54 (0,54)	1,62 (0,56)	0,07 (0,17)	KI [0,002; 0,08] p=0,04
6 Wochen	AB (n=171)	1,46 (0,48)	1,55 (0,51)	0,08 (0,20)	0,01 (0,02)
	TB (n=158)	1,54 (0,54)	1,61 (0,56)	0,06 (0,23)	KI [-0,04; 0,06] p=0,67
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * signifikanter Gruppenunterschied AB: Acclidiniumbromid, TB: Tiotropiumbromid, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall					

M/34273/23

Die Veränderung FEV₁ AUC₀₋₁₂ zur Baseline (CFB) an Tag 15 der Behandlung stellt den primären Endpunkt der Studie M/34273/23 dar. Der CFB-Wert FEV₁ AUC₀₋₁₂ am ersten Tag, sowie die Werte der zweiten Tageshälfte (FEV₁ AUC₁₂₋₂₄) und der Gesamtbetrachtung (FEV₁ AUC₀₋₂₄) stellen mit jeweils 2 Erhebungszeitpunkten (Tag 1 und Tag 15) die sekundären Endpunkte dar. Die FEV₁-Werte der Vergleichsgruppen (Acclidiniumbromid und Tiotropiumbromid) unterscheiden sich bei der Baselineerhebung mit 1,46 und 1,49 l kaum voneinander.

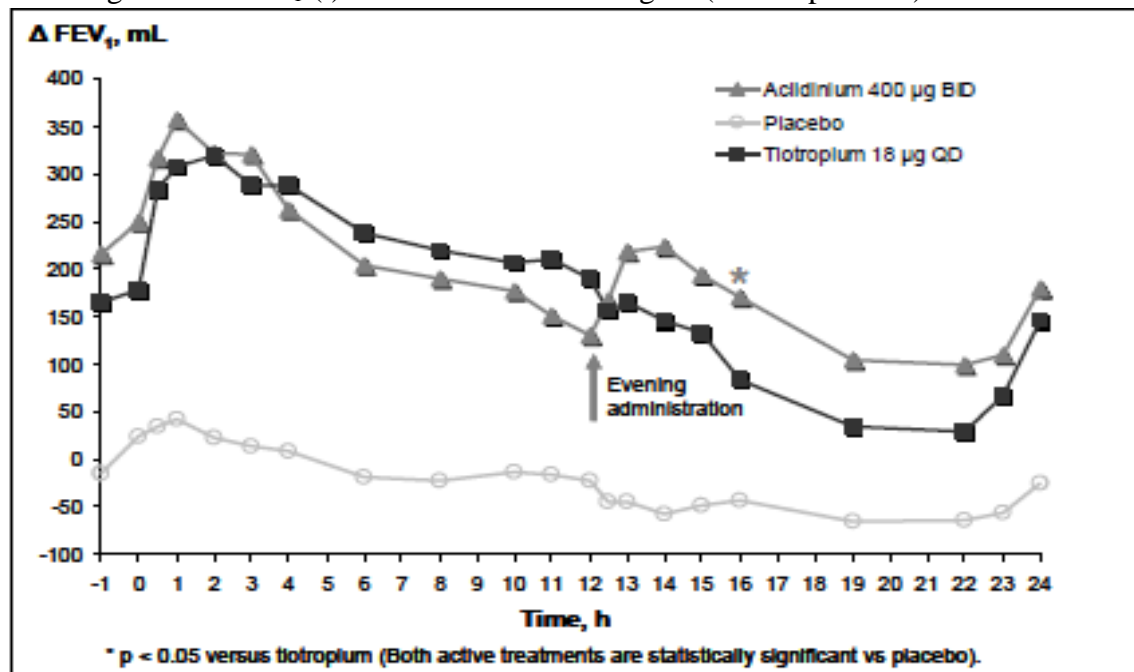
Der Unterschied von 0,03 l zwischen den beiden Therapiearmen in der Baseline Untersuchung schreibt sich beim primären Endpunkt über 15 Tage proportional fort. Nach 15 Tagen weisen beide Therapiearme eine vergleichbare, klinisch relevante [9]¹⁹ Verbesserung des FEV₁ AUC₀₋₁₂ um 0,24 bzw. 0,25 l auf. Es besteht keine Signifikanz zwischen den Gruppen (p=0,57).

Die weiteren - als sekundäre Endpunkte erhobene - FEV₁ AUC Werte an Tag 15 und Tag 1 weisen im Acclidiniumbromidarm in allen Fällen eine klinisch relevante Verbesserung gegenüber der Baseline Untersuchung auf. Anders verhält es sich im Tiotropiumbromidarm: in der durch den FEV₁ AUC₁₂₋₂₄ Wert beschriebenen zweiten Tageshälfte ist keine klinisch relevante Veränderung zu erkennen. Die Veränderung von unter 100 ml führt im Vergleich mit 200 ml bei Acclidiniumbromid zu einem klinisch bedeutenden und signifikanten (p=0,02) Vorteil von Acclidiniumbromid in der zweiten Tageshälfte.

¹⁹ Donohue 2004 definiert eine Änderung des EFV₁ Wertes um 100 ml als klinisch relevant.

Zur Veranschaulichung des Abfalls der Lungenfunktion in der zweiten Tageshälfte (ab Stunde 3 nach Inhalation) unter Tiotropiumbromid und des erneuten Anstiegs unter Acclidiniumbromid (ab Stunde 12 nach Inhalation) sind in der folgenden Abbildung 3 die FEV₁ -Werte über 24 Stunden dargestellt. Da der Endpunkt FEV₁ AUC aus den in Abbildung 3 abgetragenen FEV₁ -Werte über den Tagesverlauf ermittelt wird, wird diese Abbildung unterstützend herangezogen.

Abbildung 3: CFB FEV₁ (l) über 24 Stunden an Tag 15 (ITT-Population)



Quelle: Studienbericht zur M/34273/23 [32]

Weitere signifikante Vorteile von Acclidiniumbromid gegenüber Placebo zeigen sich in zu allen Messzeitpunkten (FEV₁ AUC₀₋₁₂, ₁₂₋₂₄, ₀₋₂₄) in der Messung am ersten Tag (p=0,04 bzw. 0,002 bzw. 0,008).

In der Betrachtung über 24 h erzielen die Patienten der Acclidiniumbromidtherapie nach 15 Tagen gegenüber der Tiotropiumbromidtherapie eine bessere Bronchodilatation, die nicht signifikant ist (p=0,10).

Zusammenfassend sorgt Acclidiniumbromid zu allen Messzeitpunkten (FEV₁ AUC₀₋₁₂, ₁₂₋₂₄, ₀₋₂₄) vom ersten (Tag 1) bis zum letzte Tag (Tag 15) der Studie für eine klinisch relevante Verbesserung der Lungenfunktion gegenüber Placebo. Hervorzuheben ist die statistisch signifikante Überlegenheit von Acclidiniumbromid gegenüber Tiotropiumbromid in der zweiten Tageshälfte (Tag 15 FEV₁ AUC₁₂₋₂₄). Die FEV₁ Werte (Tag 15) der ersten Tageshälfte sind zwischen der Therapien vergleichbar, die Betrachtung über 24 Stunden lässt einen Vorteil von Acclidiniumbromid erkennen, der jedoch nicht signifikant (p=0,1) ist.

M/34273/39

Die Veränderung $FEV_1 AUC_{0-24}$ zur Baseline (CFB) nach 6 Wochen der Behandlung stellt den primären Endpunkt der Studie M/34273/39 dar. Der CFB-Wert $FEV_1 AUC_{0-24}$ am ersten Tag, sowie die Werte der ersten ($FEV_1 AUC_{0-12}$) und zweiten Tageshälfte ($FEV_1 AUC_{12-24}$) stellen mit jeweils 2 Erhebungszeitpunkten (Tag 1 und Woche 6) die sekundären Endpunkte dar. Die FEV_1 -Werte der Vergleichsgruppen (Aclidiniumbromid und Tiotropiumbromid) unterscheiden sich bei der Baselineerhebung mit 1,46 und 1,54 l etwas voneinander.

Der Unterschied von 0,08 l zwischen den beiden Therapiearmen in der Baseline Untersuchung schreibt sich beim primären Endpunkt über 6 Wochen proportional fort. Nach 6 Wochen weisen beide Therapiearme eine vergleichbare Verbesserung des $FEV_1 AUC_{0-24}$ um 0,08 bzw. 0,06 l auf. Es besteht keine Signifikanz zwischen den Gruppen ($p=0,67$).

Die weiteren - als sekundäre Endpunkte erhobene - $FEV_1 AUC$ Werte in Woche 6 und an Tag 1 weisen im Aclidiniumbromidarm in 50% der Erhebungszeitpunkte eine klinisch relevante Verbesserung gegenüber der Baseline Untersuchung auf. Anders verhält es sich im Tiotropiumbromidarm: Dort kann nur in der ersten Tageshälfte eine klinisch relevante Verbesserung festgestellt werden. In der zweiten Tageshälfte ($FEV_1 AUC_{12-24}$) und über den ganzen Tag ($FEV_1 AUC_{0-24}$) betrachtet ist die Veränderung im Aclidiniumbromidarm am ersten Tag signifikant ($p = 0,002$ bzw. $0,04$) höher als im Tiotropiumbromidarm. Nach 6 Wochen gleichen sich beide Therapiearme an, Aclidiniumbromid ist in der zweiten Tageshälfte weiterhin numerisch überlegen.

Zusammenfassend bestätigt die Studie M/34273/39 die Ergebnisse der Studie M/34273/23, wenn auch mit kleineren Effekten. Dies könnte mit den unterschiedlichen Baselinewerten begründet werden.

Der $FEV_1 AUC$ ist in der M/34273/23 zu jedem Messzeitpunkt klinisch relevant (≥ 100 ml). In der M/34273/39 erreichen weder die Patienten unter Aclidiniumbromid, noch die unter Tiotropiumbromid so hohe Veränderungswerte wie in der M/34273/23. Die CFB der Aclidiniumbromid-Patienten ist mit der der Tiotropiumbromid-Patienten aber zu jedem Messzeitpunkt vergleichbar.

Das Bronchodilatationsprofil von Aclidiniumbromid ist in der ersten Tageshälfte und über 24 Stunden mit dem von Tiotropiumbromid vergleichbar; die CFB an Tag 15 liegt in der M/34273/23 sowohl bei Aclidiniumbromid (AB) als auch bei Tiotropiumbromid (TB) im klinisch relevanten Bereich ($CFB_{AB} FEV_1 AUC_{0-24, 0-12} = 240$ ml, $CFB_{TB} FEV_1 AUC_{0-24} = 160$ ml, $CFB_{TB} FEV_1 AUC_{0-12} = 250$). In der M/34273/39 ist dieser Effekt, wie bereits erwähnt, nicht so ausgeprägt.

In der zweiten Tageshälfte wird in der M/34273/23 durch die zweite Inhalation von Aclidiniumbromid eine signifikant überlegene Bronchodilatation erreicht ($CFB_{AB} FEV_1 AUC_{12-24} = 200$ ml, $CFB_{TB} FEV_1 AUC_{12-24} = 80$ ml; $p = 0,02$ ml).). In der M/34273/39 ist

dieser Effekt am ersten Tag signifikant (Tag1: $CFB_{AB} FEV_1 AUC_{12-24} = 90$ ml, $CFB_{TB} FEV_1 AUC_{12-24} = 20$ ml; $p = 0,002$), in der sechsten Woche numerisch ($CFB_{AB} FEV_1 AUC_{12-24} = 50$ ml, $CFB_{TB} FEV_1 AUC_{12-24} = 10$ ml; $p = 0,12$) feststellbar.

Die durch den Endpunkt $FEV_1 AUC$ analysierte Entwicklung der Lungenfunktion im Tagesverlauf belegt die Konstanz der Bronchodilatation unter Acclidiniumbromid-Behandlung und kann durch die zweimal tägliche Gabe die durch die zirkadiane Rhythmik verursachten tageszeitlichen Schwankungen (belastende COPD-Symptome in der Nacht und am Morgen) auffangen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Eine Meta-Analyse wurde aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.2 CFB Trough-FEV₁ – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-14: Operationalisierung von CFB Trough-FEV₁

Studie	Operationalisierung
M/34273/23	<p>Das forcierte Ein-Sekunden-Volumen (FEV₁) wird mit Hilfe eines Spirometers ermittelt. Dazu müssen die Patienten in ein Mundstück aus voller Lunge schnell ausatmen. Gemessen wird die Luftmenge (in Litern), die in der ersten Sekunde ausgeatmet wird.</p> <p>Der Trough-FEV₁-Wert wird vor der morgendlichen Bronchodilatation am ersten und letzten (15.) Tag der Behandlung ermittelt. Zur Berechnung der Veränderung gegenüber des Baseline-Wertes (CFB) wird der Trough-FEV₁ an Tag 1 und Tag 15 dem Trough-FEV₁-Wert vor Therapiebeginn gegenübergestellt.</p> <p>Die Baseline Werte werden als Durchschnitt zweier Messungen (eine und null Studien vor Therapiebeginn) vor jeder der drei Behandlungsperioden der drei-armigen cross-over-Studie ermittelt.</p>
LAS-MD-39	<p>Das forcierte Ein-Sekunden-Volumen (FEV₁) wird mit Hilfe eines Spirometers ermittelt. Dazu müssen die Patienten in ein Mundstück aus voller Lunge schnell ausatmen. Gemessen wird die Luftmenge (in Litern), die in der ersten Sekunde ausgeatmet wird.</p> <p>Der Trough-FEV₁- Wert wird vor der morgendlichen Bronchodilatation am ersten und letzten Tag der Behandlung (nach 6 Wochen) ermittelt. Zur Berechnung der Veränderung gegenüber des Baseline-Wertes (CFB) wird der Trough-FEV₁ an Tag 1 und nach 6 Wochen dem Trough-FEV₁-Wert vor Therapiebeginn gegenübergestellt.</p> <p>Die Baseline Werte werden als Durchschnitt zweier Messungen (eine und null Studien vor Therapiebeginn) vor Behandlungsbeginn ermittelt.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-15: Bewertung des Verzerrungspotenzials für CFB Trough-FEV₁ in RCT mit direktem Vergleich

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	niedrig	ja	ja	nein	nein	niedrig
M/34273/39	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential der beiden RCTs mit direktem Vergleich (M/34273/23 und M/34273/39) ist niedrig. Das Verzerrungspotential des erhobenen Endpunktes wird in beiden Studien ebenfalls als niedrig eingestuft; die Endpunkterheber sind verblindet, die Auswertung der Ergebnisse erfolgt nach dem ITT-Prinzip, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-16: Ergebnisse für CFB Trough-FEV₁ (l) aus RCT mit direktem Vergleich

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
M/34273/23					
Tag 1	AB (n=29)	1,46 (0,50)	1,63 (0,52)	0,17 (0,28)	0,06 (0,04) KI [-0,01; 0,14]
	TB (n=28)	1,49 (0,47)	1,59 (0,53)	0,09 (0,22)	p=0,08
Tag 15	AB (n=29)	1,46 (0,50)	1,62 (0,54)	0,16 (0,34)	0,03 (0,03) KI [-0,03; 0,10]
	TB (n=28)	1,49 (0,47)	1,59 (0,53)	0,10 (0,23)	p=0,25
M/34273/39					
Tag 1*	AB (n=171)	1,46 (0,48)	1,55 (0,50)	0,09 (0,21)	0,05 (0,02) KI [0,003; 0,09]
	TB (n=158)	1,54 (0,54)	1,58 (0,55)	0,03 (0,22)	p=0,04
6 Wochen	AB (n=171)	1,46 (0,48)	1,51 (0,50)	0,04 (0,20)	0,04 (0,03) KI [-0,01; 0,09]
	TB (n=158)	1,54 (0,54)	1,54 (0,55)	0,00 (0,26)	p=0,12
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * signifikanter Gruppenunterschied AB: Acclidiniumbromid, TB: Tiotropiumbromid, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall					

M/34273/23

Tabelle 4.3-16 zeigt, dass die Behandlung mit Acclidiniumbromid zu einem markanten Anstieg des Trough-FEV₁-Wertes führt (von 1,46 l auf 1,62 l am ersten Tag bzw. auf 1,63 l am 15. Tag). Dieser Behandlungseffekt war bei den mit Tiotropiumbromid behandelten Patienten etwas geringer ausgeprägt. Der Unterschied zur Baseline (CFB) ist dementsprechend bei Acclidiniumbromid-Patienten deutlicher zu erkennen (0,16 l bzw. 0,17 l gegenüber 0,09 l bzw. 0,10 l CFB). Die Signifikanz dieses Unterschiedes kann anhand des adjustierten ANCOVA-Modells jedoch nicht festgestellt werden (p>0,05).

M/34273/39

Tabelle 4.3-16 zeigt, dass die Behandlung mit Acclidiniumbromid im Gegensatz zu Tiotropiumbromid zu einem am ersten Tag signifikanten ($p=0,04$) und am Ende der Studie (Woche 6) zu einem numerischen Anstieg des vor-bronchodilatorischen Trough-FEV₁-Wertes führt. Der Trough-FEV₁-Wert zeigt im Gegensatz zu Acclidiniumbromid bei Tiotropiumbromid nach 6 Wochen keine Verbesserung der Lungenfunktion an (CFB=0,00).

Zusammenfassend besitzt Acclidiniumbromid durch die zweimal tägliche Gabe eine bessere Wirkung in der zweiten Tageshälfte, was sich in einer konstanten und stabileren Lungenfunktion zeigt.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Eine Meta-Analyse wurde aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

4.3.1.3.1.3 CFB Peak-FEV₁ – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-17: Operationalisierung von CFB Peak-FEV₁

Studie	Operationalisierung
M/34273/23	Das forcierte Ein-Sekunden-Volumen (FEV ₁) wird mit Hilfe eines Spirometers ermittelt. Dazu müssen die Patienten in ein Mundstück aus voller Lunge schnell ausatmen. Gemessen wird die Luftmenge (in Litern), die in der ersten Sekunde ausgeatmet wird. Der Peak-FEV ₁ -Wert ist als der innerhalb von 12 Stunden nach der Bronchodilatation höchste gemessene Wert definiert. Der Wert wird nach der morgendlichen ($\pm 9:00$ Uhr ± 1 Stunde) und abendlichen Bronchodilatation 21:00 Uhr ± 1 Stunde), sowie am ersten und letzten (15.) Tag der Behandlung ermittelt. Zur Berechnung der Veränderung gegenüber des Baseline-Wertes (CFB) wird der Peak-FEV ₁ an Tag 1 und Tag 15 dem Peak-FEV ₁ -Wert vor Therapiebeginn gegenübergestellt. Die Baseline Werte werden als Durchschnitt zweier Messungen (eine und null Studien vor Therapiebeginn) vor jeder der drei Behandlungsperioden der drei-armigen cross-over-Studie ermittelt.
M/34273/39	Das forcierte Ein-Sekunden-Volumen (FEV ₁) wird mit Hilfe eines Spirometers ermittelt. Dazu müssen die Patienten in ein Mundstück aus voller Lunge schnell ausatmen. Gemessen wird

die Luftmenge (in Litern), die in der ersten Sekunde ausgeatmet wird.

Der Peak-FEV₁-Wert ist als der innerhalb von 12 Stunden nach der Bronchodilatation höchste gemessene Wert definiert. Der Wert wird nach der morgendlichen ($\pm 9:00$ Uhr ± 1 Stunde) und abendlichen Bronchodilatation (21:00 Uhr ± 1 Stunde), sowie am ersten und letzten Tag der Behandlung (nach 6 Wochen) ermittelt. Zur Berechnung der Veränderung gegenüber des Baseline-Wertes (CFB) wird der Peak-FEV₁ an Tag 1 und nach 6 Wochen dem Peak-FEV₁-Wert vor Therapiebeginn gegenübergestellt.

Die Baseline Werte werden als Durchschnitt zweier Messungen (eine und null Stunden vor Therapiebeginn) vor Behandlungsbeginn ermittelt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-18: Bewertung des Verzerrungspotenzials für CFB Peak-FEV₁ in RCT mit direktem Vergleich

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	niedrig	ja	ja	nein	nein	niedrig
M/34273/39	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential der beiden RCTs mit direktem Vergleich (M/34273/23 M/34273/39) ist niedrig. Das Verzerrungspotential des erhobenen Endpunktes wird in beiden Studien ebenfalls als niedrig eingestuft; die Endpunkterheber sind verblindet, die Auswertung der Ergebnisse erfolgt nach dem ITT-Prinzip, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-19: Ergebnisse für CFB Peak-FEV₁ (l) aus RCT mit direktem Vergleich

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
M/34273/23					
1. Tageshälfte (± 9:00 bis ± 21:00 Uhr)					
Tag 1	AB (n=29)	1,46 (0,50)	1,83 (0,56)	0,37 (0,20)	0,48 (0,25) KI [-0,002; 0,1] p=0,059
	TB (n=28)	1,49 (0,47)	1,81 (0,56)	0,31 (0,20)	
Tag 15	AB (n=29)	1,46 (0,50)	1,88 (0,58)	0,42 (0,33)	0,026 (0,47) KI [-0,07; 0,12] p=0,59
	TB (n=28)	1,49 (0,47)	1,88 (0,61)	0,39 (0,25)	
2. Tageshälfte (± 21:00 bis ± 9:00 Uhr)					
Tag 1*	AB (n=29)	1,46 (0,50)	1,81 (0,56)	0,35 (0,26)	0,10 (0,03) KI [0,03; 0,16] p=0,007
	TB (n=28)	1,49 (0,47)	1,74 (0,57)	0,24 (0,26)	
Tag 15	AB (n=29)	1,46 (0,50)	1,79 (0,60)	0,33 (0,41)	0,05 (0,03) KI [-0,02; 0,11] p=0,16
	TB (n=28)	1,49 (0,47)	1,75 (0,58)	0,26 (0,26)	
M/34273/39					
1. Tageshälfte (± 9:00 bis ± 21:00 Uhr)					
Tag 1	AB (n=171)	1,46 (0,48)	1,72 (0,54)	0,26 (0,18)	0,01 (0,02) KI [-0,02; 0,05] p=0,42
	TB (n=158)	1,54 (0,54)	1,79 (0,59)	0,25 (0,18)	
6 Wochen	AB (n=171)	1,46 (0,48)	1,72 (0,55)	0,26 (0,23)	0,01 (0,03) KI [-0,04; 0,06] p=0,77
	TB (n=158)	1,54 (0,54)	1,79 (0,59)	0,25 (0,25)	
2. Tageshälfte (± 21:00 bis ± 9:00 Uhr)					
Tag 1*	AB (n=171)	1,46 (0,48)	1,63 (0,55)	0,17 (0,23)	0,08 (0,02) KI [0,04; 0,13] p=0,0005
	TB (n=158)	1,54 (0,54)	1,63 (0,58)	0,09 (0,21)	
6 Wochen	AB (n=171)	1,46 (0,48)	1,60 (0,52)	0,14 (0,23)	0,03 (0,03) KI [-0,03; 0,08] p=0,34
	TB (n=158)	1,54 (0,54)	1,65 (0,59)	0,11 (0,26)	
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * signifikanter Gruppenunterschied AB: Acclidiniumbromid, TB: Tiotropiumbromid, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall					

M/34273/23

Die Behandlung mit Acclidiniumbromid führte während der ersten Tageshälfte zu einem markanten Anstieg des Peak-FEV₁-Wertes (1,46 l auf 1,83 am ersten Tag bzw. auf 1,88 l am 15. Tag). Dieser Behandlungseffekt war bei den mit Tiotropiumbromid behandelten Patienten etwas geringer ausgeprägt. Der Unterschied zur Baseline (CFB) ist dementsprechend bei Acclidiniumbromid-Patienten deutlicher zu erkennen (0,37 bzw. 0,42 l gegenüber 0,31 bzw. 0,39 l CFB). Die Signifikanz dieses Unterschiedes kann anhand des adjustierten ANCOVA-Modells nicht festgestellt werden ($p > 0,05$).

In der zweiten Tageshälfte sind die CFB Peak-FEV₁-Werte der Acclidiniumbromid-Patienten gegenüber den Tiotropiumbromid-Patienten größer, am ersten Tag der Behandlung ist der Unterschied signifikant ($p = 0,007$).

M/34273/39

Die Behandlung mit Acclidiniumbromid führte auch in der 6-wöchigen Studie während der ersten Tageshälfte zu einem markanten Anstieg des Peak-FEV₁-Wertes (1,46 l auf 1,72). In der zweiten Tageshälfte ist die Veränderung zur Baseline gegenüber Tiotropiumbromid am ersten Tag signifikant ($p = 0,0005$) und nach 6 Wochen numerisch größer.

Der in beiden Studien beobachtete Effekt der besseren sofortigen Wirkung der COPD-Behandlung auf die Atemwegsobstruktion bei Acclidiniumbromidgabe ist in der zweiten Tageshälfte auf die zweimal tägliche Inhalation zurückzuführen. In der ersten Tageshälfte ist eine numerische Überlegenheit (M/34273/23) bzw. eine Gleichwertigkeit (M/34273/39) von Acclidiniumbromid festzustellen. Acclidiniumbromid besitzt somit eine mindestens gleichwertige Bronchodilatation gegenüber Tiotropiumbromid.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Eine Meta-Analyse wurde aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

4.3.1.3.1.4 CFB COPD-Symptomatik – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-20: Operationalisierung von COPD-Symptomatik

Studie	Operationalisierung
M/34273/23	<p>Tägliche Dokumentation der COPD-Symptome (Dyspnoe, Husten, Auswurf, nächtliche Symptome, vgl. Tabelle unten) durch den Patienten in einem Patiententagebuch. Nächtliche Symptome wurden bezüglich der letzten Bettruhe, die anderen COPD-Symptome bezüglich der letzten 24 h erfasst.</p> <p>Die Bewertung der COPD-Symptome erfolgt anhand von vier Kategorien; Brustenge, Husten, Auswurf und nächtliche Symptome, die an einen von Welte et al. 2006 [35] in einer klinischen Studie publizierten Score angelehnt sind. Die einzelnen Symptomausprägungen können in ihrer Schwere mit 0-4 Punkten bewertet werden (vgl. [32], S. 49).</p> <p>Zur Beurteilung der Veränderung der COPD-Symptomatik wurden die Fragebögen nach einer und nach zwei Wochen ausgewertet und der als Baseline (vor der ersten Einnahme) ermittelten Schwere der Symptomatik gegenübergestellt.</p> <p>Die CFB (Change from baseline)-Werte werden als Differenz der Baseline- und Follow-up-</p>

Werte (0-4) ermittelt. Die Gruppendifferenz wird durch ein ANCOVA-Model bestimmt.	
M/34273/39	<p>Tägliche Dokumentation der COPD-Symptome (Atemnot, Husten & Auswurf, Brustenge) durch den Patienten in einem digitalen Tagebuch mit Hilfe des EXACT- Respiratory Symptoms (E-RS) Fragebogen [12]. Dabei werden 11 Fragen (vgl. [13]) beantwortet, die anschließend zur Evaluation der Schwere der COPD-Symptome verwendet werden.</p> <p>Der E-RS stellt einen Teil des 14 Fragen enthaltenden EXACT (Exacerbations of Chronic Pulmonary Disease Tool) Fragebogens [36, 37] dar. Der EXACT Fragebogen wurde in Zusammenarbeit der FDA entwickelt und befindet sich in der Validierung²⁰.</p> <p>Die 11 vom Patienten beantworteten Fragen des E-RS werden für die drei Symptome (Atemnot, Husten & Auswurf, Brustenge) ausgewertet und zusätzlich zu einem Gesamt-E-RS- Score zusammengefasst. Je höher die Punktzahl der Einzel- und Gesamtscores sind, desto belastender werden die COPD-Symptome durch den Patienten empfunden. Im Gesamt-E-RS-Score beträgt die maximale Punktzahl 40, in den Symptom-spezifischen Kategorien können bei 5 Fragen zur Atemnot maximal 17, bei 3 Fragen zum Husten& Auswurf maximal 11 und bei 3 Fragen zur Brustenge maximal 12 Punkte erreicht werden. Zu jeder der 11 Fragen gibt es eine 5-stufige Antwortskala, die mit einer Punktzahl von 0 bis 4 Punkten verknüpft ist (bei manchen Fragen beträgt die maximale Punktzahl 3 statt 4). [13]</p> <p>Zur Beurteilung der Veränderung der COPD-Symptomatik wurden die Fragebögen nach sechs Wochen ausgewertet und der zur Baseline (vor der ersten Einnahme) ermittelten Schwere der Symptomatik gegenübergestellt.</p> <p>Die CFB (Change from baseline)-Werte werden als Differenz der Baseline- und Follow-up-Werte ermittelt. Die Gruppendifferenz wird durch ein ANCOVA-Model bestimmt.</p> <p>Außerdem werden mit einem vom Unternehmen entwickelten Fragebogen (vgl. [33] S. 86-87) vier weitere Aspekte der COPD-Symptome in einem digitalen Tagebuch erfasst:</p> <ul style="list-style-type: none"> • Schwere der nächtlichen Symptomatik (keine Symptome =1, leicht=2, mittel=3, schwer=4, sehr schwer=5) • Schwere der und morgendlichen Symptome (keine Symptome =1, leicht=2, mittel=3, schwer=4, sehr schwer=5) • Tage ohne morgendliche Symptomatik (ermittelt durch die Häufigkeit der gewählten Antwortmöglichkeit „ich hatte keine der Symptom“ bei der Frage nach morgendlichen Symptomen) • durch COPD-Symptome eingeschränkte Aktivität (gar nicht=1, ein bißchen=2, mittel=3, ziemlich=4, sehr=5). <p>Zur Beurteilung der Veränderung der COPD-Symptomatik wurden die Fragebögen nach sechs Wochen ausgewertet und der zur Baseline (vor der ersten Einnahme) ermittelten Symptomatik gegenübergestellt.</p> <p>Die CFB (Change from baseline)-Werte werden als Differenz der Baseline- und Follow-up-Werte ermittelt. Die Gruppendifferenz wird durch ein ANCOVA-Model bestimmt.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

²⁰ <http://www.exactproinitiative.com/default.php>

Tabelle 4.3-21: Bewertung des Verzerrungspotenzials für COPD-Symptomatik in RCT mit direktem Vergleich

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	niedrig	ja	ja	nein	nein	niedrig
M/34273/39	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential der beiden RCTs mit direktem Vergleich (M/34273/23 und LAS-MD-39) ist niedrig. Das Verzerrungspotential des erhobenen Endpunktes wird in beiden Studien ebenfalls als niedrig eingestuft; die Endpunkterheber sind verblindet, die Auswertung der Ergebnisse erfolgt nach dem ITT-Prinzip, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-22: Ergebnisse für COPD-Symptomatik aus RCT mit direktem Vergleich

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
M/34273/23					
Atemnot					
1 Woche	AB (n=29)	1,58 (1,03)	1,30 (0,95)	-0,28 (0,68)	-0,10 (0,13) KI [-0,36; 0,16]
	TB (n=28)	1,53 (1,01)	1,37 (1,00)	-0,15 (0,84)	p=0,45
2 Wochen	AB (n=29)	1,58 (1,03)	1,27 (0,96)	-0,31 (0,79)	-0,08 (0,16) KI [-0,40; 0,23]
	TB (n=28)	1,53 (1,01)	1,34 (1,03)	-0,19 (0,89)	p=0,59
Husten					
1 Woche	AB (n=29)	1,32 (1,00)	1,17 (0,92)	-0,14 (0,60)	-0,13 (0,13) KI [-0,40; 0,14]
	TB (n=28)	1,26 (0,97)	1,27 (0,96)	0,01 (0,72)	p=0,33
2 Wochen	AB (n=29)	1,32 (1,00)	1,17 (0,95)	-0,14 (0,81)	-0,13 (0,15) KI [-0,44; 0,18]
	TB (n=28)	1,26 (0,97)	1,28 (0,93)	0,02 (0,73)	p=0,40
Auswurf					
1 Woche	AB (n=29)	0,89 (0,78)	0,88 (0,74)	-0,02 (0,41)	-0,11 (0,19) KI [-0,28; 0,07]
	TB (n=28)	0,82 (0,69)	0,92 (0,74)	0,11 (0,35)	p=0,23
2 Wochen	AB (n=29)	0,89 (0,78)	0,88 (0,81)	-0,01 (0,55)	-0,18 (0,12) KI [-0,41; 0,05]
	TB (n=28)	0,82 (0,69)	1,01 (0,78)	0,19 (0,46)	β=0,13
Nächtliche Symptome					
1 Woche	AB (n=29)	0,66 (0,70)	0,50 (0,70)	-0,16 (0,44)	-0,13 (0,09) KI [-0,31; 0,04]
	TB (n=28)	0,66 (0,71)	0,64 (0,77)	-0,02 (0,54)	p=0,13
2 Wochen	AB (n=29)	0,66 (0,70)	0,52 (0,74)	-0,13 (0,53)	-0,09 (0,11) KI [-0,30; 0,13]
	TB (n=28)	0,66 (0,71)	0,62 (0,74)	-0,05 (0,57)	p=0,41
M/34273/39					
Atemnot (E-RS)					
6 Wochen	AB (n=171)	6,50 (3,30)	5,50 (3,40)	-0,90 (2,00)	-0,40 (0,22) KI [-0,80; 0,10]
	TB (n=158)	6,00 (3,20)	5,60 (3,50)	-0,40 (1,90)	p=0,09
Husten & Auswurf (E-RS)					

6 Wochen	AB (n=171)	3,90 (1,60)	3,20 (1,70)	-0,70 (1,20)	-0,20 (0,12) KI [-0,40; 0,00]
	TB (n=158)	3,60 (1,50)	3,20 (1,50)	-0,40 (1,10)	p=0,08
Brustenge (E-RS)					
6 Wochen	AB (n=171)	3,30 (1,70)	2,70 (1,80)	-0,50 (1,20)	-0,10 (0,13) KI [-0,40; 0,10]
	TB (n=158)	2,90 (1,60)	2,70 (1,70)	-0,30 (1,10)	p=0,26
Gesamt E-RS Score					
6 Wochen	AB (n=171)	13,60 (5,80)	11,50 (6,30)	-2,00 (4,00)	-0,70 (0,42) KI [-1,60; 0,10]
	TB (n=158)	12,50 (5,80)	11,50 (6,20)	-1,10 (3,70)	p=0,07
Tage ohne morgendliche Symptome					
6 Wochen	AB (n=171)	11,13 (23,05)	17,48 (29,61)	6,61 (20,89)	3,37 (2,09) KI [-0,75; 7,49]
	TB (n=158)	13,95 (24,91)	16,83 (28,95)	2,77 (18,42)	p=0,11
Schwere der morgendlichen Symptome					
6 Wochen	AB (n=171)	2,36 (0,61)	2,16 (0,68)	-0,19 (0,46)	-0,09 (0,05) KI [-0,19; 0,00]
	TB (n=158)	2,25 (0,62)	2,17 (0,65)	-0,08 (0,40)	p=0,052
Schwere der nächtlichen Symptome					
6 Wochen	AB (n=171)	1,99 (0,69)	1,82 (0,70)	-0,17 (0,43)	-0,07 (0,05) KI [-0,16 ; 0,02]
	TB (n=158)	1,90 (0,65)	1,81 (0,68)	-0,08 (0,43)	p=0,13
Grad der Einschränkung der Aktivität durch COPD-Symptome					
6 Wochen*	AB (n=171)	2,01 (0,70)	1,86 (0,74)	-0,14 (0,43)	-0,10 (0,05) KI [-0,19 ; -0,01]
	TB (n=158)	1,94 (0,72)	1,91 (0,76)	-0,03 (0,39)	p=0,04
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * signifikanter Gruppenunterschied AB: Acclidiniumbromid, TB: Tiotropiumbromid, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall					

M/34273/23

Während der zweiwöchigen Behandlung mit Acclidiniumbromid konnten - ausgehend von der Baselineerhebung – drei der vier Ausprägungen der COPD-Symptome (Atemnot, Husten, nächtliche Symptome) deutlich verbessert werden (CFB=-0,31 bzw. -0,14 bzw. -0,13). Bei der Symptomkategorie ‚Auswurf‘ war im Acclidiniumbromidarm keine Veränderung (CFB=-0,01) zu beobachten. Im Tiotropiumbromidarm war ein Anstieg der Schwere dieser Symptomkategorie zu verzeichnen (CFB=0,19). Nahezu keine Veränderung zeigten sich bei der Behandlung mit Tiotropiumbromid für den Husten (CFB=0,02) und die nächtlichen Symptome (CFB=-0,05).

Bei der Gegenüberstellung der Therapieeffekte besitzt Acclidiniumbromid in jeder der vier Kategorien einen numerischen Vorteil, der nicht signifikant ($p > 0,05$) ist.

M/34273/39

Die Ergebnisse der 6-wöchigen Studie bestätigen die Ergebnisse der M/34273/23. Die Unterschiede zwischen den Therapiearmen sind noch deutlicher ausgeprägt. Die Therapie mit Acclidiniumbromid ist in allen Symptomkategorien deutlich numerisch überlegen; dies gilt für die Kategorien des E-RS Fragebogens und die spezifischen Fragen zur morgendlichen und nächtlichen Symptomatik. Der eindeutige numerische Vorteil von Acclidiniumbromid hinsichtlich aller Aspekte der COPD-Symptomatik spiegelt sich im signifikanten ($p = 0,04$) Grad der Einschränkung der körperlichen Aktivität durch COPD-Symptome wieder. Diese Verbesserung stellt einen bedeutenden Zusatznutzen für den Patienten dar.

Zusammenfassend können gegenüber der Tiotropiumbromidbehandlung unter der Acclidiniumbromidbehandlung in beiden Studien numerische Vorteile bei der Verbesserung der COPD-Symptomatik beobachtet werden. Ein signifikanter Unterschied liegt beim Grad der Einschränkung der körperlichen Aktivität durch COPD-Symptome vor.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Eine Meta-Analyse wurde aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

4.3.1.3.2 Gesundheitsbezogene Lebensqualität - RCT

4.3.1.3.2.1 Behandlungszufriedenheit- RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-23: Operationalisierung der Behandlungszufriedenheit

Studie	Operationalisierung
M/34273/23	<p>Zum Studienende wurden alle Patienten aufgefordert 11 Fragen zu ihrem Eindruck und Erfahrungen mit den beiden während der Studie verwendeten Inhalationssystemen zu beantworten. Die Abschätzung der die Behandlungszufriedenheit betreffenden Fragen umfasste die Einfachheit der Bedienung und der Vorbereitung der Inhalation (fünfstufige Skala von sehr einfach bis sehr schwer), sowie die Bevorzugung eines der beiden Inhalationssysteme (Genuair[®], HandiHaler[®], keine Präferenz).</p> <p>Aufgrund des duple-dummy-Designs verwendet jeder Patient beide Inhalatoren und kann dessen Handhabung unabhängig von der Wirksamkeit des enthaltenen Arzneimittels bewerten. Bedingt durch das cross-over-Design der Studie wird der Patient vor jeder Behandlungsperiode bezüglich der Verwendung der Inhalatoren geschult.</p>
M/34273/39	<p>Zum Studienende wurden alle Patienten aufgefordert 8 Fragen zu ihrer Präferenz (Genuair[®], HandiHaler[®], keine Präferenz) bezüglich der beiden während der Studie verwendeten Inhalationssystemen zu beantworten. Die Fragen betrafen die Vorbereitung der Inhalation, die Bedienung, die Handhabung, die Durchführung, die Erlernbarkeit, der Bevorzugung bezüglich des Feedbacksystems und die allgemeine Bevorzugung eines der beiden Inhalationssysteme. Außerdem wird die Bereitschaft der Patienten ermittelt, den Inhalatoren weiterzuverwenden. Dazu wird eine Skala von 0-100 (0=keine Bereitschaft, 100=eindeutige Bereitschaft) vorgegeben, auf dem die Patienten beide Inhalatoren einordnen müssen.</p> <p>Aufgrund des duple-dummy-Designs verwendet jeder Patient beide Inhalatoren und kann dessen Handhabung unabhängig von der Wirksamkeit des enthaltenen Arzneimittels bewerten. Der Patient wird zu Beginn der Studie bezüglich der Verwendung der Inhalatoren geschult.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-24: Bewertung des Verzerrungspotenzials für die Behandlungszufriedenheit in RCT mit direktem Vergleich

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	niedrig	ja	ja	nein	nein	niedrig
M/34273/39	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential der beiden RCTs mit direktem Vergleich (M/34273/23 und M/34273/39) ist niedrig. Das Verzerrungspotential des erhobenen Endpunktes wird in beiden Studien ebenfalls als niedrig eingestuft; die Endpunkterheber sind verblindet, die Auswertung der Ergebnisse erfolgt nach dem ITT-Prinzip, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-25: Ergebnisse für die Behandlungszufriedenheit aus RCT mit direktem Vergleich

Studie	Frage	Antwort	Genuair®	Handihaler®	keine Präferenz
M/34273/23 (n=30)					
Einfache Bedienung (in %)					
		Sehr einfach	24 (80,00%)	16 (53,33%)	-
		Einfach	5 (16,67%)	9 (30,00%)	-
		Normal	0 (0,00%)	5 (16,67%)	-
		Schwer	0 (0,00%)	0 (0,00%)	-
		Sehr schwer	1 (3,33%)	0 (0,00%)	-
Einfache Vorbereitung der Inhalation (in %)					
		Sehr einfach	25 (83,33%)	14 (46,67%)	-
		Einfach	4 (13,33%)	12 (40,00%)	-
		Normal	1 (3,33%)	4 (13,33%)	-
		Schwer	0 (0,00%)	0 (0,00%)	-
		Sehr schwer	0 (0,00%)	0 (0,00%)	-
Bevorzugung (in %)					
		Ich präferiere maßgeblich	9 (30,00%)	2 (6,67%)	12 (40%)
		Ich präferiere etwas	6 (20,00%)	2 (3,33%)	
M/34273/39 (n=414)					
Bevorzugung aufgrund der einfachen Bedienung (in %)					
		Ich präferiere	357 (86,70%)*	37 (9,00%)	18 (4,40%)
Bevorzugung aufgrund der einfachen Vorbereitung der Inhalation (in %)					
		Ich präferiere	356 (86,40%)*	32 (7,80%)	24 (5,80%)
Bevorzugung aufgrund der einfachen Handhabung (in %)					
		Ich präferiere	334 (81,10%)*	38 (9,20%)	40 (9,70%)
Bevorzugung aufgrund der einfachen Durchführung (in %)					
		Ich präferiere	329 (79,90%)*	35 (8,50%)	48 (11,70%)

Bevorzugung bezüglich des Feedbacksystems (in %)				
	Ich präferiere	314 (76,20%)*	39 (9,50%)	59 (14,30%)
Bevorzugung aufgrund der einfachen Erlernbarkeit (in %)				
	Ich präferiere	326 (79,10%)*	33 (8,00%)	53 (12,90%)
Bevorzugung aufgrund des Komforts (in %)				
	Ich präferiere	360 (87,40%)*	32 (7,80%)	20 (4,90%)
Bevorzugung (in %)				
	Ich präferiere	330 (80,10%)*	44 (10,70%)	38 (9,20%)
Bereitschaft der Patienten, den Inhalatoren weiterzuverwenden; Score ¹ (SD)				
	Ich präferiere	88,80 (22,1)*	45,40 (35,5)	-
* signifikanter (p<0,001) Unterschied zum HandiHaler [®]				
¹ Mittelwert; 0=keine Bereitschaft, 100=eindeutige Bereitschaft				

In der Studie M/34273/23 stuften 80% der Patienten die Bedienung des Genuair[®]-Inhalators als sehr einfach ein. Beim HandiHaler[®] waren es nur 53%. Ähnlich verhielt es sich mit der Vorbereitung der Inhalation (83% vs. 47%). Bei der Festlegung auf ein Inhalationssystem präferierten 30% maßgeblich und 20% der Patienten etwas den Genuair[®]-Inhalator. Insgesamt bevorzugen damit 50 % der Patienten den Genuair-Inhalator. 40% der Patienten präferierten keinen der beiden Inhalatoren.

In der Studie M/34273/39 ist die Präferenz für den Genuair[®]-Inhalator noch deutlicher zu erkennen: Durchschnittlich 80% der Patienten präferierten den Genuair[®]-Inhalator – sowohl im Rahmen der sieben verschiedenen Aspekte, als auch in der Frage nach der Gesamtpräferenz. Die Bereitschaft, den Genuair[®]-Inhalator weiterzuverwenden ist im Gegensatz zur Weiterverwendung des HandiHaler[®] ebenfalls eindeutig (88,0%).

Im Vergleich der Studienergebnisse wird deutlich, dass der HandiHaler[®] erst nach mehrmaliger Schulung (M/34273/23) der Patienten eine mittlere Akzeptanz bei den Patienten erreicht. Bei einmaliger Schulung (M/34273/39) liegt die Bevorzugung hingegen bei maximal 10%. Dies kann auf die komplexe Handhabung des HandiHaler[®] zurückgeführt werden. Bei dem Genuair[®]-Inhalator ist hingegen bereits nach einmaliger Schulung einer sehr hohe Präferenz von 80% zu beobachten. Da Anwendungsfehler, wie in Modul 3 beschrieben, ein relevantes Problem in der optimalen Versorgung von COPD-Patienten darstellen, ist die beschriebene Überlegenheit des Genuair[®]-Inhalator ein klinisch höchst relevantes Ergebnis.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Eine Meta-Analyse wurde aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

4.3.1.3.3 Unerwünschte Arzneimittelwirkungen – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-26: Operationalisierung von unerwünschten Arzneimittelwirkungen

Studie	Operationalisierung
M/34273/23	<p>Die AEs (unerwünschte Ereignisse) wurden vom Unterschreiben der Einwilligungserklärung bis zum Abschluss der Studie (inkl. der Folgeuntersuchung) oder bis zum vorzeitigen Verlassen der klinischen Studie durch den Patienten aufgezeichnet. Unerwünschte Ereignisse wurden durch nicht-suggestive Fragen (z. B. Wie geht/ging es Ihnen?) an die Patienten und durch deren tägliche Aufzeichnungen von AEs auf der Patienten-Tagebuchkarte zwischen den Visiten in Erfahrung gebracht. Alle auf der Patienten-Tagebuchkarte festgehaltenen AEs wurden durch die Untersucher auf der AE-Berichtsbogen (Case Record Form, CRF) Seite in medizinische Terminologie übertragen.</p> <p>Im Verlauf der klinischen Studie mussten nur die Anomalien in Labortests, EKGs und körperlicher Untersuchung als AEs festgehalten werden, die von dem Untersucher als klinisch relevant befunden wurden oder eine Intervention, einen Diagnostest oder ein Absetzen des IMP erforderten. Zum Zeitpunkt der Unterzeichnung der Einwilligungserklärung angegebene Erkrankungen wurden nur dann als AEs berücksichtigt, wenn sie sich nach dieser Zeit verschlimmert hatten.</p>

Alle von dem Untersucher gesammelten und von Patienten während der definierten AE Aufzeichnungsperiode berichteten AEs mussten auf der AE-CRF-Seite festgehalten werden. Wenn im betreffenden Zeitraum keine AEs auftraten, musste das entsprechende Kontrollkästchen auf der AE-CRF-Seite angekreuzt werden. Wenn ein AE die Kriterien eines schweren AEs (SAE, Serious Adverse Event) erfüllte, musste dies außerdem auf dem SAE-Formular festgehalten und im vorgegebenen zeitlichen Rahmen gemeldet werden.

Ein **SAE** war jedes AE, das:

- tödlich war
- lebensbedrohlich war.
- eine stationäre Behandlung oder einen verlängerten bestehenden Krankenhausaufenthalt erforderte
- zu einer bleibenden oder schwerwiegenden Behinderung oder Invalidität geführt hat
- zu einer angeborenen Anomalie/einem Geburtsdefekt führte
- eine anderes medizinisch wichtiges Ereignis war, das den Patienten gefährdet haben könnte, oder einen Eingriffe zur Verhinderung eines der anderen oben genannten Ergebnisse erfordert hat.

„Lebensbedrohlich“ bedeutet in diesem Zusammenhang, dass der an der Studie teilnehmende Patient zum Zeitpunkt des Ereignisses in unmittelbarer Todesgefahr schwebte. Das bedeutete nicht, dass das Ereignis hypothetisch den Tod verursacht haben könnte, wenn es schwerer gewesen wäre oder länger gedauert hätte.

Ein Krankenhausaufenthalt wurde als eine Übernachtung im Krankenhaus oder der Notaufnahme definiert. Eine Verlängerung des Krankenhausaufenthaltes wurde definiert als jede Verlängerung eines stationären Krankenhausaufenthaltes über den erwarteten/erforderlichen Aufenthalt hinaus bezogen auf den ursprünglichen Grund für die erstmalige Einweisung, die durch den Untersucher oder den behandelnden Arzt festgelegt wurde.

Studienabbrüche aufgrund von Interventionsbedingten AEs (TAEs) und schweren TAEs (TSAEs) wurden als solche erfasst.

Die **Kausalität** zwischen einem AE und der Intervention wurde von dem Untersucher nach den unten aufgeführten Informationen in 2 Kategorien (Zusammenhang oder kein Zusammenhang) bewertet:

Zusammenhang: AEs mit guter Begründung und hinreichenden Informationen (z. B. plausible zeitliche Abfolge, Dosis-Wirkungs-Beziehung, Pharmakologie, positive Reaktion auf Absetzen der Medikation und/oder Reprovokation), die auf einen kausalen Zusammenhang mit der Intervention hindeuten, der plausibel, vorstellbar oder wahrscheinlich ist.

Kein Zusammenhang: AEs mit guter Begründung und hinreichenden Informationen (z. B. unplausible zeitliche Abfolge und/oder Rückführbarkeit auf parallele Erkrankungen oder andere Medikamente), die einen kausalen Zusammenhang mit der Intervention ausschließen.

M/34273/39	<p>Ein AE war definiert als jegliches unerwünschte medizinische Vorkommnis bei einem Patienten einer klinischen Studie (unabhängig von der Intervention und des kausalen Zusammenhang mit dem berichteten Ereignis).</p> <p>Daher konnte ein AE jegliches unerwünschte und nicht absichtlich herbeigeführte medizinische Vorkommnis während der Teilnahme des Patienten an der Studie sein, was die Verschlechterung einer zuvor bestehenden Erkrankung, ungewöhnliche Werte bei einer Laboruntersuchung, Anomalien im EKG, der Herzfrequenz und des Blutdrucks oder einen Befund mit Anomalien bei der körperlichen Untersuchung beinhaltet.</p> <p>AEs müssen mit der Teilnahme des Patienten an der Studie zeitlich in Verbindung gestanden haben, d. h. sie müssen zwischen der Unterzeichnung der Einwilligungserklärung und dem letzten planmäßigen Kontakt mit dem Patienten aufgetreten sein. Zum Zeitpunkt des Auftretens eines AE musste die Intervention daher noch nicht eingeleitet worden sein. Wenn sie bereits eingeleitet wurde, musste dies nicht unbedingt in einer kausalen Beziehung zu dem</p>
------------	---

Ereignis stehen.

Ein **SAE** war jedes AE, das:

- tödlich war
- lebensbedrohlich war.
- eine stationäre Behandlung oder einen verlängerten bestehenden Krankenhausaufenthalt erforderte
- zu einer bleibenden oder schwerwiegenden Behinderung oder Invalidität geführt hat
- eine angeborene Anomalie/ein Geburtsdefekt war
- ein anderes medizinisch wichtiges Ereignis war, das den Patienten gefährdet haben könnte, oder einen Eingriff zur Verhinderung eines der anderen oben genannten Ergebnisse erfordert hat.

Lebensbedrohlich in diesem Zusammenhang bedeutet, dass der an der Studie teilnehmende Patient zum Zeitpunkt des Ereignisses in unmittelbarer Todesgefahr schwebte. Das bedeutete nicht, dass das Ereignis hypothetisch den Tod verursacht haben könnte, wenn es schwerer war oder länger gedauert hat.

Ein Krankenhausaufenthalt wurde als eine Übernachtung im Krankenhaus oder der Notaufnahme definiert. Eine Verlängerung des Krankenhausaufenthaltes wurde definiert als jede Verlängerung eines stationären Krankenhausaufenthaltes über den erwarteten/erforderlichen Aufenthalt bezogen auf den ursprünglichen Grund für die erstmalige Einweisung, die durch den Untersucher oder den behandelnden Arzt festgelegt wurde. Die Analysen von Safety-Outcomes wurden über die Safety-Population durchgeführt.

Studienabbrüche aufgrund von Interventionsbedingten AEs (TAEs) und schweren TAEs (TSAEs) wurden als solche erfasst.

Der Untersucher stellte eine Bewertung der **kausalen Beziehung** zur Studienmedikation zur Verfügung. Die Kausalitätsbewertung wurde auf der entsprechenden Seite für AE-Aufzeichnungen im elektronischen CRF des Patienten festgehalten. Die kausale Beziehung wurde von dem Untersucher unter Beantwortung der folgenden Frage bewertet: Liegt es im Bereich des realistisch Möglichen, dass das Medikament das Ereignis ausgelöst hat?

Ja: Es gab eine mögliche oder wahrscheinliche Beziehung; d. h. es gab eine mittlere oder starke zeitliche Beziehung und die Ereignisse standen wahrscheinlich nicht mit anderen Arzneien, parallelen Erkrankungen oder anderen Faktoren im Zusammenhang.

Nein: Die Beziehung war unwahrscheinlich oder nicht nachvollziehbar; d. h. es gab keine starke zeitliche Beziehung und/oder Beziehung zur Verwendung anderer Arzneien, zu parallelen Erkrankungen oder anderen Faktoren, die plausible Erklärungen für das Auftreten des Ereignisses geliefert hätten.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-27: Bewertung des Verzerrungspotenzials für unerwünschte Arzneimittelwirkungen in RCT mit direktem Vergleich

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	niedrig	ja	ja	nein	nein	niedrig
LAS-MD-39	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential der beiden RCTs mit direktem Vergleich (M/34273/23 und LAS-MD-39) ist niedrig. Das Verzerrungspotential des erhobenen Endpunktes wird in beiden Studien ebenfalls als niedrig eingestuft; die Endpunkterheber sind verblindet, die Auswertung der Ergebnisse erfolgt nach dem ITT-Prinzip, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-28: Ergebnisse für unerwünschte Arzneimittelwirkungen aus RCT mit direktem Vergleich

Studie	M/34273/23			M/34273/39	
	Vor der Behandlung	Acridinium- bromid	Tiotropium- bromid	Acridinium- bromid	Tiotropium- bromid
Ausprägung	n (%)	n (%)	n (%)	n (%)	n (%)
	N=30	N=29	N=28	N=171	N=158
Patienten mit unerwünschten Arzneimittelwirkungen (AEs)					
AEs die nicht auf die Intervention zurückzuführen sind (non TEAEs)	1 (3,3%)	NR	NR	12 (7,0%)	12 (7,6%)
AEs die auf die Intervention zurückzuführen sind (TEAEs)	NR	7 (24,14%)	3 (10,71%)	47 (27,5%)	47 (29,7%)
TEAEs, die zum Abbruch der Studie führten	NR	0	0	3 (1,8%)	2 (1,3%)
Patienten mit schweren unerwünschten Arzneimittelwirkungen (SAEs)					
schwere AEs die nicht auf die Intervention zurückzuführen sind (non TESAEs)	0	0	0	1 (0,6%)	0
Schwere AEs die auf die Intervention zurückzuführen sind (TESAEs)	NR	0	0	3 (1,8%)	4 (2,5%)
TESAEs, die zum Abbruch der Studie führten	NR	0	0	0	1 (0,6%)
NR: nicht berichtet, da vor Behandlungsbeginn					

Auf die Therapie zurückzuführende unerwünschte Arzneimittelwirkungen traten in der 15-tägigen Studie (M/34273/23) bei 10 bzw. 25% der mit Acridiniumbromid oder Tiotropiumbromid behandelten Patienten auf. In der 6-wöchigen Studie (M/34273/39) lag der Anteil in beiden Gruppen bei durchschnittlich 28%. Bei durchschnittlich 1,5% der Patienten führte dies in der Studie M/34273/39 zu einem Therapieabbruch. Bei der 15-tägigen Studie M/34273/23 wurde bei 7% (Acridiniumbromidarm) bzw. 3% (Tiotropiumbromidarm) der Patienten ein AE berichtet.

Schwere Unerwünschte Ereignisse (SAEs) traten nur in der längeren der beiden Studien mit direktem Vergleich auf. Zwei Ereignisse (schwere Schluckstörungen und schwere Aspiration) bei einem Patienten waren nicht auf die Therapie zurückzuführen. Exacerbationen führten bei einem Patienten zum Studienabbruch. SAEs, die auf die Therapie zurückzuführen waren, wurden in beiden Vergleichsgruppen der Studie M/34273/39 bei durchschnittlich zwei Prozent der Patienten dokumentiert.

Die Anteile der Patienten mit AEs, SAEs, sowie aufgrund von AEs und SAEs stattgefundenen Studienabbrüchen lassen auf keinen Unterschied im Sicherheitsprofil der beiden gegenübergestellten Arzneimittel schließen.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Eine Meta-Analyse wurde aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.4 Mortalität - RCT

4.3.1.3.4.1 COPD-assoziierte kardiovaskuläre Morbidität und Mortalität - RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-29: Operationalisierung von COPD-assoziierte kardiovaskuläre Morbidität und Mortalität

Studie	Operationalisierung
M/34273/23	Schwerwiegende Nebenwirkungen, die als kardiale (Herzerkrankungen wie z. B. instabile Angina, ventrikuläre Extrasystole; Myokardinfarkt) oder vaskuläre (Gefäßerkrankungen, wie z. B. periphere Durchblutungsstörungen; Schlaganfall) Ereignisse definiert sind, sowie Todesfälle, die mit kardialen oder vaskulären Ereignissen in Verbindung stehen.
M/34273/39	Schwerwiegende Nebenwirkungen, die als kardiale (Herzerkrankungen wie z. B. instabile Angina, ventrikuläre Extrasystole; Myokardinfarkt) oder vaskuläre (Gefäßerkrankungen, wie z. B. periphere Durchblutungsstörungen; Schlaganfall) Ereignisse definiert sind, sowie Todesfälle, die mit kardialen oder vaskulären Ereignissen in Verbindung stehen.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-30: Bewertung des Verzerrungspotenzials für COPD-assoziierte kardiovaskuläre Morbidität und Mortalität in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	niedrig	ja	ja	nein	nein	niedrig
M/34273/39	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential der beiden RCTs mit direktem Vergleich (M/34273/23 und M/34273/39) ist niedrig. Das Verzerrungspotential des erhobenen Endpunktes wird in beiden Studien ebenfalls als niedrig eingestuft; die Endpunkterheber sind verblindet, die Auswertung der Ergebnisse erfolgt nach dem ITT-Prinzip, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-31: Ergebnisse für COPD-assoziierte kardiovaskuläre Morbidität und Mortalität aus RCT mit einem direkten Vergleich

Studie	M/34273/23		M/34273/39	
	Aclidinium-bromid n (%) N=29	Tiotropium-bromid n (%) N=28	Aclidinium-bromid n (%) N=171	Tiotropium-bromid n (%) N=158
Kardiale SAEs (nicht tödlich)	0	0	1 (0,6%)	1 (0,6)
Vaskuläre SAEs (nicht tödlich)	0	0	1 (0,6%)	1 (0,6)
Kardiale oder vaskuläre tödliche Ereignisse	0	0	0	0

Schwere kardiovaskuläre Ereignisse traten in einer der Studien (M/34273/39) sehr selten und gleichverteilt zwischen den zwei Therapiearmen auf. In der Studie (M/34273/23) wurden aufgrund der kurzen Studiendauer keine entsprechenden Ereignisse registriert. Die geringen Fallzahlen lassen keine gesicherte Aussage aber einen Hinweis zur vergleichbaren COPD-assoziierten kardiovaskuläre Morbidität zu. Aussagen zur COPD-assoziierten kardiovaskulären Mortalität sind aufgrund ausschließlich nicht tödlicher SAEs nicht möglich.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Eine Meta-Analyse wurde aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

4.3.1.3.4.2 COPD bedingte Letalität, Gesamtmortalität - RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammengefasst und als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Werden Studien in der Meta-Analyse nicht berücksichtigt, ist das zu begründen.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-32: Operationalisierung von COPD bedingte Letalität, Gesamtmortalität

Studie	Operationalisierung
M/34273/23	Todesfall im Studienzeitraum, ggf. Verknüpfung zur COPD
M/34273/39	Todesfall im Studienzeitraum, ggf. Verknüpfung zur COPD

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-33: Bewertung des Verzerrungspotenzials für COPD bedingte Letalität, Gesamtmortalität in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
M/34273/23	niedrig	ja	ja	nein	nein	niedrig
M/34273/39	niedrig	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential der beiden RCTs mit direktem Vergleich (M/34273/23 und M/34273/39) ist niedrig. Das Verzerrungspotential des erhobenen Endpunktes wird in beiden Studien ebenfalls als niedrig eingestuft; die Endpunkterheber sind verblindet, die Auswertung der Ergebnisse erfolgt nach dem ITT-Prinzip, es gibt keine Hinweise auf ergebnisgesteuerte Berichterstattung oder sonstige das Verzerrungspotential beeinflussende Punkte.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

In beiden Studien (M/34273/23 und M/34273/39) wurde kein Todesfall während der Studiendauer dokumentiert. Eine Auswertung des Endpunktes ‚COPD assoziierte Letalität, Gesamtmortalität‘ ist somit weder auf Studienebene, noch im Rahmen der Meta-Analyse möglich.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Eine Meta-Analyse wurde aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

4.3.1.4 Subgruppenanalysen – RCT

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Der im Beratungsgespräch empfohlenen Berücksichtigung des Stufenschemas der Nationalen VersorgungsLeitlinie (NVL) [1] soll mit der in Abschnitt 4.2.5.5 beschriebenen Subgruppenanalyse Rechnung getragen werden. Danach sollen nach Möglichkeit neben den Schweregraden der COPD auch die Anzahl der Exazerbationen in den letzten 12 Monaten vor Studienbeginn berücksichtigt werden.

Für die Darstellung der Subgruppenergebnisse für die Studie M/34273/23 musste aufgrund von fehlenden Exazerbationsdaten von dem im Abschnitt 4.2.5.5 beschriebenen Vorgehen abgewichen werden. In der Studie M/34273/23 ist in der Anamnese des bisherigen Verlaufs der COPD lediglich das Datum der letzten Exazerbation dokumentiert. Die Anzahl der anamnestischen Exazerbationen wurde hingegen nicht dokumentiert.

Die Verwendung der anamnestischen ICS-Medikation stellt keine hinreichenden Surrogatparameter für die Anamnese der COPD dar. In der Studie befinden sich 14 Patienten mit anamnestischer ICS-Medikation, jedoch nur 6 Patienten mit mindestens einer Exazerbation im Jahr vor Studieneinschluss.

Im Anschluss an die Beschreibung der Subgruppenanalysen für M/34273/39 werden die Ergebnisse der Subgruppenanalyse der Studie M/34273/23, wie in Abschnitt 4.2.5.5 beschrieben, nach den Schweregraden II und II dargestellt²¹.

Für die Darstellung der Subgruppenergebnisse für die Studie M/34273/39 musste aufgrund zu kleiner Patientenzahlen in den im Abschnitt 4.2.5.5 definierten Subgruppen ebenfalls von dem geplanten Vorgehen abgewichen werden. Wie in Tabelle 4.3-34 dargestellt, befinden sich

²¹ Es bedingt sich nur ein Patient der Stufe IV in der Studie.

keine Patienten der COPD-Stufe IV (Subgruppe 4 und 5) und lediglich 2 Patienten der Subgruppe 3 in der Studie M/34273/39.

Tabelle 4.3-34: Darstellung der Patientenpopulation der M/34273/39 nach definierten Subgruppen

Subgruppen		Placebo	Acidinium- bromid 400	Tiotropium- bromid	Gesamt
1	Stufe II	n=58	n=108	n=104	n=270
2	Stufe III mit ≤ 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=27	n=61	n=53	n=141
3	Stufe III mit > 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=0	n=2	n=0	n=2
4+5	Stufe IV	n=0	n=0	n=0	n=0
Gesamt					

Der Grund für die geringe Anzahl von Patienten mit mehr als zwei Exazerbationen im Jahr vor Studieneinschluss kann durch die Definition der Ausschlusskriterien begründet sein. Diese schließen Patienten aus, die innerhalb der letzten 6 Wochen vor Studieneinschluss Exazerbationen hatten. So soll auf das Ziel der Studie, die Wirksamkeit von Acidiniumbromid bei stabilen COPD-Patienten nachzuweisen, fokussiert werden. Unter leitlinienkonformen Versorgungsaspekten ist noch darauf hinzuweisen, dass die Patienten, die schon mindestens 4 Wochen vor Run-In-Phase zusätzliche ICS-Medikation erhielten, diese auch während der Studie fortführen konnten (vgl. Tabelle 4.3-6). Die Anforderungen der NVL bzgl. der zusätzlichen Gabe von ICS werden somit im Studiendesign berücksichtigt.

Im Folgenden werden die Subgruppenergebnisse der Studien M/34273/23 und M/34273/39, wie in Abschnitt 4.2.5.5 beschrieben, nach Schweregraden dargestellt.

Tabelle 4.3-35: Subgruppenergebnisse für FEV₁ AUC (l) aus RCT mit direktem Vergleich

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
M/34273/23					
FEV₁ AUC₀₋₁₂					
Tag 1					
Stufe II	AB (n=19)	1.63 (0,44)	1,89 (0,52)	0,27 (0,19)	0,05 (0,03) KI [-0,02; 0,11] p=0,15
	TB (n=19)	1.63 (0,43)	1,85 (0,50)	0,22 (0,18)	
Stufe III	AB (n=10)	1.15 (0,47)	1,33 (0,39)	0,18 (0,22)	0,05 (0,05) KI [-0,05; 0,15] p=0,31
	TB (n=9)	1.20 (0,41)	1,28 (0,34)	0,08 (0,19)	
<i>Interaktion p=0,61</i>					
Tag 15					
Stufe II	AB (n=19)	1.63 (0,44)	1,92 (0,56)	0,30 (0,30)	0,01 (0,04) KI [-0,06; 0,09] p=0,71
	TB (n=19)	1.63 (0,43)	1,92 (0,59)	0,29 (0,29)	
Stufe III	AB (n=10)	1.15 (0,47)	1,28 (0,22)	0,13 (0,42)	-0,11 (0,10) KI [-0,33; 0,12] p=0,30
	TB (n=9)	1.20 (0,41)	1,37 (0,40)	0,17 (0,14)	
<i>Interaktion p=0,16*</i>					
FEV₁ AUC₁₂₋₂₄					
Tag 1					
Stufe II	AB (n=18)	1.63 (0,44)	1,83 (0,52)	0,23 (0,22)	0,09 (0,03) KI [0,02; 0,16] p=0,02
	TB (n=18)	1.63 (0,43)	1,83 (0,51)	0,16 (0,22)	
Stufe III	AB (n=10)	1.15 (0,47)	1,28 (0,34)	0,13 (0,35)	0,10 (0,05) KI [-0,02; 0,22] p=0,09
	TB (n=9)	1.20 (0,41)	1,15 (0,23)	-0,06 (0,27)	
<i>Interaktion p=0,44</i>					
Tag 15					
Stufe II	AB (n=19)	1.63 (0,44)	1,82 (0,55)	0,19 (0,33)	0,07 (0,04) KI [-0,02; 0,16] p=0,11
	TB (n=19)	1.63 (0,43)	1,76 (0,54)	0,13 (0,25)	
Stufe III	AB (n=9)	1.15 (0,47)	1,23 (0,26)	0,21 (0,16)	0,06 (0,04) KI [-0,03; 0,15] p=0,18
	TB (n=9)	1.20 (0,41)	1,17 (0,25)	-0,04 (0,22)	
<i>Interaktion p=0,79</i>					

FEV₁ AUC₀₋₂₄					
Tag 1					
Stufe II	AB (n=19)	1,63 (0,44)	1,87 (0,52)	0,25 (0,20)	0,06 (0,03)
	TB (n=19)	1,63 (0,43)	1,82 (0,51)	0,19 (0,19)	KI [-0,004; 0,12] p=0,06
Stufe III	AB (n=10)	1,15 (0,47)	1,30 (0,36)	0,15 (0,28)	0,07 (0,05)
	TB (n=9)	1,20 (0,41)	1,21 (0,28)	0,01 (0,22)	KI [-0,03; 0,17] p=0,15
<i>Interaktion p=0,54</i>					
Tag 15					
Stufe II	AB (n=19)	1,63 (0,44)	1,87 (0,55)	0,25 (0,31)	0,04 (0,04)
	TB (n=19)	1,63 (0,43)	1,84 (0,56)	0,21 (0,26)	KI [-0,03; 0,12] p=0,26
Stufe III	AB (n=9)	1,15 (0,47)	1,25 (0,24)	0,23 (0,15)	0,07 (0,05)
	TB (n=9)	1,20 (0,41)	1,27 (0,32)	0,06 (0,15)	KI [-0,04; 0,17] p=0,18
<i>Interaktion p=0,78</i>					

M/34273/39					
FEV₁ AUC₀₋₁₂					
Tag 1					
Stufe II	AB (n=108)	1,65 (0,49)	1,81 (0,52)	0,16 (0,20)	0,01 (0,02)
	TB (n=104)	1,76 (0,49)	1,90 (0,52)	0,14 (0,19)	KI [-0,03; 0,06] p=0,56
Stufe III	AB (n=63)	1,14 (0,26)	1,26 (0,33)	0,11 (0,16)	0,01 (0,03)
	TB (n=53)	1,11 (0,32)	1,21 (0,34)	0,10 (0,14)	KI [-0,04; 0,06] p=0,68
<i>Interaktion p=0,18*</i>					
6 Wochen					
Stufe II	AB (n=108)	1,65 (0,49)	1,78 (0,52)	0,13 (0,24)	0,000 (0,03)
	TB (n=104)	1,76 (0,49)	1,88 (0,53)	0,12 (0,25)	KI [-0,06; 0,06] p=1,00
Stufe III	AB (n=63)	1,14 (0,26)	1,24 (0,33)	0,10 (0,18)	-0,04 (0,04)
	TB (n=53)	1,11 (0,32)	1,24 (0,34)	0,13 (0,20)	KI [-0,10; 0,03] p=0,32
<i>Interaktion p=0,54</i>					
FEV₁ AUC₁₂₋₂₄					
Tag 1					
Stufe II	AB (n=107)	1,65 (0,49)	1,75 (0,52)	0,10 (0,23)	0,07 (0,03)
	TB (n=104)	1,76 (0,49)	1,78 (0,52)	0,02 (0,22)	KI [0,01; 0,13] p=0,02
Stufe III	AB (n=62)	1,14 (0,26)	1,22 (0,32)	0,08 (0,17)	0,07 (0,03)
	TB (n=52)	1,11 (0,32)	1,12 (0,30)	0,01 (0,13)	KI [0,02; 0,12] p=0,01
<i>Interaktion p=0,36</i>					
6 Wochen					
Stufe II	AB (n=108)	1,65 (0,49)	1,70 (0,50)	0,05 (0,22)	0,04 (0,03)
	TB (n=104)	1,76 (0,49)	1,75 (0,52)	-0,01 (0,27)	KI [-0,02; 0,11] p=0,18
Stufe III	AB (n=62)	1,14 (0,26)	1,18 (0,31)	0,04 (0,16)	0,04 (0,03)
	TB (n=53)	1,11 (0,32)	1,12 (0,29)	0,01 (0,19)	KI [-0,03; 0,10] p=0,25
<i>Interaktion p=0,83</i>					
FEV₁ AUC₀₋₂₄					
Tag 1					
Stufe II	AB (n=107)	1,65 (0,49)	1,78 (0,51)	0,13 (0,20)	0,04 (0,03)
	TB (n=104)	1,76 (0,49)	1,84 (0,51)	0,08 (0,19)	KI [-0,01; 0,09] p=0,12

Stufe III	AB (n=62)	1,14 (0,26)	1,24 (0,32)	0,10 (0,16)	0,04 (0,03) KI [-0,01; 0,09]
	TB (n=53)	1,11 (0,32)	1,16 (0,31)	0,05 (0,12)	p=0,10
<i>Interaktion p=0,25</i>					
6 Wochen					
Stufe II	AB (n=108)	1,65 (0,49)	1,74 (0,51)	0,09 (0,22)	0,02 (0,03) KI [-0,04; 0,08]
	TB (n=104)	1,76 (0,49)	1,82 (0,53)	0,06 (0,25)	p= 0,49
Stufe III	AB (n=62)	1,14 (0,26)	1,21 (0,32)	0,07 (0,17)	0,001 (0,03) KI [-0,06; 0,06]
	TB (n=53)	1,11 (0,32)	1,18 (0,31)	0,07 (0,19)	p=0,98
<i>Interaktion p=0,68</i>					
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion, der aufgrund der Gleichgerichtetheit der Effekte in den Subgruppen verworfen werden kann AB: Acclidiniumbromid, TB: Tiotropiumbromid, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall					

Tabelle 4.3-36: Subgruppenergebnisse für Trough-FEV₁ (l) aus RCT mit direktem Vergleich

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
M/34273/23					
Tag 1					
Stufe II	AB (n=19)	1,63 (0,44)	1,82 (0,51)	0,19 (0,22)	0,04 (0,05) KI [-0,06; 0,13] p=0,45
	TB (n=19)	1,63 (0,43)	1,79 (0,51)	0,16 (0,18)	
Stufe III	AB (n=10)	1,15 (0,47)	1,28 (0,30)	0,12 (0,39)	0,07 (0,06) KI [-0,06; 0,20] p=0,25
	TB (n=9)	1,20 (0,41)	1,16 (0,24)	-0,04 (0,25)	
<i>Interaktion p=0,62</i>					
Tag 15					
Stufe II	AB (n=19)	1,63 (0,44)	1,80 (0,55)	0,17 (0,30)	0,04 (0,04) KI [-0,04; 0,12] p=0,34
	TB (n=19)	1,63 (0,43)	1,78 (0,53)	0,15 (0,23)	
Stufe III	AB (n=10)	1,15 (0,47)	1,28 (0,30)	0,13 (0,42)	-0,002 (0,05) KI [-0,12; 0,11] p=0,97
	TB (n=9)	1,20 (0,41)	1,20 (0,27)	-0,001 (0,22)	
<i>Interaktion p=0,77</i>					
M/34273/39					
Tag 1					
Stufe II	AB (n=108)	1,65 (0,49)	1,73 (0,50)	0,08 (0,22)	0,04 (0,03) KI [-0,03; 0,10] p=0,26
	TB (n=104)	1,76 (0,49)	1,80 (0,52)	0,04 (0,26)	
Stufe III	AB (n=63)	1,15 (0,26)	1,24 (0,32)	0,09 (0,19)	0,07 (0,03) KI [0,02; 0,13] p=0,01
	TB (n=53)	1,11 (0,32)	1,13 (0,30)	0,02 (0,13)	
<i>Interaktion p=0,35</i>					
6 Wochen					
Stufe II	AB (n=108)	1,65 (0,49)	1,68 (0,51)	0,04 (0,21)	0,03 (0,03) KI [-0,03; 0,10] p=0,34
	TB (n=104)	1,76 (0,49)	1,75 (0,52)	-0,006 (0,28)	
Stufe III	AB (n=63)	1,14 (0,26)	1,20 (0,32)	0,06 (0,19)	0,06 (0,03) KI [-0,01; 0,13] p=0,08
	TB (n=53)	1,11 (0,32)	1,11 (0,32)	-0,002 (0,20)	
<i>Interaktion p=0,90</i>					
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell AB: Acclidiniumbromid, TB: Tiotropiumbromid, SD: Standardabweichung, SE: Standardfehler, CFB:					

Change from Baseline, KI: Konfidenzintervall

Tabelle 4.3-37: Subgruppenergebnisse für Peak-FEV₁ (l) aus RCT mit direktem Vergleich

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
M/34273/23					
1. Tageshälfte (± 9:00 bis ± 21:00 Uhr)					
Tag 1					
Stufe II	AB (n=19)	1,63 (0,44)	2,02 (0,53)	0,40 (0,20)	0,04 (0,03) KI [-0,02; 0,10] p=0,22
	TB (n=19)	1,63 (0,43)	1,99 (0,54)	0,36 (0,19)	
Stufe III	AB (n=10)	1,15 (0,47)	1,46 (0,45)	0,31 (0,20)	0,05 (0,05) KI [-0,05; 0,16] p=0,31
	TB (n=9)	1,20 (0,41)	1,42 (0,40)	0,22 (0,18)	
<i>Interaktion p=0,80</i>					
Tag 15					
Stufe II	AB (n=19)	1,63 (0,44)	2,10 (0,58)	0,47 (0,31)	0,07 (0,04) KI [-0,02; 0,16] p=0,13
	TB (n=19)	1,63 (0,43)	2,05 (0,60)	0,42 (0,29)	
Stufe III	AB (n=10)	1,15 (0,47)	1,46 (0,27)	0,31 (0,35)	-0,08 (0,09) KI [-0,27; 0,12] p=0,40
	TB (n=9)	1,20 (0,41)	1,51 (0,46)	0,31 (0,14)	
<i>Interaktion p=0,08**</i>					
2. Tageshälfte (± 21:00 bis ± 9:00 Uhr)					
Tag 1					
Stufe II	AB (n=19)	1,63 (0,44)	2,00 (0,54)	0,38 (0,25)	0,06 (0,03) KI [-0,01; 0,12] p=0,10
	TB (n=19)	1,63 (0,43)	1,95 (0,57)	0,32 (0,23)	
Stufe III	AB (n=10)	1,15 (0,47)	1,45 (0,40)	0,29 (0,30)	0,15 (0,08) KI [-0,02; 0,31] p=0,07
	TB (n=9)	1,20 (0,41)	1,29 (0,24)	0,08 (0,25)	
<i>Interaktion p=0,23</i>					
Tag 15					
Stufe II	AB (n=19)	1,63 (0,44)	2,00 (0,61)	0,37 (0,38)	0,07 (0,04) KI [-0,006; 0,15] p=0,07
	TB (n=19)	1,63 (0,43)	1,94 (0,58)	0,31 (0,27)	

Stufe III	AB (n=10)	1,15 (0,47)	1,39 (0,32)	0,23 (0,46)	-0,04 (0,06) KI [-0,17; 0,10]
	TB (n=9)	1,20 (0,41)	1,35 (0,32)	0,14 (0,19)	p=0,58
<i>Interaktion p=0,30</i>					
M/34273/39					
1. Tageshälfte (± 9:00 bis ± 21:00 Uhr)					
Tag 1					
Stufe II	AB (n=108)	1,65 (0,49)	1,94 (0,52)	0,29 (0,19)	0,03 (0,02) KI [-0,01; 0,08]
	TB (n=104)	1,76 (0,49)	2,02 (0,54)	0,27 (0,19)	p=0,18
Stufe III	AB (n=63)	1,14 (0,26)	1,36 (0,34)	0,21 (0,16)	-0,01 (0,03) KI [-0,07; 0,04]
	TB (n=53)	1,11 (0,32)	1,33 (0,38)	0,22 (0,15)	p=0,61
<i>Interaktion p=0,26</i>					
6 Wochen					
Stufe II	AB (n=108)	1,65 (0,49)	1,93 (0,53)	0,28 (0,25)	0,03 (0,03) KI [-0,04; 0,09]
	TB (n=104)	1,76 (0,49)	2,01 (0,55)	0,25 (0,26)	p=0,38
Stufe III	AB (n=63)	1,14 (0,26)	1,36 (0,36)	0,21 (0,20)	-0,02 (0,04) KI [-0,10; 0,05]
	TB (n=53)	1,11 (0,32)	1,34 (0,39)	0,23 (0,22)	p=0,54
<i>Interaktion p=0,46</i>					
2. Tageshälfte (± 21:00 bis ± 9:00 Uhr)					
Tag 1					
Stufe II	AB (n=108)	1,65 (0,49)	1,84 (0,54)	0,19 (0,25)	0,09 (0,03) KI [0,03; 0,16]
	TB (n=104)	1,76 (0,49)	1,86 (0,54)	0,10 (0,23)	p=0,003
Stufe III	AB (n=63)	1,14 (0,26)	1,28 (0,34)	0,13 (0,19)	0,06 (0,03) KI [-0,002; 0,13]
	TB (n=53)	1,11 (0,32)	1,18 (0,33)	0,07 (0,17)	p=0,06
<i>Interaktion p=0,18*</i>					
6 Wochen					
Stufe II	AB (n=108)	1,65 (0,49)	1,80 (0,51)	0,15 (0,25)	0,04 (0,04) KI [-0,03; 0,11]
	TB (n=104)	1,76 (0,49)	1,87 (0,56)	0,11 (0,28)	p=0,27
Stufe III	AB (n=63)	1,14 (0,26)	1,26 (0,33)	0,12 (0,19)	0,02 (0,04) KI [-0,05; 0,09]
	TB (n=53)	1,11 (0,32)	1,21 (0,33)	0,10 (0,21)	p=0,57
<i>Interaktion p=0,63</i>					
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion, ** Beleg für					

Interaktion (kann in diesem Fall aufgrund der Gleichgerichtetheit der Effekte ignoriert werden)
 AB: Acclidiniumbromid, TB: Tiotropiumbromid, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall

Tabelle 4.3-38: Subgruppenergebnisse für COPD-Symptome aus RCT mit direktem Vergleich

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
M/34273/23					
Atemnot					
1 Woche					
Stufe II	AB (n=19)	1,65 (1,05)	1,32 (1,03)	-0,33 (0,69)	-0,01 (0,18) KI [-0,39; 0,36] p=0,95
	TB (n=19)	1,65 (1,05)	1,32 (1,10)	-0,33 (0,94)	
Stufe III	AB (n=10)	1,44 (1,02)	1,26 (0,82)	-0,18 (0,70)	-0,15 (0,17) KI [-0,50; 0,21] p=0,39
	TB (n=9)	1,27 (0,91)	1,48 (0,79)	0,21 (0,46)	
<i>Interaktion p=0,73</i>					
2 Wochen					
Stufe II	AB (n=19)	1,65 (1,05)	1,23 (1,04)	-0,42 (0,81)	-0,13 (0,22) KI [-0,58; 0,31] p=0,55
	TB (n=19)	1,65 (1,05)	1,34 (1,13)	-0,31 (1,00)	
Stufe III	AB (n=10)	1,44 (1,02)	1,35 (0,84)	-0,09 (0,76)	0,05 (0,22) KI [-0,41; 0,52] p=0,81
	TB (n=9)	1,27 (0,91)	1,33 (0,83)	0,06 (0,57)	
<i>Interaktion p=0,72</i>					
Husten					
1 Woche					
Stufe II	AB (n=19)	1,34 (0,98)	1,15 (0,95)	-0,19 (0,68)	-0,09 (0,18) KI [-0,46; 0,28] p=0,64
	TB (n=19)	1,34 (0,98)	1,24 (1,06)	-0,10 (0,79)	
Stufe III	AB (n=10)	1,28 (1,09)	1,21 (0,90)	-0,07 (0,45)	-0,16 (0,20) KI [-0,58; 0,26] p=0,43
	TB (n=9)	1,09 (0,97)	1,33 (0,78)	0,25 (0,54)	
<i>Interaktion p=0,66</i>					
2 Wochen					
Stufe II	AB (n=19)	1,34 (0,98)	1,07 (1,00)	-0,27 (0,80)	-0,24 (0,21) KI [-0,67; 0,20] p=0,27
	TB (n=19)	1,34 (0,98)	1,31 (1,00)	-0,03 (0,81)	
Stufe III	AB (n=10)	1,28 (1,09)	1,38 (0,84)	0,10 (0,83)	0,17 (0,21) KI [-0,28; 0,63] p=0,43
	TB (n=9)	1,09 (0,97)	1,22 (0,80)	0,13 (0,55)	

<i>Interaktion p=0,56</i>					
Auswurf					
1 Woche					
Stufe II	AB (n=19)	0,86 (0,77)	0,89 (0,82)	0,02 (0,45)	-0,06 (0,11)
	TB (n=19)	0,86 (0,77)	0,91 (0,81)	0,05 (0,32)	KI [-0,27; 0,16] p=0,59
Stufe III	AB (n=10)	0,93 (0,84)	0,83 (0,62)	-0,10 (0,33)	-0,27 (0,17)
	TB (n=9)	0,72 (0,52)	0,95 (0,60)	0,24 (0,39)	KI [-0,63; 0,09] p=0,13
<i>Interaktion p=0,18*</i>					
2 Wochen					
Stufe II	AB (n=19)	0,86 (0,77)	0,82 (0,83)	-0,05 (0,53)	-0,31 (0,13)
	TB (n=19)	0,86 (0,77)	1,08 (0,88)	0,22 (0,48)	KI [-0,58; -0,04] p=0,03
Stufe III	AB (n=10)	0,93 (0,84)	1,00 (0,80)	0,07 (0,62)	-0,05 (0,23)
	TB (n=9)	0,72 (0,52)	0,85 (0,50)	0,14 (0,41)	KI [-0,52; 0,42] p=0,82
<i>Interaktion p=0,10*</i>					
Nächtliche Symptome					
1 Woche					
Stufe II	AB (n=19)	0,78 (0,72)	0,57 (0,72)	-0,21 (0,52)	-0,14 (0,12)
	TB (n=19)	0,78 (0,72)	0,72 (0,83)	-0,06 (0,61)	KI [-0,39; 0,11] p=0,25
Stufe III	AB (n=10)	0,43 (0,64)	0,36 (0,67)	-0,07 (0,22)	-0,11 (0,10)
	TB (n=9)	0,43 (0,68)	0,48 (0,66)	0,05 (0,38)	KI [-0,32; 0,10] p=0,28
<i>Interaktion p=0,49</i>					
2 Wochen					
Stufe II	AB (n=19)	0,78 (0,72)	0,61 (0,78)	-0,17 (0,63)	-0,11 (0,14)
	TB (n=19)	0,78 (0,72)	0,72 (0,83)	-0,06 (0,61)	KI [-0,39; 0,18] p=0,46
Stufe III	AB (n=10)	0,43 (0,64)	0,37 (0,68)	-0,06 (0,24)	-0,08 (0,15)
	TB (n=9)	0,43 (0,68)	0,41 (0,50)	-0,02 (0,50)	KI [-0,40; 0,25] p=0,63
<i>Interaktion p=0,97</i>					
M/34273/39					
Atemnot (E-RS)					
6 Wochen					
Stufe II	AB (n=104)	6,04 (3,08)	5,12 (3,14)	-0,86 (1,89)	-0,27 (0,25)
	TB (n=99)	5,14 (3,02)	4,79 (3,24)	-0,42 (1,84)	KI [-0,77; 0,24] p=0,30
Stufe III	AB (n=62)	7,21 (3,47)	6,27 (3,82)	-0,94 (2,30)	-0,64 (0,40)
	TB (n=53)	7,71 (3,04)	7,35 (3,45)	-0,35 (2,00)	KI [-1,43; 0,16] p=0,12

						<i>Interaktion p=0,38</i>
Husten & Auswurf (E-RS)						
6 Wochen						
Stufe II	AB (n=104)	3,87 (1,60)	3,12 (1,55)	-0,75 (1,15)	-0,18 (0,14)	KI [-0,46; 0,11] p=0,22
	TB (n=99)	3,34 (1,55)	2,92 (1,47)	-0,39 (1,14)		
Stufe III	AB (n=62)	3,97 (1,64)	3,47 (1,82)	-0,50 (1,23)	-0,25 (0,21)	KI [-0,67; 0,17] p=0,24
	TB (n=53)	4,00 (1,44)	3,73 (1,52)	-0,27 (1,03)		
						<i>Interaktion p=0,71</i>
Brustenge (E-RS)						
6 Wochen						
Stufe II	AB (n=104)	3,13 (1,63)	2,56 (1,66)	-0,55 (1,12)	-0,11 (0,15)	KI [-0,41; 0,18] p=0,45
	TB (n=99)	2,65 (1,52)	2,37 (1,56)	-0,31 (1,13)		
Stufe III	AB (n=62)	3,46 (1,91)	3,07 (2,06)	-0,39 (1,30)	-0,23 (0,23)	KI [-0,68; 0,22] p=0,31
	TB (n=53)	3,39 (1,68)	3,23 (1,95)	-0,16 (1,13)		
						<i>Interaktion p=0,87</i>
Gesamt E-RS Score						
6 Wochen						
Stufe II	AB (n=104)	13,05 (5,65)	10,79 (5,82)	-2,16 (3,80)	-0,58 (0,50)	KI [-1,57; 0,41] p=0,25
	TB (n=99)	11,12 (5,53)	10,07 (5,69)	-1,12 (3,72)		
Stufe III	AB (n=62)	14,64 (6,07)	12,81 (6,85)	-1,83 (4,41)	-1,10 (0,75)	KI [-2,59; 0,39] p=0,15
	TB (n=53)	15,10 (5,35)	14,31 (6,22)	-0,78 (3,56)		
						<i>Interaktion p=0,76</i>
Tage ohne morgendliche Symptome						
6 Wochen						
Stufe II	AB (n=103)	10,36 (20,87)	16,17 (27,65)	6,16 (19,21)	1,85 (2,69)	KI [-3;45; 7,15] p=0,49
	TB (n=97)	17,94 (28,02)	20,82 (31,82)	3,01 (21,20)		
Stufe III	AB (n=62)	12,42 (26,41)	19,76 (32,86)	7,34 (23,55)	5,74 (3,50)	KI [-1,18; 12,66] p=0,10
	TB (n=53)	6,93 (16,03)	8,84 (20,67)	1,92 (11,81)		
						<i>Interaktion p=0,48</i>
Schwere der morgendlichen Symptome						
6 Wochen						
Stufe II	AB (n=103)	2,30 (0,55)	2,10 (0,61)	-0,19 (0,43)	-0,08 (0,06)	KI [-0,19; 0,04] p=0,18
	TB (n=97)	2,11 (0,60)	2,04 (0,62)	-0,07 (0,40)		
Stufe III	AB (n=62)	2,45 (0,69)	2,27 (0,77)	-0,19 (0,51)	-0,13 (0,09)	

	TB (n=53)	2,50 (0,60)	2,43 (0,64)	-0,06 (0,39)	KI [-0,30; 0,04] p=0,12 <i>Interaktion p=0,80</i>
Schwere der nächtlichen Symptome					
6 Wochen					
Stufe II	AB (n=103)	1,93 (0,65)	1,77 (0,65)	-0,18 (0,38)	-0,04 (0,05)
	TB (n=97)	1,81 (0,61)	1,71 (0,62)	-0,10 (0,43)	KI [-0,15 ; 0,06] p=0,41
Stufe III	AB (n=62)	2,08 (0,75)	1,92 (0,77)	-0,16 (0,50)	-0,13 (0,08)
	TB (n=53)	2,05 (0,70)	2,02 (0,75)	-0,03 (0,40)	KI [-0,30 ; 0,03] p=0,12 <i>Interaktion p=0,15</i>
Grad der Einschränkung der Aktivität durch COPD-Symptome					
6 Wochen*					
Stufe II	AB (n=103)	1,91 (0,67)	1,74 (0,67)	-0,17 (0,40)	-0,11 (0,05)
	TB (n=97)	1,76 (0,66)	1,75 (0,72)	-0,04 (0,37)	KI [-0,21 ; -0,00] p=0,04
Stufe III	AB (n=62)	2,18 (0,73)	2,07 (0,80)	-0,10 (0,48)	-0,10 (0,09)
	TB (n=53)	2,26 (0,74)	2,25 (0,75)	-0,01 (0,43)	KI [-0,27 ; 0,08] p=0,27 <i>Interaktion p=0,54</i>
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion AB: Aclidiniumbromid, TB: Tiotropiumbromid, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall					

Tabelle 4.3-39: Subgruppenergebnisse für die Behandlungszufriedenheit aus RCT mit einem direkten Vergleich

Studie				
Frage	Antwort	Genuair®	Handihaler®	keine Präferenz
M/34273/23				
Einfache Bedienung (in %)				
Stufe II (n=19)				
	Sehr einfach	16 (84,2%)	10 (52,6%)	-
	Einfach	2 (10,5%)	5 (26,3%)	-
	Normal	0 (0,0%)	4 (21,1%)	-
	Schwer	0 (0,0%)	0 (0,0%)	-
	Sehr schwer	1 (5,3%)	0 (0,0%)	-
	<i>Interaktion</i>	<i>p=0,62</i>	<i>p=0,76</i>	
Stufe III (n=10)				
	Sehr einfach	8 (80,0%)	6 (60,0%)	-
	Einfach	2 (20,0%)	3 (30,0%)	-
	Normal	0 (0,0%)	1 (10,0%)	-
	Schwer	0 (0,0%)	0 (0,0%)	-
	Sehr schwer	0 (0,0%)	0 (0,0%)	-
	<i>Interaktion</i>	<i>p=0,37</i>	<i>p=0,34</i>	
Einfache Vorbereitung der Inhalation (in %)				
Stufe II (n=19)				
	Sehr einfach	17 (89,5%)	8 (42,1%)	-
	Einfach	2 (10,5%)	9 (47,4%)	-
	Normal	0 (0,0%)	2 (10,5%)	-
	Schwer	0 (0,0%)	0 (0,0%)	-
	Sehr schwer	0 (0,0%)	0 (0,0%)	-
Stufe III (n=10)				
	Sehr einfach	8 (80,0%)	6 (60,0%)	-
	Einfach	1 (10,0%)	2 (20,0%)	-
	Normal	1 (10,0%)	2 (20,0%)	-
	Schwer	0 (0,0%)	0 (0,0%)	-
	Sehr schwer	0 (0,0%)	0 (0,0%)	-
	<i>Interaktion</i>	<i>p=0,34</i>		

Bevorzugung (in %)				
Stufe II (n=19)				
	Ich präferiere maßgeblich	5 (26,3%)	1 (5,3%)	10 (52,6%)
	Ich präferiere etwas	3 (15,8%)	0 (0,0%)	
Stufe III (n=10)				
	Ich präferiere maßgeblich	3 (30,0%)	1 (10,0%)	2 (20,0%)
	Ich präferiere etwas	3 (30,0%)	1 (10,0%)	
<i>Interaktion p=0,35</i>				
M/34273/39				
Bevorzugung aufgrund der einfachen Bedienung (in %)				
Stufe II (n=268)				
	Ich präferiere	227 (84,7%)*	26 (9,7%)	15 (5,6%)
Stufe III (n=143)				
	Ich präferiere	129 (90,2%)*	11 (7,7%)	3 (2,1%)
<i>Interaktion p=19**</i>				
Bevorzugung aufgrund der einfachen Vorbereitung der Inhalation (in %)				
Stufe II (n=268)				
	Ich präferiere	230 (85,8%)*	21 (7,8%)	17 (6,3%)
Stufe III (n=143)				
	Ich präferiere	125 (87,4%)*	11 (7,7%)	7 (4,9%)
<i>Interaktion p=83</i>				
Bevorzugung aufgrund der einfachen Handhabung (in %)				
Stufe II (n=268)				
	Ich präferiere	215 (80,2%)*	27 (10,1%)	26 (9,7%)
Stufe III (n=143)				
	Ich präferiere	118 (82,5%)*	11 (7,7%)	14 (9,8%)
<i>Interaktion p=0,73</i>				
Bevorzugung aufgrund der einfachen Durchführung (in %)				
Stufe II (n=268)				
	Ich präferiere	212 (79,1%)*	21 (7,8%)	35 (13,1%)
Stufe III (n=143)				
	Ich präferiere	117 (81,8%)*	14 (9,8%)	12 (8,4%)
<i>Interaktion p=0,32</i>				
Bevorzugung bezüglich des Feedbacksystems (in %)				
Stufe II (n=268)				
	Ich präferiere	204 (76,1%)*	25 (9,3%)	39 (14,6%)
Stufe III (n=143)				
	Ich präferiere	109 (76,2%)*	14 (9,8%)	20 (14,0%)

<i>Interaktion p=98</i>				
Bevorzugung aufgrund der einfachen Erlernbarkeit (in %)				
Stufe II (n=268)	Ich präferiere	209 (78,0%)*	20 (7,5%)	39 (14,6%)
Stufe III (n=143)	Ich präferiere	117 (81,8%)*	13 (9,1%)	13 (9,1%)
<i>Interaktion p=0,26</i>				
Bevorzugung aufgrund des Komforts (in %)				
Stufe II (n=268)	Ich präferiere	231 (86,2%)*	22 (8,2%)	15 (5,6%)
Stufe III (n=143)	Ich präferiere	128 (89,5%)*	10 (7,0%)	5 (3,5%)
<i>Interaktion p=0,57</i>				
Bevorzugung (in %)				
Stufe II (n=268)	Ich präferiere	214 (79,9%)*	29 (10,8%)	25 (9,3%)
Stufe III (n=143)	Ich präferiere	115 (80,4%)*	15 (10,5%)	13 (9,1%)
<i>Interaktion p=0,99</i>				
Bereitschaft der Patienten, den Inhalatoren weiterzuverwenden; Score¹ (SD)				
Stufe II (n=267)	Ich präferiere	89,3 (21,6)*	47,6 (35,5)	-
Stufe III (n=142)	Ich präferiere	88,0 (23,1)*	41,0 (35,1)	-
<i>Interaktion p=0,25</i>				
* signifikanter (p<0,001) Unterschied zum HandiHaler [®] , **Hinweis auf Interaktion				
¹ Mittelwert; 0=keine Bereitschaft, 100=eindeutige Bereitschaft				

Tabelle 4.3-40: Subgruppenergebnisse für unerwünschte Arzneimittelwirkungen aus RCT mit einem direkten Vergleich

Studie	M/34273/23		M/34273/39	
	Aclidinium- bromid n/n _{ges} (%)	Tiotropium- bromid n/n _{ges} (%)	Aclidinium- bromid n/n _{ges} (%)	Tiotropium- bromid n/n _{ges} (%)
Unerwünschte Arzneimittelwirkungen (AEs)				
AEs vor der Therapie				
Stufe II	1/19 (5,3%)		6/108 (5,6%)	7/104 (6,7%)
Stufe III	0	0	6/63 (9,5%)	4/53 (7,5%)
AEs während der Therapie				
Stufe II	4/19 (21,1%)	1/19 (5,3%)	25/108 (23,1%)	31/104 (29,8%)
Stufe III	3/10 (30,0%)	2/9 (22,2%)	22/63 (34,9%)	15/53 (28,3%)
<i>Interaktion p=0,24</i>				
AEs, die zum Abbruch der Therapie führten				
Stufe II	0	0	1/108 (0,9%)	1/104 (1,0%)
Stufe III	0	0	2/63 (3,2%)	1/53 (1,9%)
<i>Interaktion p=0,89</i>				
Schwere Unerwünschte Arzneimittelwirkungen (SAEs)				
SAEs vor der Therapie				
Stufe II	0	0	0	0
Stufe III	0	0	1/63 (1,6%)	0
SAEs während der Therapie				
Stufe II	0	0	2/108 (1,9%)	2/104 (1,9%)
Stufe III	0	0	1/63 (1,6%)	2/53 (3,8%)
SAEs, die zum Abbruch der Therapie führten				
Stufe II	0	0	0	0
Stufe III	0	0	0	1/53 (1,9%)

Wie in den vorangegangenen Tabellen dargestellt, zeigten sich für einen Großteil der verfügbaren Daten keine unterschiedlichen Effekte in den Subgruppen. Es ergaben sich acht Fallkonstellationen, die Hinweise (p-Wert zwischen 0,05 und 0,2) auf unterschiedliche Effekte erbrachten. Die in den Tabellen gekennzeichneten Fälle hatten keine Auswirkungen auf das Fazit der Nutzen- und Zusatznutzenbewertung.

In sechs Fällen handelte es sich um Hinweise auf Interaktionen, die durch die Gleichgerichtetheit der Effekte in den Subgruppen und der Gesamtpopulation (jeweils nicht signifikant) verworfen werden konnten. Der Hinweis auf eine Interaktion (p=0,18) beim Endpunkt Peak-FEV₁ am ersten Tag in der zweiten Tageshälfte der Studie M/34273/39 mit leicht entgegengesetzten (p=0,0003 bzw. 0,06) Effekten kann zum Messzeitpunkt nach 6 Wochen nicht mehr festgestellt werden (p=0,65) und wird als Ausreißer betrachtet.

Der Hinweis auf eine Interaktion durch den Schweregrad beim COPD-Symptom ‚Auswurf‘ in der zweiten Woche der Studie M/34273/23 ist auf entgegengesetzte Effekte in den Subgruppen zurückzuführen. Da die Effekte in der ersten Woche gleichgerichtet sind wird hier von einem Ausreißer und nicht von einer systematischen Effektmodifikation ausgegangen. Des Weiteren ist die Interaktion nur in der Studie M/34273/23 zu beobachten und wird nicht von dem Ergebnis des vergleichbaren COPD-Symptoms ‚Husten & Auswurf‘ der größeren Studie M/34273/39 bestätigt (p=0,71).

Für die Endpunktkategorie Mortalität können aufgrund der geringen Zahl von Ereignissen keine Subgruppenergebnisse dargestellt werden. Für die unerwünschten Arzneimittelwirkungen (M/34273/23) und die schweren unerwünschten Ereignisse (M/34273/39) können aufgrund der geringen Zahl von Ereignissen keine Interaktionstests durchgeführt werden. Beim Vergleich der Ereignisse in den Subgruppen können keine größeren numerischen Unterschiede festgestellt werden.

A priori geplante Subgruppenanalysen wurden in den direkten Vergleichsstudien nicht durchgeführt und können nicht dargestellt werden.

4.3.1.5 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen aus randomisierten kontrollierten Studien geben. Die Zusammenfassung soll Aussagen zu allen in Abschnitt 4.3.1 präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus randomisierten kontrollierten Studien zusammen.

Mit Acridiniumbromid wird durch die zweimal tägliche Gabe eine konstante Bronchodilatation über 24 Stunden ermöglicht. Die Veränderung zur Baseline (CFB) der Fläche unter der Kurve (AUC) des FEV₁-Wertes ist in der M/34273/23 zu jedem Messzeitpunkt (Tag 1 und 15, 0-12, 12-24 und 0-24 Stunden) klinisch relevant (≥ 100 ml). In der Studie M/34273/39 erreichen weder die Patienten unter Acridiniumbromid, noch die unter Tiotropiumbromid so hohe Veränderungswerte wie in der Studie M/34273/23. Die CFB der Acridiniumbromid-Patienten ist mit der der Tiotropiumbromid-Patienten aber zu jedem Messzeitpunkt mindestens gleichwertig.

Das Bronchodilatationsprofil von Acridiniumbromid ist in der ersten Tageshälfte und über 24 Stunden mit dem von Tiotropiumbromid gleichwertig; die CFB an Tag 15 liegt in der M/34273/23 sowohl bei Acridiniumbromid (AB) als auch bei Tiotropiumbromid (TB) im klinisch relevanten Bereich (CFB_{AB} FEV₁ AUC_{0-24, 0-12} = 240 ml, CFB_{TB} FEV₁ AUC₀₋₂₄ = 160 ml, CFB_{TB} FEV₁ AUC₀₋₁₂ = 250. In der Studie M/34273/39 ist dieser Effekt, wie bereits erwähnt, nicht so stark sichtbar.

In der zweiten Tageshälfte wird in der Studie M/34273/23 durch die zweite Inhalation von Acridiniumbromid eine signifikant überlegene Bronchodilatation erreicht (CFB_{AB} FEV₁ AUC₁₂₋₂₄ = 200 ml, CFB_{TB} FEV₁ AUC₁₂₋₂₄ = 80 ml; p = 0,02). In der Studie M/34273/39 ist dieser Effekt am ersten Tag signifikant (Tag 1: CFB_{AB} FEV₁ AUC₁₂₋₂₄ = 90 ml, CFB_{TB} FEV₁ AUC₁₂₋₂₄ = 20 ml; p = 0,002), in der sechsten Woche numerisch (CFB_{AB} FEV₁ AUC₁₂₋₂₄ = 50 ml, CFB_{TB} FEV₁ AUC₁₂₋₂₄ = 1 ml; p = 0,12) feststellbar. Die durch den Endpunkt FEV₁ AUC analysierte Entwicklung der Lungenfunktion im Tagesverlauf belegt die Konstanz der Bronchodilatation unter der Acridiniumbromid-Behandlung. Sie kann durch die zweimal tägliche Gabe die durch die zirkadiane Rhythmik verursachte stärkere Ausprägung der belastenden COPD-Symptome in der Nacht und am Morgen auffangen.

Bessere prä-bronchodilatatorische Trough-FEV₁-Werte (Unterschied am ersten Tag in der Studie M/34273/39 signifikant, in der Studie M/34273/23 numerisch, am Ende der beiden Studien jeweils numerisch) zeigen eine Verbesserung der Lungenfunktion.

Die zu zwei Messzeitpunkten signifikanten und ansonsten numerisch größeren Peak-FEV₁-Werte an Tag 1 zeigen gegenüber Tiotropiumbromid zeigen die sofortige Wirkung der Acridiniumbromid-Behandlung ab dem ersten Tag.

Während der zwei- bzw. sechswöchigen Behandlung mit Acclidiniumbromid konnten - ausgehend von der Baselineerhebung - alle Ausprägungen der COPD-Symptome verringert werden. Gegenüber Tiotropiumbromid wurden die Veränderungen als numerisch überlegen beobachtet. Die Einschränkung der körperlichen Aktivität durch COPD-Symptome ist bei Acclidiniumbromid-Patienten signifikant ($p=0,04$) geringer als bei Tiotropiumbromid-Patienten. Dies stellt einen Nutzen für den Patienten und einen Zusatznutzen gegenüber Tiotropiumbromid dar.

Unerwünschte Arzneimittelwirkungen traten in beiden, schwere unerwünschte Arzneimittelwirkungen nur in der längeren Studie auf. Das Nebenwirkungsprofil ist mit dem von Tiotropiumbromid vergleichbar. Aussagen zur kardiovaskulären Morbidität und Mortalität waren aufgrund der relativ kurzen Studiendauer kaum bzw. nicht möglich. Die berichteten Ereignisse lassen auf ein vergleichbares Sicherheitsprofil schließen.

In den zwei- bzw. sechswöchigen Studien wurde ein wichtiger Aspekt der Lebensqualität - die Behandlungszufriedenheit - mit Hilfe der Patientenpräferenz für das eine oder andere Inhalationssystem beschrieben. Dabei wurden u. a. die einfache Handhabung und die allgemeine Bevorzugung eines Inhalators erfragt. Dabei zeigt sich ein signifikanter Vorteil für den Genuair[®] Inhalator gegenüber dem HandiHaler[®]: 80% der Patienten bevorzugten das Inhalationssystem Genuair[®].

Die Ergebnisse gelten aufgrund der im Abschnitt 4.3.1.4 festgestellten Vergleichbarkeit der Subgruppen für die gesamte Zielpopulation der COPD-Patienten Stufe II-IV.

Meta-Analysen wurden aufgrund der heterogenen Studiencharakteristik (Studiendesign, Studiendauer, Patientenzahl) nicht durchgeführt.

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

Die zwei in Abschnitt 4.3.1 herangezogenen RCTs mit einem direkten Vergleich des zu bewertenden Arzneimittels und der Vergleichstherapie erlauben Aussagen zur Bewertung des Zusatznutzens. Zur Erhöhung der Aussagesicherheit (vgl. Abschnitt 4.5.1) wird zusätzlich ein indirekter Vergleich durchgeführt.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

4.3.2.1.1.1 Studien des pharmazeutischen Unternehmers

Benennen Sie in der nachfolgenden Tabelle alle Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war.

Von den randomisierten kontrollierten Studien, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden, sollen die Studien, deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle randomisierten kontrollierten Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.

Benennen Sie nur randomisierte, kontrollierte Studien, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

In der folgenden Tabelle 4.3-41 sind alle Studien des pharmazeutischen Unternehmers im Anwendungsgebiet COPD aufgelistet. Im ersten Teil der Tabelle sind die im Studienprogramm zuerst durchgeführten Studien mit der einmal täglichen Gabe aufgelistet. Im zweiten Abschnitt der Tabelle sind die Studien mit einer kombinierten Gabe von Aclidiniumbromid und Formoterol dargestellt. In den darauf folgenden Abschnitten sind die im Rahmen der Weiterentwicklung des Studienprogramms durchgeführten Studien mit der (zumindest in einem Therapiearm) zulassungskonformen zweimal täglichen Gabe gelistet. Die Tabelle schließt mit der Auflistung der Device Studien.

Tabelle 4.3-41: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studien- bezeichnung	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
<Studie 1>	ja	abgeschlossen	12 Monate	Medikament A, Medikament B, Placebo
Studien mit einmal täglicher Einnahme von Acridiniumbromid				
M/34273/00	nein	abgeschlossen	1 Tag	Acridiniumbromid 50, 300, 600 µg QD vs. Placebo (via Cycobaler®)
M/34273/01	nein	abgeschlossen	5 Tage	Acridiniumbromid 600-6000 µg QD vs. Placebo (via Cycobaler®)
M/34273/04	nein	abgeschlossen	1 Woche	Acridiniumbromid 400 µg QD [phenyl-U- ¹⁴ C] (IV) vs. Acridiniumbromid 400 µg QD [glycolyl-U- ¹⁴ C] (IV)
M/34273/05	nein	abgeschlossen	1 Woche	Teil I: Acridiniumbromid 25- 400 µg (IV) QD vs. Placebo (IV) Teil II: Acridiniumbromid 200 µg (IV), 200 µg QD (via Genuair®)
M/34273/06	nein	abgeschlossen	5 Tage	Acridiniumbromid 200, 400, 800 µg QD vs. Placebo (via Genuair®)
M/34273/08	nein	abgeschlossen	1 Tag	Acridiniumbromid 400 µg QD (via Genuair®) bei 4 Gruppen mit normaler Nierenfunktion, sowie leichter, mittlerer und schwerer Niereninsuffizienz
M/34273/09	nein	abgeschlossen	3 Tage	Acridiniumbromid 200, 400 µg QD (via Genuair®)
M/34273/11	nein	abgeschlossen	3 Tage	Acridiniumbromid 200, 800 µg QD (via Genuair®)
M/34273/21	nein	abgeschlossen	4 Tage	Acridiniumbromid 100, 300, 900 µg QD vs. Placebo (via Genuair®)
M/34273/22	nein	abgeschlossen	4 Wochen	Acridiniumbromid 25, 50, 100, 200, 400 µg QD vs. Placebo (via Genuair®)
M/34273/24	nein	abgeschlossen	1 Tag	Acridiniumbromid 200 µg QD (via Genuair®) vs. Placebo (via Genuair®) vs. Tiotropiumbromid 18 µg QD

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studien- bezeichnung	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
				(via HandiHaler®)
M/34273/25	nein	abgeschlossen	1 Woche	Acidiniumbromid 200 µg QAM/QPM vs. Placebo (via Genuair®)
LAS-MD-26	nein	abgeschlossen	6 Wochen	Acidiniumbromid 200 µg QD vs. Placebo (via Genuair®)
M/34273/30 (ACCLAIM I)	nein	abgeschlossen	52 Wochen	Acidiniumbromid 200 µg QD vs. Placebo (via Genuair®)
M/34273/31 (ACCLAIM II)	nein	abgeschlossen	52 Wochen	Acidiniumbromid 200 µg QD vs. Placebo (via Genuair®)
Studien mit der Kombinationstherapie aus Acidiniumbromid und Formoterol				
M/273FO/23	nein	abgeschlossen	4 Wochen	Acidiniumbromid 200 µg und Formoterol 6 µg QD vs. Acidiniumbromid 200 µg und Formoterol 12 µg QD vs. Acidiniumbromid 200 µg und Formoterol 18 µg QD vs. Acidiniumbromid 200 µg vs. Formoterol 12 µg QD vs. Placebo
M/40464/26	nein	abgeschlossen	4 Wochen	Acidiniumbromid 200 µg und Formoterol 6 µg BID vs. Acidiniumbromid 200 µg und Formoterol 12 µg BID vs. Acidiniumbromid 200 µg vs. Formoterol 12 µg BID vs. Placebo
LAC-MD-27	nein	abgeschlossen	2 Wochen	Acidiniumbromid 400 µg BID und Formoterol 6 µg BID vs. Acidiniumbromid 400 µg BID und Formoterol 12 µg BID vs. Acidiniumbromid 400 µg BID vs. Formoterol 12 µg BID vs. Placebo
Studien mit zweimal täglicher Einnahme				
M/34273/23	nein	abgeschlossen	15 Tage	Acidiniumbromid 400 µg BID (via Genuair®) vs. Placebo (via Genuair®) vs. Tiotropiumbromid 18 µg QD (via HandiHaler®)
M/34273/29	nein	abgeschlossen	1 Woche	Acidiniumbromid 100, 200,

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studien- bezeichnung	Zulassungs- studie (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
				400 µg BID (via Genuair [®]) vs. Placebo (via Genuair [®]) vs. Formoterol 12 µg BID (via Aerolizer [®])
LAS-MD-33 (ACCORD I)	ja	abgeschlossen	12 Wochen	Acclidiniumbromid 200, 400 µg BID vs. Placebo (via Genuair [®])
M/34273/34 (ATTAIN)	ja	abgeschlossen	24 Wochen	Acclidiniumbromid 200, 400 µg BID vs. Placebo (via Genuair [®])
LAS-MD-38 A (ACCORD II)	ja	abgeschlossen	12 Wochen	Acclidiniumbromid 200, 400 µg BID vs. Placebo (via Genuair [®])
M/34273/39	nein	abgeschlossen	6 Wochen	Acclidiniumbromid 400 µg BID (via Genuair [®]) vs. Placebo (via Genuair [®]) vs. Tiotropiumbromid 18 µg QD (via HandiHaler [®])
Langzeitstudien				
LAS-MD-35	nein	abgeschlossen	52 Wochen	Acclidiniumbromid 200, 400 µg BID (via Genuair [®])
LAS-MD-36	nein	abgeschlossen	52 Wochen	Acclidiniumbromid 200, 400 µg BID (via Genuair [®])
LAS-MD-38 B	nein	abgeschlossen	40 Wochen	Acclidiniumbromid 400 µg BID (via Genuair [®])
Laufende Studien				
M/34273/40	nein	laufend	3 Wochen	Acclidiniumbromid 400 µg BID vs. Placebo (via Genuair [®])
Device Studien				
M/34273/03	nein	abgeschlossen	1 Tag	Acclidiniumbromid 200 µg (via Genuair [®])
M/34273/07	nein	abgeschlossen	1 Tag	Placebo (via Genuair [®]) vs. Placebo (via HandiHaler [®])
M/34273/32	nein	abgeschlossen	1 Tag	Placebo (via Genuair [®]) vs. Placebo (via HandiHaler [®])
QD = einmal täglich, BID = zweimal täglich, QAM = einmal morgens, QPM = einmal abends, IV = intravenös, µg = Mikrogramm				

Geben Sie an, welchen Stand die Information in Tabelle 4.3-41 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus sollte nicht mehr als 3 Monate vor dem für die Einreichung des Dokuments maßgeblichen Zeitpunkt liegen.

Die Informationen in Tabelle 4.3-41 stellen den Stand vom 09.07.2012 dar.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4.3-41 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Unter den Studien des pU (Tabelle 4.3-41) befinden sich Studien, die alle Einschlusskriterien-inklusive der empfohlenen Studiendauer von ≥ 12 Wochen- erfüllen. Deshalb werden, wie in Abschnitt 4.2.2 geplant, kürzere Studien von der Nutzenbewertung mit Hilfe des indirekten Vergleiches ausgeschlossen.

Die folgende Tabelle 4.3-42 zeigt die Studien des pharmazeutischen Unternehmers, die nicht für den indirekten Vergleich berücksichtigt wurden. Bei der Begründung der Nichtberücksichtigung der Studie ist bei Zutreffen mehrerer Ausschlussgründe lediglich der primäre Grund aufgeführt.

Tabelle 4.3-42: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel für indirekte Vergleiche

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
M/34273/00	Population (es handelt sich um männliche gesunde Freiwillige)
M/34273/01	Population (es handelt sich um männliche gesunde Freiwillige)
M/34273/03	Prüf- und Vergleichstherapie (es handelt sich um eine Device Studie)
M/34273/04	Population (es handelt sich um männliche gesunde Freiwillige)
M/34273/05	Population (es handelt sich um männliche gesunde Freiwillige)
M/34273/06	Population (es handelt sich um gesunde Freiwillige)
M/34273/07	Prüf- und Vergleichstherapie (es handelt sich um eine Device Studie)
M/34273/08	Vergleichstherapie (Ermittlung der Dositoleranz bei Niereninsuffizienz)
M/34273/09	Vergleichstherapie (es handelt sich um eine Dosisfindungsstudie)
M/34273/11	Population (es handelt sich um gesunde Freiwillige)
LAS-PK-12	Population (es handelt sich um gesunde Freiwillige)
M/34273/21	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/34273/22	Studiendauer (mit 4 Wochen kürzer als erforderlich)
M/34273/23	Studiendauer (mit 15 Tagen kürzer als erforderlich)
M/273FO/23	Studiendauer (mit 4 Wochen kürzer als erforderlich)
M/34273/24	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/34273/25	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/40464/26	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
LAS-MD-26	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
LAC-MD-27	Studiendauer (mit 2 Wochen kürzer als erforderlich)
M/34273/29	Studiendauer (mit 1 Woche kürzer als erforderlich)
M/34273/30	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/34273/31	Intervention (Dosierung von 200 µg QD entspricht nicht der Zulassung)
M/34273/32	Prüf- und Vergleichstherapie (es handelt sich um eine Device Studie)
LAS-MD-35	Vergleichstherapie (Studie ist nicht kontrolliert)
LAS-MD-36	Vergleichstherapie (Studie ist nicht kontrolliert)
LAS-MD-38 B	Vergleichstherapie (Studie ist nicht kontrolliert)
M/34273/39	Studiendauer (mit 6 Wochen kürzer als erforderlich)
M/34273/40	Laufende Studie - es liegen noch keine Ergebnisse vor
QD = einmal täglich, µg = Mikrogramm	

Nach Ausschluss der oben gelisteten Studien verbleiben 3 Studien im Studienpool. Dabei handelt es sich um die Studien LAS-MD-33 (ACCORD I), M/34273/34 (ATTAIN) und LAS-

MD-38A (ACCORD II). Die Akronyme der Studien werden im Folgenden zur Bezeichnung der Studien verwendet.

Stellen Sie die abgebrochenen und laufenden Studien, die nicht in die Bewertung eingehen, in Anhang 4-D und Anhang 4-E dar.

4.3.2.1.1.2 Studien aus der bibliografischen Literaturrecherche

Sofern eine bibliografische Literaturrecherche durchgeführt wurde, beschreiben Sie nachfolgend das Ergebnis dieser Recherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-C.

Falls die Recherche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel gemeinsam mit der Recherche zu einer anderen Fragestellung (z. B. zu indirekten Vergleichen) durchgeführt wurde, unterteilen Sie die Angaben zu relevanten Treffern und Studien entsprechend diesen Fragestellungen.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.

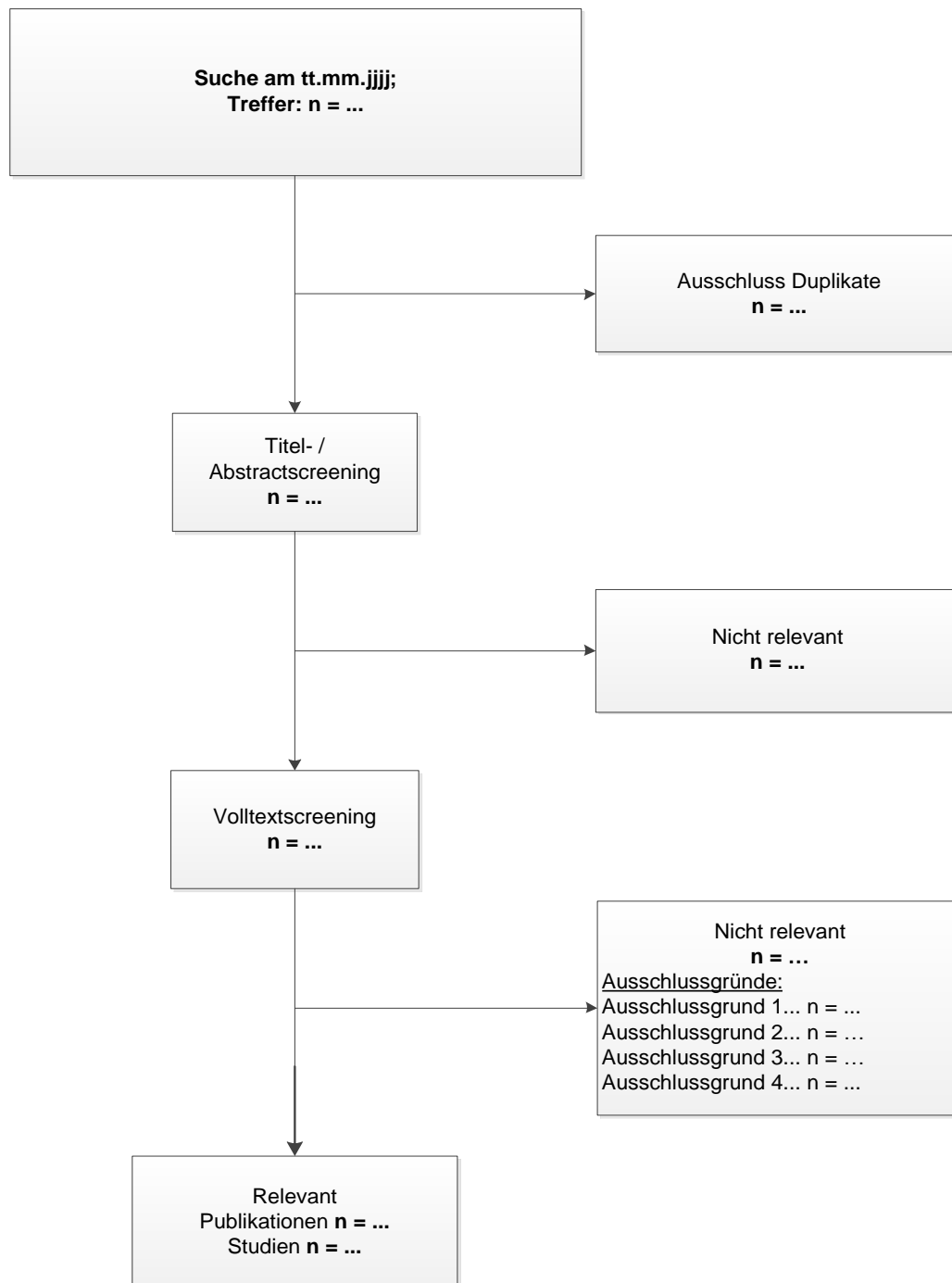


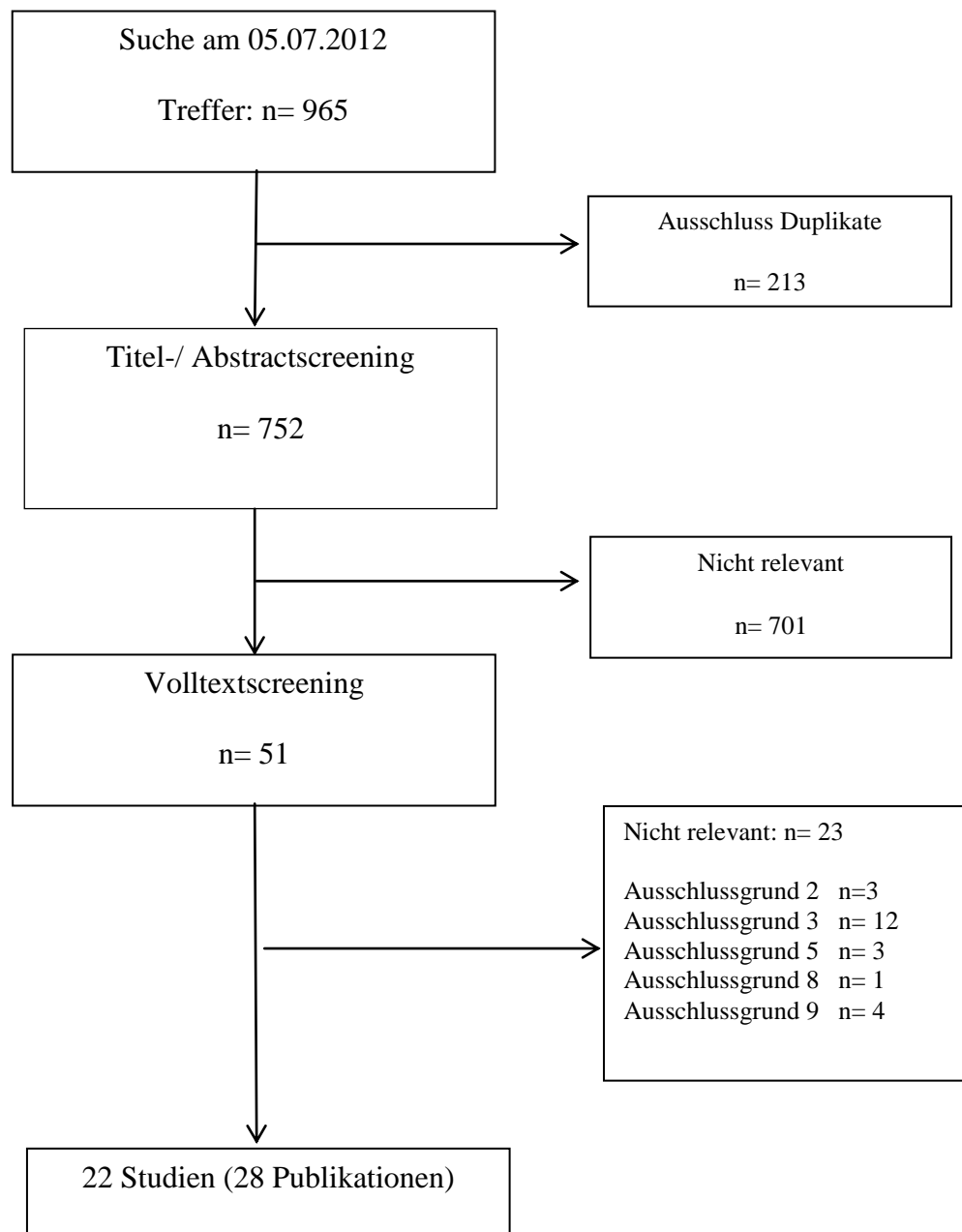
Abbildung 4: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Die bibliografische Literaturrecherche wurde in den Datenbanken MEDLINE und EMBASE sowie in den Cochrane-Datenbanken durchgeführt. Es wurden placebokontrollierte RCTs mit dem zu bewerteten oder mit dem zu vergleichenden Arzneimittel gesucht. Die Suchstrategie ist in Anhang 4-B dargestellt.

Von den insgesamt erzielten 965 Treffern wurden infolge des Dubletten-Checks 213 Ergebnisse entfernt (vgl. Abbildung 2).

Nach dem Titel-/Abstractscreening wurden anhand der Ausschlusskriterien (Abschnitt 4.2.2) 701 Treffer ausgeschlossen. Die verbleibenden 51 Treffer wurden im Volltext gelesen und anhand der Ein- und Ausschlusskriterien in relevante und nicht relevante Treffer eingestuft.

Abbildung 5: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach RCTs für indirekte Vergleiche



In der bibliografischen Literaturrecherche nach placebokontrollierten RCTs mit dem zu bewertenden oder mit dem zu vergleichenden Arzneimittel wurden 22 Studien und dazugehörige 28 Publikationen identifiziert.

4.3.2.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4.3-43) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde (sofern eine solche durchgeführt wurde). Fügen Sie für jede Studie eine neue Zeile ein.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4.3-43: Relevante Studien aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel und der Vergleichstherapie für den indirekten Vergleich

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d.b)
<Studie 1>	clinicaltrials.gov [Zitate] clinicalstudyresults.org [Zitate]	ja	nein
placebokontrollierte Studien mit dem zu bewertenden Arzneimittel			
ATTAIN (M/34273/34)	clinicaltrials.gov [NCT01001494],	ja	nein ²²
ACCORD I (LAS-MD-33)	clinicaltrials.gov [NCT00891462] International Clinical Trials Registry Platform [NCT00891462]	ja	ja
ACCORD II (LAS-MD-38)	clinicaltrials.gov [NCT01045161] International Clinical Trials Registry Platform [NCT01045161]	ja	nein
placebokontrollierte Studien mit der Vergleichstherapie			
Spiriva® Assessment of FEV ₁ (SAFE)	clinicaltrials.gov [NCT00277264]	nein	ja
Spiriva® Assessment of FEV ₁ (SAFE Portugal)	clinicaltrials.gov [NCT00239408]	nein	ja
Effects of 9-mth Treatment of SPIRIVA on Health Related Quality of Life in Patients With COPD (TIPHON)	clinicaltrials.gov [NCT00274053]	nein	ja
Evaluation of the Long- Term Effects of Spiriva on Lung Function in COPD Patients (UPLIFT)	clinicaltrials.gov [NCT00144339]	nein	ja
Formoterol Certihaler, Tiotropium HandiHaler and Tiotropium HandiHaler in Combination With Formoterol Certihaler in Patients With Stable COPD	clinicaltrials.gov [NCT00134979]	nein	ja
SPIRIVA in Usual Care	clinicaltrials.gov	nein	ja

²² Bei der bibliografischen Literatursuche wurde lediglich der Abstract gefunden. Durch eine daraufhin stattfindende freie Suche wurde der nur online publizierte Volltext („Epub ahead of print“) gefunden und in den Studienpool aufgenommen.

Studienbezeichnung	Identifikationsorte (Name der Studienregister und Angabe der Zitate)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein / n. d.b)
(SPRUCE)	[NCT00274079]		
Efficacy and Safety (Including 24-hour Holter Monitoring) of Tiotropium Inhalation Capsules in Patients With COPD	clinicaltrials.gov [NCT00239460]	nein	ja
Exacerbation of Chronic Obstructive Pulmonary Disease (COPD) in the Veterans Affairs (VA) Setting	clinicaltrials.gov [NCT00274547]	nein	ja
26 Week Efficacy, Safety and Tolerability Study of Indacaterol in Patients With Chronic Obstructive Pulmonary Disease (COPD)	clinicaltrials.gov [NCT00463567]	nein	ja
Measure of the Long Term Influence of SPIRIVA® in Acute Respiratory Disorders	clinicaltrials.gov [NCT00274014]	nein	ja
Effect of Tiotropium on Inflammation and Exacerbations in COPD	clinicaltrials.gov [NCT00405236]	nein	ja
Tiotropium / Respimat One-Year Study	clinicaltrials.gov [NCT00168844] [NCT00168831]	nein	ja
12 Week Comparison of 5 Mcg and 10 Mcg of Tiotropium / Respimat, Placebo and Ipratropium MDI in COPD	clinicaltrials.gov [NCT00239473] [NCT00240435]	nein	ja
Tiotropium / Respimat One Year Study in COPD	clinicaltrials.gov [NCT00387088]	nein	ja
Effects of Once Daily Tiotropium in Patients With Chronic Obstructive Pulmonary Disease (COPD) of Different Severity	clinicaltrials.gov [NCT00274573]	nein	ja
12 Week Efficacy of Tiotropium Versus Placebo in Patients With Mild COPD According to Swedish Guidelines (SPIRIMILD)	clinicaltrials.gov [NCT00144196]	nein	ja
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. b: n. d. = bibliografische Literaturrecherche nicht durchgeführt.			

4.3.2.1.1.4 Resultierender Studienpool: Studien für indirekte Vergleiche

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.2.1.1.1, 4.3.2.1.1.2, 4.3.2.1.1.3) resultierenden Pool relevanter Studien für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle für die Nutzenbewertung berücksichtigten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Datenquellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4.3-44: Studienpool – RCT mit dem zu bewertenden Arzneimittel und der Vergleichstherapie für den indirekten Vergleich

Studie ^e	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponsernte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrollierte Studien mit dem zu bewertenden Arzneimittel						
ATTAIN	ja	ja	nein ^d	ja [38]	ja [NCT01001494]	ja, [39] ²³
ACCORD I	ja	ja	nein	ja [40]	ja [NCT00891462]	ja, [41]
ACCORD II	ja	ja	nein	ja [19]	ja [NCT01045161]	nein
placebokontrollierte Studien mit der Vergleichstherapie						
SAFE (Chann 2007)	nein	nein	ja	nein	ja [NCT00277264]	ja [42]
SAFE Portugal (Moita 2008)	nein	nein	ja	nein	ja [NCT00239408]	ja [43]
TIPHON (Tonnel 2008)	nein	nein	ja	nein	ja [NCT00274053]	ja [44]
UPLIFT (Tashkin 2008)	nein	nein	ja	nein	ja [NCT00144339]	ja [45] weitere: [46-49]
Vogelmeier 2008	nein	nein	ja	nein	ja [NCT00134979]	ja [50]
SPRUCE (Freeman 2007)	nein	nein	ja	nein	ja [NCT00274079]	ja [21]
Verkindre 2006	nein	nein	ja	nein	nein	ja [51]
Covelli 2005	nein	nein	ja	nein	ja [NCT00239460]	ja [52]
Niewoehner 2005	nein	nein	ja	nein	ja [NCT00274547]	ja [53] weitere: [54]

²³ Bei der bibliografischen Literatursuche wurde lediglich der Abstract gefunden. Durch eine daraufhin stattfindende freie Suche wurde der nur online publizierte Volltext („Epub ahead of print“) gefunden und in den Studienpool aufgenommen.

Studie ^e	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponsernte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
Brusasco 2003	nein	nein	ja	nein	nein	ja [55]
Casaburi 2000	nein	nein	ja	nein	nein	ja [56]
Donohue 2002	nein	nein	ja	nein	nein	ja [57]
Casaburi 2002	nein	nein	ja	nein	nein	ja [58]
INHANCE (Donohue 2010)	nein	nein	ja	nein	ja [NCT00463567]	ja [59]
MISTRAL (Dusser 2006)	nein	nein	ja	nein	ja [NCT00274014]	ja [60]
Powrie 2007	nein	nein	ja	nein	ja [NCT00405236]	ja [61]
Bateman 2010a	nein	nein	ja	nein	ja [NCT00168844, NCT00168831]	ja [62]
Voshaar 2008	nein	nein	ja	nein	ja [NCT00239473, NCT00240435]	ja [63]
Bateman 2010b	nein	nein	ja	nein	ja [NCT00387088]	ja [64]
Beeh 2006	nein	nein	ja	nein	ja [NCT00274573]	ja [65]
Johansson 2008	nein	nein	ja	nein	ja [NCT00144196]	ja [20]

a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.

b: Studie, für die der Unternehmer (Almirall, S.A.) Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war.

c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

d: Studie wurde in Kooperation mit Forrest Laboratories als Lizenzinhaber in den USA durchgeführt

e: Der hier eingeführten Studiennamen werden im Folgenden ohne zusätzliche Bezeichnung des Autors der Hauptpublikation verwendet. Besitzt die Studie kein Akronym wird die Studie mit dem Erstautor und Jahr der Hauptpublikation aufgeführt und die Bezeichnung im Folgenden so verwendet.

Der resultierende Studienpool beinhaltet 21 Studien, die Tiotropiumbromid gegen Placebo vergleichen und 3 Studien, die Acclidiniumbromid gegen Placebo vergleichen. Die Zahl der Tiotropiumbromid-Studien ist nicht identisch mit den in der Nutzenbewertung des IQWiG zu Tiotropiumbromid [22] erfassten 20 Studien (mit den Behandlungsarmen Tiotropiumbromid

und Placebo). Dieser Unterschied ist durch den Einschluss kürzerer Studien (< 6 Monate), der Definition der patientenrelevanten Endpunkte und den Ausschluss von Studien aufgrund der Definition der Intervention oder aufgrund des Studienziels zu begründen. In zwei der vom IQWiG eingeschlossenen Studien (Kesten 2008; Studien-ID: 205.230 und Ambrosino 2008; Studien-ID: 205.247) handelt es sich um die Bewertung einer Kombinationstherapie von Tiotropiumbromid und Rehabilitation. Im Fokus der beiden Studien steht der zusätzlich zur Rehabilitation beobachtete Effekt von Tiotropiumbromid auf den primären Endpunkt der körperlichen Belastbarkeit. In den Einschlusskriterien des vorliegenden Dossiers ist als Intervention „Monotherapien mit Tiotropiumbromid“ definiert. Ebenfalls ausgeschlossen wurde die Publikation von Cooper 2010 (Studien-ID: 205.368). Bei dieser Studie handelt es sich um eine Validierungsstudie des Belastungstests auf dem Laufband zur Bewertung der körperlichen Belastbarkeit. Studien, die ihren Fokus nicht auf den Wirksamkeitsvergleich des Arzneimittels gegen Placebo oder das jeweilige andere Arzneimittel legen, werden nach den in 4.2.2 beschriebenen Ausschlusskriterien nicht in den Studienpool des vorliegenden Dossiers aufgenommen.

4.3.2.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – Studien für indirekte Vergleiche

4.3.2.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-F zu hinterlegen.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4.3-45: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel für den indirekten Vergleich

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen^b (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
<Studie 1>	<i>RCT, doppelblind, parallel</i>	<i>Jugendliche und Erwachsene, leichtes bis mittelschweres Asthma</i>	<Gruppe 1> (n=354) <Gruppe 2> (n=347)	<i>Run-in: 2 Wochen Behandlung: 6 Monate</i>	<i>Europa (Deutschland, Frankreich, Polen) 9/2003 – 12/2004</i>	<i>FEV₁; Asthma-Symptome, Exazerbationen, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse</i>
placebokontrollierte Studien mit dem zu bewertenden Arzneimittel						
ATTAIN	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	Patienten ≥ 40 Jahren, mittelschwere bis schwere COPD, eine Raucherhistorie ≥ 10 Packungsjahren. Nach Gabe eines Bronchodilatators FEV ₁ /FVC <70% FEV ₁ $\geq 30\%$ bis <80% des Sollwerts (Spirometrie)	Acidiniumbromid 400 μg (n=272); davon relevante Patienten ^a : n=269 Placebo (n=276); davon relevante Patienten ^a : n=273	Run-in: 2 Wochen Behandlung: 24 Wochen Follow-up: 2 Wochen	Europa (Deutschland, Spanien, Frankreich, Italien, Polen, Ungarn, Tschechien, Russland, Ukraine, Südafrika, Peru 10/2009 – 11/2010	Primärer Endpunkt: Trough- FEV ₁ Änderung gegenüber dem Baselinewert; Sekundäre Endpunkte: Peak- FEV ₁ Änderung gegenüber dem Baselinewert; TDI-Responders: Nr. Patienten mit ≥ 1 Punkt Verbesserung) SGRQ-Responders: Nr. Patienten mit ≥ 4 Punkte Verbesserung.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

ACCORD I	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	Patienten ≥ 40 Jahren, mittelschwere bis schwere COPD, eine Raucherhistorie ≥ 10 Packungsjahren. Nach Gabe eines Bronchodilatators $FEV_1/FVC < 70\%$ $FEV_1 \geq 30\%$ aber $< 80\%$ des Sollwerts (Spirometrie)	Acclidiniumbromid 400 μg (n=190); davon relevante Patienten ^a : n=190 Placebo (n=186); davon relevante Patienten ^a : n=185	Run-in: 2 Wochen Behandlung: 12 Wochen Follow-up: 2 Wochen	Nordamerika (Vereinigte Staaten, Kanada) 09/2009 – Nov 2009	Primärer Endpunkt: Trough- FEV_1 Änderung gegenüber dem Baselinewert; Sekundärer Endpunkt: Peak- FEV_1 Änderung gegenüber dem Baselinewert;
ACCORD II	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	Patienten ≥ 40 Jahren, mittelschwere bis schwere COPD, eine Raucherhistorie ≥ 10 Packungsjahren. Nach Gabe eines Bronchodilatators $FEV_1/FVC < 70\%$ $FEV_1 \geq 30\%$ aber $< 80\%$ des Sollwerts (Spirometrie)	Acclidiniumbromid 400 μg (n=178); davon relevante Patienten ^a : n=177 Placebo (n=182); davon relevante Patienten ^a : n=182	Run-in: 2 Wochen Behandlung: 12 Wochen	Nordamerika (Vereinigte Staaten, Kanada) Dez 2009 – Sept 2010	Primärer Endpunkt: Trough- FEV_1 Änderung gegenüber dem Baselinewert; Sekundärer Endpunkt: Peak- FEV_1 Änderung gegenüber dem Baselinewert;
placebokontrollierte Studien mit der zweckmäßigen Vergleichstherapie						
<Studie 1>	RCT, doppelblind, parallel	Jugendliche und Erwachsene, leichtes bis mittelschweres Asthma	<Gruppe 1> (n=354) <Gruppe 2> (n=347)	Run-in: 2 Wochen Behandlung: 6 Monate	Europa (Deutschland, Frankreich, Polen) 9/2003 – 12/2004	FEV_1 ; Asthma-Symptome, Exazerbationen, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse
SAFE	RCT, placebokontrolliert, doppelblind, multizentrisch,	$FEV_1 \leq 65\%$; $FEV_1/FVC \leq 70\%$; ≥ 1 Exazerbation im	Tiotropiumbromid 18 μg QD (n=608); davon relevante Patienten ^a : n=587	48 Wochen	Kanada	Primärer Endpunkt: Trough- FEV_1 bei Studienende

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

	parallel	Vorjahr, jedoch nicht innerhalb 6 Wochen vor Randomisierung	Placebo (n=305); davon relevante Patienten ^a : n=287			Sekundärer Endpunkt: FVC, FEV ₆ , SGRQ, Exazerbationen
SAFE Portugal	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ ≤ 70%; FEV ₁ /FVC ≤ 70%; ausgeschlossen, falls ≥ 3 Exazerbationen im Vorjahr oder Exazerbation innerhalb 6 Wochen vorher	Tiotropiumbromid 18µg QD (n=147); davon relevante Patienten ^a : n=144 Placebo (n=164); davon relevante Patienten ^a : n=160	12 Wochen	Portugal	Primärer Endpunkt: Änderung Trough-FEV ₁ -Wert nach 12 Wochen. Sekundäre Endpunkte: Trough-FEV ₁ -Wert nach 6 Wochen, Trough-FVC-Wert nach 6 und 12 Wochen, EQ-5D-Fragebogen und Verwendung von Notfallmedikation zur Tages-/Nachtzeit.
TIPHON	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ 20-70% FEV ₁ /SVC ≤ 70%;	Tiotropiumbromid 18µg QD (n=266); davon relevante Patienten ^a : n=247 Placebo (n=288); davon relevante Patienten ^a : n=245	9 Monate	Frankreich Mai 2002 – Juni 2003 Nachbeobachtung: August 2002 – April 2004	Primärer Endpunkt: Anteil an Patienten, die am Studienende eine Reduktion von mindestens 4 Punkten im SGRQ-Gesamtwert erreichen. Sekundäre Endpunkte: Exazerbationen, AEs
UPLIFT	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ ≤ 70%; FEV ₁ /FVC ≤ 70%; Ausschluss bei Exazerbation 4 Wochen vor	Tiotropiumbromid 18µg QD (n=2987); davon relevante Patienten ^a : n= variiert je nach	4 Jahre	490 Zentren/ 37 Länder Januar 2003 – Februar 2008	Primärer Endpunkt: Zusammengesetzt aus Rückgangsrate des mittleren FEV ₁ -

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

	Studieneignis	Ergebnis				Wertes vor und nach Bronchodilatation
		Placebo (n=3006); davon relevante Patienten ^a : n= variiert je nach Ergebnis				Sekundäre Endpunkte: FVC, SGRQ-Fragebogen, Exazerbationen und Mortalität.
Vogelmeier 2008	RCT, placebokontrolliert, open label, multizentrisch, parallel	Stabile COPD FEV ₁ <70%, FEV ₁ /FVC <70%; ≥40 Jahre ≥10 Packungsjahre	Tiotropiumbromid 18µg QD (n=221); davon relevante Patienten ^a : n=221 Placebo (n=209); davon relevante Patienten ^a : n=209	24 Wochen	Europa (Deutschland, Italien, Niederlande, Russland, Polen, Tschechien, Spanien und Ungarn) Oktober 2004 – November 2005	Primärer Endpunkt: FEV ₁ 2 h nach der Dosis nach 24 Wochen Sekundäre Endpunkte: „Schlechte Tage“ aufgrund von COPD, Symptome, Verwendung von Notfallmedikation und expiratorischer Spitzenfluss sowie Aspekte der gesundheitsbezogenen Lebensqualität.
SPRUCE	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	Stabile COPD im Sinne der britischen Richtlinien; FEV ₁ : 30-65%, FEV ₁ /FVC ≤70%.	Tiotropiumbromid 18µg QD (n=200); davon relevante Patienten ^a : n=191 Placebo (n=195); davon relevante Patienten ^a : n=183	12 Wochen	Großbritannien (England, Schottland, Wales) Oktober 2002 – Oktober 2003	Primärer Endpunkt: Trough-FEV ₁ -Wert und Sicherheit. Sekundäre Endpunkte: Trough-FVC-Wert, Verwendung eines schnell wirkenden β-Agonisten für Notfälle sowie Exazerbationen von >3 Tagen Dauer, die

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

						eine Änderung der Behandlung erfordern.
Verkindre 2006	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ ≤ 50%; FEV ₁ /FVC ≤ 70%; Verbleibendes Volumen ≥ 125% Im Falle von instabilen, 6 Wochen vorher oral verabreichten Dosen Corticosteroid Patienten ≥ 40 Jahren, Raucherhistorie ≥ 10 Packungsjahren.	Tiotropiumbromid 18µg QD (n=46); davon relevante Patienten ^a : n= variiert je nach Ergebnis Placebo (n=54); davon relevante Patienten ^a : n= variiert je nach Ergebnis	12 Wochen	Frankreich (10 Zentren)	Primäre Endpunkte: FVC und Lungenemphysem Sekundäre Endpunkte (u.A.): körperliche Belastbarkeit, Dyspnoe und gesundheitsbezogene Lebensqualität (HRQoL, Health-Related Quality of Life)
Covelli 2005	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ ≤ 60%; FEV ₁ /FVC ≤ 70% Ausschluss falls Exazerbation in letzten 6 Wochen Patienten ≥ 40 Jahren, Raucherhistorie ≥ 10 Packungsjahren.	Tiotropiumbromid 18µg QD (n=100); davon relevante Patienten ^a : n=194 Placebo (n=96); davon relevante Patienten ^a : n=84	12 Wochen	Vereinigte Staaten (12 Zentren)	Primärer Endpunkt: Trough-FEV ₁ -Wert Sekundäre Endpunkte (u. a.): Trough-FVC-Wert, Trough-FEV ₁ und FVC, allgemeine Bewertung der COPD durch Mediziner und Patienten, Ergebnisse des EQ-5D-Fragebogens und Behandlung mit Notfallmedikation.
Niewoehener 2005	RCT, placebokontrolliert, doppelblind, multizentrisch,	FEV ₁ ≤ 60%; FEV ₁ /FVC ≤ 70% Ausschluss, wenn nicht ≥ 30 Tage	Tiotropiumbromid 18µg QD (n=914); davon relevante	6 Monaten	Vereinigte Staaten (26 Zentren)	Primärer Endpunkt: Prozentsatz von Patienten mit einer

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

	parallel	vorher von Exazerbation erholt Patienten ≥ 40 Jahren, Raucherhistorie ≥ 10 Packungsjahren.	Patienten ^a : n=914 Placebo (n=915); davon relevante Patienten ^a : n=915			COPD-Exazerbation und dem Prozentsatz der Patienten mit einer COPD-bedingten Einweisung ins Krankenhaus.
Brusasco 2003	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ $\leq 65\%$; FEV ₁ /FVC $\leq 70\%$; Patienten ≥ 40 Jahren, Raucherhistorie ≥ 10 Packungsjahren.	Tiotropiumbromid 18 μ g QD (n=402); davon relevante Patienten ^a : n=386 Placebo (n=400); davon relevante Patienten ^a : n=362	24 Wochen	18 Länder	Endpunkte (nicht weiter differenziert): Exazerbationen, Nutzung medizinischer Versorgung, Dyspnoe (bewertet mit TDI), gesundheitsbezogene Lebensqualität (bewertet mit SGRQ) und Spirometrie (Trough-FEV ₁ -Wert, FVC).
Casaburi 2000	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ $\leq 65\%$; FEV ₁ /FVC $\leq 70\%$; ≥ 40 Jahre alt; COPD-Diagnose nach ATS-Definition; Tabakkonsum von ≥ 10 Packungsjahren	Tiotropiumbromid 18 μ g QD (n=279); davon relevante Patienten ^a : n=276 Placebo (n=191); davon relevante Patienten ^a : n=188	13 Wochen	Vereinigte Staaten (25 Zentren)	Primärer Endpunkt: Trough-FEV ₁ -Wert bei der letzten Behandlungsvisite.
Donohue 2002	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ $\leq 60\%$; FEV ₁ /FVC $\leq 70\%$; ≥ 40 Jahre > 10 Packungsjahre	Tiotropiumbromid 18 μ g QD (n=209); davon relevante Patienten ^a : n= variiert je nach Ergebnis	24 Wochen	39 Zentren/ 12 Länder	Endpunkte (nicht weiter differenziert): 12-stündiger Beobachtung der Spirometrie (FEV ₁ ,

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

			Placebo (n=201); davon relevante Patienten ^a : n= variiert je nach Ergebnis			FVC),Transition Dyspnoe Index (TDI), SGRQ
Casaburi 2002	Zwei RCTs, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ ≤ 65%; FEV ₁ /FVC ≤ 70%; ≥40 Jahre ≥ 10 Packungsjahre	Tiotropiumbromid 18µg QD (n=550); davon relevante Patienten ^a : n=507 Placebo (n=371); davon relevante Patienten ^a : n=332	1 Jahr	50 Zentren	Primärer Endpunkt: Trough- FEV ₁ -Wert, Veränderungen der Dyspnoe wurden mit dem Transition Dyspnoe Index gemessen, während der Gesundheitszustand mit dem krankheitsspezifische n SGRQ und dem allgemeinen SF-36 gemessen wurde. Medikationseinsatz und unerwünschte Nebenwirkungen wurden aufgezeichnet.
Bateman 2010a	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ ≤ 60%; FEV ₁ /FVC ≤ 70%; ≥40 Jahre COPD-Diagnose ≥ 10 Packungsjahre	Tiotropiumbromid 5µg QD (n=670); davon relevante Patienten ^a : n=variiert je nach Ergebnis Placebo (n=653); davon relevante Patienten ^a : n= variiert je nach Ergebnis	48 Wochen	Süd Afrika, Großbritannien, Deutschland, Kanada	Primärer Endpunkt: CFB Trough-FEV ₁ -Wert Sekundäre Endpunkte: SGRQ-Gesamtwert Zentralwert des TDI Als unerwünschte Nebenwirkungen eingestufte COPD- Exazerbationen pro

						Patientenjahr
INHANCE	RCT, placebokontrolliert, open-label, multizentrisch, parallel	Moderate bis schwere COPD (GOLD) FEV ₁ <80% und ≥30% FEV ₁ /FVC <70%; ≥40 Jahre alt; ≥20 Packungsjahre;	Tiotropiumbromid 18µg QD (n=420); davon relevante Patienten ^a : n=415 Placebo (n=425); davon relevante Patienten ^a : n=418	26 Wochen	Vereinigte Staaten, Europa (Schweden, die Türkei, Deutschland)	Primärer Endpunkt: Trough-FEV ₁ -Wert nach Woche 12. Zusätzliche Analysen (nicht für vielfaches Auftreten angepasst) berücksichtigten u. A. den Transition Dyspnoe Index (TDI), den Gesundheitszustand (St George's Respirationsfragebogen, SGRQ), und Exazerbationen.
Voshaar 2008	Publikation zu zwei identischen RCTs; placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ ≤ 60%; FEV ₁ /FVC ≤ 70%; ≥40 Jahre COPD-Diagnose ≥ 10 Packungsjahre	Tiotropiumbromid 5µg QD (n=180); davon relevante Patienten ^a : n=175 Placebo (n=181); davon relevante Patienten ^a : n=171	12 Wochen	RCT 1: Deutschland, Italien, Südafrika, Schweiz RCT 2: USA, Kanada	Primärer Endpunkt: Trough-FEV ₁ nach 12 Wochen Behandlung. Sekundäre Endpunkte: u. A. weitere Spirometriemessungen und Verwendung von Notfallmedikation.
Bateman 2010b	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ ≤ 60%; FEV ₁ /FVC ≤ 70%; ≥40 Jahre ≥ 10 Packungsjahre	Tiotropiumbromid 5µg QD (n=1989); davon relevante Patienten ^a : n=1952 Placebo (n=2002); davon relevante Patienten ^a : n=1965	48 Wochen	336 Zentren, 31 Länder	Primärer Endpunkt: Kombination aus Trough-FEV ₁ -Wert und die Zeit bis zur ersten Exazerbation. Sekundäre

						Endpunkte: weitere Spirometriemessungen, Exazerbationen, SGRQ
Beeh 2006	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	COPD-Patienten; FEV ₁ /FVC < 70%; FEV ₁ ≤ 70% vom Sollwert; Alter ≥ 40 Jahre; ≥ 10 Packungsjahre	Tiotropiumbromid 18µg QD (n=1236); davon relevante Patienten ^a : n=1236 Placebo (n=403); davon relevante Patienten ^a : n=403	12 Wochen + 1 Woche „run-in“	Deutschland	Primäre Endpunkte: Verbesserung der Lungenfunktion 23-24 Stunden nach der vorhergehenden Inhalation (Trough-Wert) und 2 Stunden nach Inhalation der Studienmedikation. Sekundäre Endpunkte: FVC, IVC, Exazerbationen, unerwünschte Ereignisse
MISTRAL	RCT, placebokontrolliert, doppelblind, multizentrisch, parallel	FEV ₁ 30-65% vor Bronchodilatator; FEV ₁ /SVC ≤ 70% vor Bronchodilatator; ≥ 40 Jahre alt; ≥ 10 Packungsjahre; ≥ 1 Exazerbation im letzten Jahr; Ausschluss falls Exazerbation in letzten 6 Wochen	Tiotropiumbromid 18µg QD (n=500); davon relevante Patienten ^a : n=500 Placebo (n=510); davon relevante Patienten ^a : n=510	Run-in: 2 Wochen Behandlung: 48 Wochen + 3 Wochen Untersuchung + 2 Wochen	Frankreich (177 Zentren)	Endpunkte (nicht weiter differenziert): Exazerbationen (Auftreten, Krankenhausaufenthalt); Nutzung der medizinischen Ressourcen; Begrenzung des Luftstroms (FEV ₁ , FVC, SVC und Einatemkapazität)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

<p>Johansson 2008</p>	<p>RCT, placebokontrolliert, doppelblind, multizentrisch, parallel</p>	<p>Ambulante Patienten im Alter von >40 Jahren mit diagnostizierter, milder COPD im Sinne der Schwedischen Richtlinien von 2003 (FEV₁/FVC <70% und FEV₁ >60% nach Bronchodilatator vorhergesagt; vorhergegangener Tabakkonsum von >10 Packungsjahren; Dyspnoewert >2 (Messung mit Medical Research Council (MRC))</p>	<p>Tiotropiumbromid 18µg QD (n=107); davon relevante Patienten^a: n=107 Placebo (n=117); davon relevante Patienten^a: n=117</p>	<p>Run-in: 2 Wochen Behandlung: 1 Jahr</p>	<p>Schweden (27 Zentren) Schweden (27 Zentren) März 2004 – Juli 2005</p>	<p>Primäre Endpunkte: Änderung der FEV₁-Fläche unter der Kurve vom Vordosiswert (zeitlicher Nullpunkt) zum Wert 2 Stunden nach der Dosis (AUC₀₋₂); vom Ausgangswert über 12 Wochen. Weitere Endpunkte: Trough-FEV₁- und -FVC-Werte, Verwendung von Notfallmedikation und unerwünschte Nebenwirkungen.</p>
<p>Powrie 2007</p>	<p>RCT, placebokontrolliert, doppelblind, monozentrisch, parallel</p>	<p>FEV₁ ≤80% vor Bronchodilatator; FEV₁/FVC ≤70% vor Bronchodilatator; ≥40 Jahre alt; ≥10 Packungsjahre</p>	<p>Tiotropiumbromid 18µg QD (n=69); davon relevante Patienten^a: n=69 Placebo (n=73); davon relevante Patienten^a: n=73</p>	<p>Run-in: 2 Wochen Behandlung: 1 Jahr</p>	<p>London, Großbritannien (44 Zentren)</p>	<p>Primärer Endpunkt: Konzentration von Interleukin (IL)-6 im Sputum. Sekundäre Endpunkte (u.A.): Konzentrationen von IL-8 und Myeloperoxidase (MPO) im Sputum, Konzentrationen von Serum IL-6 und C-</p>

	reaktiven Proteinen (CRP), bakterielle Besiedlung des Sputums, FEV ₁ und Exazerbationsfrequenz.
--	--

a: Intention to Treat (ITT) population

b: Es werden nur die für den indirekten Vergleich relevanten Therapiearme dargestellt (Tiotropiumbromid 5 oder 18µg, Placebo)

Tabelle 4.3-46: Charakterisierung der Interventionen – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel für den indirekten Vergleich

Studie	<Gruppe 1>	<Gruppe 2>	<Gruppe 3>	Begleitende Medikation
<Studie 1>	xxx 250 µg, 1 Inhalation bid + Placebo 2 Inhalationen bid	yyy 200 µg, 2 Inhalationen bid + Placebo 1 Inhalation bid		Vorbehandlung: zzz 1000 µg pro Tag, 4 Wochen vor Studienbeginn Bedarfsmedikation: aaa
placebokontrollierte Studien mit dem zu bewertenden Arzneimittel				
ATTAIN	Acclidiniumbromid 200 µg, 2 Inhalationen bid (morgens/abends, via Genuair®)	Acclidiniumbromid 400 µg, 2 Inhalationen bid (morgens/abends, via Genuair®)	Placebo, 2 Inhalationen bid (morgens/abends, via Genuair®)	<u>Bedarfsmedikation:</u> Salbutamol als <u>zugelassene</u> atemwegserweiternde Medikamente, unterbrochen ≥6 h vor dem Studienbesuch <u>Begleitende Medikation:</u> inhalative Corticosteroide (ICS) oder orale Theophylline oder systemische Corticosteroide, dauerhafte Nutzung ≥ 4 Wochen vor Run-In Sauerstofftherapie, dauerhafte Nutzung ≥ 4 Wochen vor Run-In (≤15 h/ täglich)
ACCORD I	Acclidiniumbromid 200 µg, 2 Inhalationen bid (morgens/abends, via Genuair®)	Acclidiniumbromid 400 µg, 2 Inhalationen bid (morgens/abends, via Genuair®)	Placebo, 2 Inhalationen bid (morgens/abends, via Genuair®)	<u>Bedarfsmedikation:</u> Salbutamol als <u>zugelassenes</u> atemwegserweiterndes Medikament, unterbrochen ≥6 h vor dem Studienbesuch; inhalative Corticosteroide (ICS), systemische Corticosteroide entspricht Prednison ≤10 mg/täglich oder 20 mg alle zwei Tage, und oralem Theophyllin. ICS und Theophyllin unterbrochen vor dem Studienbesuch.
ACCORD II	Acclidiniumbromid 200 µg, 2 Inhalationen bid (morgens/abends, via Genuair®)	Acclidiniumbromid 400 µg, 2 Inhalationen bid (morgens/abends, via Genuair®)	Placebo, 2 Inhalationen bid (morgens/abends, via Genuair®)	<u>Bedarfsmedikation:</u> Die zur Behandlung von episodischen oder chronischen Erkrankungen <u>zugelassene</u> Medikamente beinhalteten zusätzliche Gabe von Sauerstoff; nicht zur Behandlung von episodischen oder chronischen Erkrankungen <u>zugelassene</u> Medikamente beinhalteten kurz- und langwirksame Anticholinergika, langwirksame β ₂ -Agonisten (LABAs), Kombinationstherapien, Methylxanthine, Cromone, Leukotrien-Modifikatoren und

			nichtselektive-blockierende Wirkstoffe. Die zur Behandlung von episodischen, nicht aber chronischen Erkrankungen <u>zugelassenen</u> Medikamente beinhalteten SABAs; die zur Behandlung von chronischen, nicht aber episodischen Erkrankungen <u>zugelassenen</u> Medikamente beinhalteten ICSs, orale oder parentale Corticosteroide, H ₁ -Antihistamine und selektive, α_1 -blockierende Wirkstoffe.
placebokontrollierte Studien mit der zweckmäßigen Vergleichstherapie			
SAFE	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen</u> : Feste Dosis oraler Corticosteroide, ICS, Theophyllinpräparate, schleimlösende Präparate (die keine Bronchodilatoren enthalten), LABAs
SAFE Portugal	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen</u> : LABAs, Theophyllin, Schleimlöser, ICS, feste Dosis oraler Corticosteroide. Vorübergehend erhöhte Dosis von Theophyllin oder oraler Corticosteroide bei Exazerbationen; <u>Nicht zugelassen</u> : Theophyllin 24-Stunden-Präparate
TIPHON	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen</u> : Feste Dosen von Theophyllinpräparaten (außer 24-Stunden-Präparate), Schleimlöser, ICS und orale Steroide
UPLIFT	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen</u> : Alle Atemwegstherapeutika außer zu inhalierende, anticholinerge Medikamente
Vogelmeier 2008	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen</u> : Salbutamol, ICS-Monotherapie
SPRUCE	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen</u> : LABAs, ICS, LABA/ICS
Verkindre	Tiotropiumbromid	Placebo,	<u>Zugelassen</u> : Feste Dosen oraler

2006	18µg, 1 Inhalation od (morgens via HandiHaler®)	1 Inhalation od (morgens via HandiHaler®)	Corticosteroiden, ICS, Theophyllinpräparate, schleimlösende Wirkstoffe; <u>Nicht zugelassen:</u> Verwendung von SABAs, oraler β2-Agonisten oder LABAs
Covelli 2005	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen:</u> ICS, LABAs und Theophylline <u>Nicht zugelassen:</u> Cromone, Leukotrienantagonisten und zu inhalierende Anticholinergika
Niewoehener 2005	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen:</u> Alle anderen respiratorischen Medikationen (inklusive ICS und LABAs) <u>Nicht zugelassen:</u> Unverblindeter, anticholinerg Bronchodilatator
Brusasco 2003	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	Nicht berichtet
Casaburi 2000	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen:</u> feste Dosis Theophyllin, ICS, orales Prednison <u>Nicht zugelassen:</u> Andere zu inhalierende oder orale Bronchodilatoren
Donohue 2002	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen:</u> Gebräuchliche ICS und orale Steroide; <u>Nicht zugelassen:</u> Zu inhalierende, anticholinergische LABAs
Casaburi 2002	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Zugelassen:</u> feste Dosis Theophyllin, ICS, orales Prednison
Bateman 2010a	Tiotropiumbromid 5µg, 2 Inhalationen [2.5µg] od (morgens via Respimat®)	Placebo, 2 Inhalationen od (morgens via Respimat®)	Orale (bis zu 10 mg Prednison pro Tag) und zu inhalierende Corticosteroide (ICS), Theophyllinpräparate, schleimlösende Wirkstoffe und Antileukotriene waren <u>zugelassen</u> , wenn sie vor und während der Studie mindestens für 6 Wochen stabilisiert waren. <u>Nicht zugelassen:</u> Patienten, die LABAs und ICS einnahmen, wechselten vor der Einleitungsphase zu einem einzelnen Produkt (zu inhalierendes Corticosteroid). Salbutamol MDI wurde als Notfallmedikation verwendet.

INHACE	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	Behandlungen mit ICS- Monotherapien, Salbutamol, anticholinergen Bronchodilatoren oder β2- Agonisten wurden unter Beachtung der angezeigten Auswaschung abgebrochen; Patienten, die eine feste Kombination von β2-Agonist/ICS einnahmen, wechselten zur ICS- Monotherapie mit der äquivalenten Dosis.
Voshaar 2008	Tiotropiumbromid 5µg, 2 Inhalationen [2,5µg] od (morgens via Respimat®)	Placebo, 2 Inhalationen od (morgens via Respimat®)	Notfallmedikation (Salbutamol pMDI) war während der Studie je nach Erforderlichkeit <u>zugelassen</u> . Orale Corticosteroide (äquivalent zu <10 mg Prednison pro Tag), oral inhalierte Corticosteroide, Theophylline und Schleimlöser waren <u>zugelassen</u> , wenn sie vor und während der Studie mindestens für 6 Wochen stabilisiert waren. Orale b- Adrenergika und andere experimentelle Medikamente waren im Zeitraum von mindestens einem Monat vor der Einleitungsphase <u>nicht</u> <u>zugelassen</u> . Cromolynnatrium und Nedocromilnatrium waren im Zeitraum von mindestens drei Monaten vor der Einleitungsphase <u>nicht</u> <u>zugelassen</u> . Außerdem waren Anticholinergika, zu inhalierende b-Adrenergika außer Salbutamol und Kombinationsinhalatoren mit festen Dosen während des Behandlungszeitraumes <u>nicht</u> <u>zugelassen</u> .
Bateman 2010b	Tiotropiumbromid 5µg, 2 Inhalationen [2,5µg] od (morgens via Respimat®)	Placebo, 2 Inhalationen od (morgens via Respimat®)	<u>Zugelassen</u> : Salbutamol pMDI wurde allen Patienten als Notfallmedikation zur Verwendung zu jedem Zeitpunkt der Studie zur Verfügung gestellt. Mit <u>Ausnahme</u> von zu inhalierenden Anticholinergika, waren während der Studie alle respiratorischen Medikationen zugelassen.
Beeh 2006	Tiotropiumbromid 18µg, 1 Inhalation od (morgens via HandiHaler®)	Placebo, 1 Inhalation od (morgens via HandiHaler®)	<u>Nicht zugelassen</u> : Langwirksame β-Mimetika und andere (kurzwirksame) Anticholinergika als das Prüfpräparat wurden zu Beginn der Studie als nicht erlaubte

			Begleitmedikation abgesetzt. Kurzwirksame β -Mimetika zur Inhalation wurden einheitlich durch Fenoterol Dosieraerosol zur bedarfsweisen Anwendung ersetzt.
MISTRAL	Tiotropiumbromid 18 μ g, 1 Inhalation od (morgens via HandiHaler [®])	Placebo, 1 Inhalation od (morgens via HandiHaler [®])	<u>Zugelassen:</u> kurzwirksame β 2-Agonisten, ICS und orale Steroide (bei fester Dosis oder Äquivalent von <10mg Prednison pro Tag) <u>Nicht zugelassen:</u> Anticholinergika, langwirksame β 2-Agonisten, Theophyllin
Johansson 2008	Tiotropiumbromid 18 μ g, 1 Inhalation od (morgens via HandiHaler [®])	Placebo, 1 Inhalation od (morgens via HandiHaler [®])	<u>Zugelassen:</u> Die Verwendung eines Salbutamol-Dosierinhalators war den Patienten zur Linderung akuter Symptome je nach Bedarf als Notfallmedikation mit einer Auswaschperiode von 8 Stunden vor der Spirometrie gestattet. Zur Behandlung von COPD-Exazerbationen konnten die Forscher jedoch Antibiotika und orale Corticosteroide (für <2 Wochen) oder Theophyllin (für \leq 7 Tage) verschreiben. <u>Nicht zugelassen:</u> Die Verwendung von kurzwirksamen Anticholinergika, β 2-Agonisten (außer Notfallmedikation), oralen Corticosteroiden und Theophyllin war nicht zugelassen.
Powrie 2007	Tiotropiumbromid 18 μ g, 1 Inhalation od (morgens via HandiHaler [®])	Placebo, 1 Inhalation od (morgens via HandiHaler [®])	<u>Nicht zugelassen:</u> Anticholinergika

Tabelle 4.3-47: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel für den indirekten Vergleich

Studie Gruppe	N ^a	Alter (Jahre) ^b	Geschlecht w / m (%)	Schweregrad ^c (mittel / schwer)%	Rauchen [Packungs- jahre] ^b	Zeit seit COPD- Diagnose [Jahre] ^b	FEV ₁ ^d zu Studienbeginn [%Sollwert] ^b	FEV ₁ ^d zu Studienbeginn [% FVC] ^b	Studienabbruch n (%)
<Studie 1>									
<Gruppe 1>									
<Gruppe 2>									
placebokontrollierte Studien mit dem zu bewertenden Arzneimittel									
ATTAIN	819	62,4 ± 8,0	32,6 / 67,4	68,1 / 31,9	40,2 (19,8)	6,8 (6,2)	56,8 (12,8)	49,6 (10,3)	82 (9,9)
AB 200 µg	277	62,3 ± 7,8	34,7 / 65,3	69,6 / 30,4	40,0 (19,8)	7,0 (6,5)	57,6 (13,2)	49,7 (9,9)	24 (8,6)
AB 400 µg	269	62,9 ± 8,4	32,3 / 67,7	68,7 / 31,3	41,7 (21,1)	7,2 (6,7)	56,2 (12,2)	49,7 (10,2)	17 (6,3)
Placebo	273	62,0 ± 8,0	30,8 / 69,2	65,9 / 34,1	38,9 (18,3)	6,4 (5,4)	56,6 (12,8)	49,3 (10,7)	41 (14,9)
ACCORD I	560	64,3 ± 9,4	47,0 / 53,0	58,4 / 39,5	54,3 (26,8)	8,6 (6,3)	53,8 (13,4)	51,7 (10,5)	94 (16,8)
AB 200 µg	184	63,1 ± 9,5	45,1 / 54,9	53,3 / 43,5	53,0 (23,3)	8,3 (6,1)	52,8 (13,7)	50,9 (10,6)	33 (17,8)
AB 400 µg	190	64,9 ± 9,5	47,4 / 52,6	62,1 / 35,8	57,2 (28,5)	8,9 (6,4)	54,1 (12,9)	51,5 (10,2)	24 (12,6)
Placebo	186	65,1 ± 9,2	48,4 / 51,6	59,7 / 39,2	52,7 (28,1)	8,5 (6,5)	54,6 (13,5)	52,7 (10,5)	37 (19,9)
ACCORD II	542	62,8 ± 8,9	46,9 / 53,1	53,0 / 45,8	53,8 (29,1)	8,0 (6,1)	52,5 (13,4)	51,2 (10,9)	90 (16,5)
AB 200 µg	183	63,4 ± 8,5	45,9 / 54,1	51,9 / 46,4	54,7 (31,2)	7,9 (6,2)	52,0 (13,7)	51,0 (10,7)	29 (15,8)
AB 400 µg	177	63,2 ± 9,0	49,7 / 50,3	44,6 / 54,2	54,2 (27,7)	8,5 (6,2)	50,2 (13,1)	49,9 (10,3)	30 (16,9)
Placebo	182	61,7 ± 9,3	45,1 / 54,9	62,1 / 36,8	52,6 (28,4)	7,7 (6,0)	55,2 (13,2)	53,3 (11,2)	31 (17,0)
placebokontrollierte Studien mit der zweckmäßigen Vergleichstherapie									
SAFE									
TIO 18 µg	608	67,0 ± 8,7	41 / 59	NR / NR	50,2 (22,6)	9,9 (8,1)	39,4 (13,0)	46,4 (12,0)	135 (22,2%)
Placebo	305	67,0 ± 9,1	39 / 61	NR / NR	51,0 (26,3)	9,9 (7,9)	39,3 (14,0)	46,3 (12,0)	84 (27,5%)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

SAFE Portugal ^e (Nichtraucher)									
TIO 18 µg	104	65,7 ± 8,6	3 / 97	NR / NR	54,3 (27,1)	14,0 (10,9)	38,4 (12,8)	44,4 (11,0)	11 (7,5%)
Placebo	120	65,7 ± 9,0	5 / 95	NR / NR	54,3 (26,7)	13,7 (10,5)	42,3 (15,3)	45,2 (11,5)	11 (6,6%)
SAFE Portugal ^e (Raucher)									
TIO 18 µg	40	61,6 ± 9,8	7 / 93	NR / NR	59,9 (23,4)	9,4 (6,8)	44,4 (13,9)	47,9 (13,7)	11 (7,5%)
Placebo	40	64,0 ± 7,2	7 / 93	NR / NR	55,3 (17,3)	12,3 (8,5)	40,4 (14,5)	46,7 (13,4)	11 (6,6%)
TIPHON									
TIO 18 µg	266	65 ± 9,7	13 / 87	42,9 / 57,1	44,4 (21,3)	7,9 (7,6)	47,5 (13,3)	55,3 (11,3)	39 (14,7%)
Placebo	288	64 ± 10,1	15 / 85	38,3 / 61,7	43,0 (22,5)	8,0 (7,9)	46,2 (12,4)	54,6 (11,3)	74 (25,7%)
UPLIFT									
TIO 18 µg	2987	65 ± 8,4	25 / 75	46,0 / 44,0	49 (28)	9,9 (7,6)	39,5 (12,0)	42,4 (11,0)	1.099 (36,8%)
Placebo	3006	65 ± 8,5	26 / 74	45,0 / 44,0	48,4 (27,9)	9,7 (7,4)	39,3 (12,0)	42,1 (11,0)	1.358 (45,2%)
Vogelmeier 2008									
TIO 18 µg	221	63 ± 9,5	21 / 79	NR / 44	38,6 (19,3)	6,9 (6,3)	51,6 (11,2)	54,4 (9,6)	29 (13,1%)
Placebo	209	62 ± 8,6	22 / 78	NR / 46	40,1 (22,8)	6,7 (6,1)	51,1 (11,0)	53,5 (10,0)	30 (14,4%)
SPRUCE									
TIO 18 µg	200	65 ± 9,0	50 / 50	49,5 / 2,9	50,8 / 4,2	36,9 (16,9)	NR	47,0 (10,4)	55,0 (9,7)
Placebo	195	65 ± 9,3	41 / 59	48,1 / 1,6	37,9 (17,7)	NR	49,0 (10,7)	55,0 (10,0)	33 (16,9%)
Verkindre 2006									
TIO 18 µg	46	61 ± 9,5	6 / 94	NR / NR	45,6 (23,1)	9,7 (6,9)	34,7 (9,4)	NR	1 (2,2%)
Placebo	54	60 ± 10,2	6 / 94	NR / NR	41,8 (18)	8,8 (6,6)	35,8 (9,0)	NR	8 (14,8%)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Covelli 2005									
TIO 18 µg	100	66 ± 8,9	44 / 66	NR / NR	66 (35,6)	10,1 (8,1)	40,2 (13,0)	NR	10 (10,0%)
Placebo	96	63 ± 9,2	51 / 49	NR / NR	65 (31,2)	10,4 (7,7)	38,6 (13,8)	NR	17 (17,7%)
Niewoehener 2005									
TIO 18 µg	914	68 ± 8,7	2 / 98	NR / NR	67,4 (35,4)	12,2 (10,4)	35,6 (12,6)	47,9 (11,5)	78 (8,5%)
Placebo	915	68 ± 8,5	1 / 99	NR / NR	69,4 (36,6)	11,9 (10,5)	35,6 (12,6)	47,7 (11,1)	111 (12,1%)
Brusasco 2003									
TIO 18 µg	402	64 ± 8,0	33 / 77	NR / 82	44,1 (22,9)	9,0 (7,3)	39,2 (11,6)	43,7 (9,7)	62 (15,4%)
Placebo	400	65 ± 8,6	34 / 76	NR / 82	42,4 (22,7)	9,8 (7,4)	38,7 (12,1)	42,3 (9,2)	103 (25,7%)
Casaburi 2000									
TIO 18 µg	279	65 ± 8,6	33 / 67	NR / NR	64,5 (33,1)	9,3 (8,0)	39,0 (13,8)	46,0 (11,8)	17 (6,1%)
Placebo	191	65 ± 9,0	37 / 63	NR / NR	60,5 (30,2)	8,6 (6,9)	38,0 (14,1)	46,0 (11,5)	21 (11,0%)
Donohue 2002									
TIO 18 µg	209	65 ± 7,9	26 / 74	NR / 59	47 (25)	9,2 (7,8)	41,0 (39,0)	43,6 (9,8)	121 (57,9%)
Placebo	201	66 ± 7,8	25 / 75	NR / 60	46 (24)	9,7 (7,9)	41,0 (36,0)	41,3 (8,7)	129 (64,2%)
Casaburi 2002									
TIO 18 µg	550	65 ± 9,0	33 / 67	NR / NR	63 (31)	8,6 (7,4)	39,1 (13,7)	45,8 (11,6)	NR (18,7%)
Placebo	371	65 ± 9,0	37 / 63	NR / NR	59 (30)	8,1 (6,8)	38,1 (14,1)	45,5 (11,6)	NR (27,8%)
Bateman 2010a									
TIO 5 µg	670	65 ± 8,6	27 / 73	NR / NR	NR	8,3 (6,4)	38,0 (11,7)	42,4 (11,5)	NR (17,2%)
Placebo	653	65 ± 8,7	25 / 75	NR / NR	NR	9,5 (7,5)	37,5 (11,6)	42,1 (11,0)	NR (31,4%)
INHANCE									
TIO 18 µg	420	64 ± 8,7	35 / 65	NR / NR	50,0 (25,1)	NR	53,9 (15,6)	52,7 (10,1)	89 (21,2%)
Placebo	425	63 ± 8,9	39 / 61	NR / NR	49,7 (23,9)	NR	56,1 (14,3)	53,4 (10,1)	131 (30,8%)
Voshaar 2008									
TIO 5 µg	180	64 ± 9,0	31 / 69	41 / 47	52 (30)	10 (8)	40,0 (12,0)	47,0 (11,0)	16 (8,9%)
Placebo	181	63 ± 9,0	31 / 69	52 / 37	51 (30)	9 (7)	42,0 (12,0)	48,0 (11,0)	22 (12,2%)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Bateman 2010b									
TIO 5 µg	1989	64,8 ± 9,1	22 / 78	NR / NR	46 (26,1)	8,3 (7,0)	47,2 (10,8) ^f	48,5 (11,3)	318 (16,0%)
Placebo	2002	64,8 ± 9,0	23 / 77	NR / NR	45 (26,5)	8,1 (6,5)	45,4 (13,6) ^f	47,9 (11,3)	373 (18,6%)
Beeh 2006				35,8 / 26,8					
TIO 18 µg	1236	62 ± 8,6	25 / 75	NR	36,1 (19,5)	9,7 (7,9)	45,3 (14,9)	57,2 (12,2)	NR (11,4%)
Placebo	403	62 ± 8,7	23 / 77	NR	35,0 (19,4)	9,9 (7,6)	45,7 (15,0)	56,8 (12,0)	NR (16,4%)
MISTRAL									
TIO 18 µg	500	64 ± 9,1	11 / 89	NR / NR	NR	8,2 (7,8)	48,2 (12,8)	55,0 (12,0)	NR (23,4%)
Placebo	510	65 ± 9,5	13 / 87	NR / NR	NR	8,5 (7,7)	47,6 (12,5)	54,0 (12,0)	NR (28,8%)
Johansson 2008				68,6 / 3,2					
TIO 18 µg	107	61 ± 8,3	47 / 53	NR	31,4 (11,9)	4,6 (3,6)	73,6 (12,2)	62,1 (6,4)	NR
Placebo	117	62 ± 8,3	57 / 43	NR	31,6 (12,2)	4,9 (4,8)	73,2 (12,9)	61,3 (6,8)	NR
Powrie 2007									
TIO 18 µg	69	66 ± 8,1	30 / 70	NR / NR	54,6 (25,5)	NR	50,9 (14,8)	NR	21 (30,4%)
Placebo	73	66 ± 9,8	44 / 56	NR / NR	55,7 (28,0)	NR	49,2 (15,6)	NR	21 (28,8%)

FEV₁: Einsekundenkapazität; FVC: forcierte Vitalkapazität; NR: nicht berichtet, µg: Mikrogramm

a: Zahl der Patienten, für die in den Studien Charakteristika zu Studienbeginn dargestellt wurden (Sicherheits-Population; alle randomisierten und behandelten Patienten)

b: Mittelwert und Standardabweichung

c: Nach GOLD Leitlinie 2010 [34] oder Leirlinien die vergleichbare Grenzwerte verwenden; die Anteile der Patienten mit leichter oder sehr schwerer COPD sind nicht aufgelistet, deshalb ergeben die Anteile nicht 100%

d: Postbronchodilatatorisch

e: Patienteneigenschaften wurden pro Raucher-Untergruppe (Raucher und Nichtraucher) für jeden Behandlungsarm angegeben.

f: Die Werte wurden bei der Auswahl der Patienten und nicht bei der Erstuntersuchung, die sieben Tage später stattfand, erfasst.

Beschreiben Sie die Studien zusammenfassend. Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

Studiencharakteristika

Die Tabelle 4.3-45 gibt einen Überblick über die Studiencharakteristika der eingeschlossenen Studien. Die Charakteristika der durchgeführten Interventionen werden in Tabelle 4.3-46 beschrieben.

Bei den eingeschlossenen Studien handelt es sich um randomisierte kontrollierte Studien (RCTs), die alle - mit Ausnahme von zwei Tiotropiumbromid-Studien (INHANCE; Vogelmeier, 2008) - doppelblind durchgeführt wurden. Die Studien berichteten Daten nach 12 und/oder 24 Wochen, jedoch lieferten zu diesen Zeitpunkten nicht alle Studien Daten zu allen Endpunkten. Das Studiendesign unterscheidet sich in Bezug auf den Studienort (geographischer Standort) und die Anzahl der beteiligten Studienzentren (multizentrisch, monozentrisch).

In den meisten Studien wurden Patienten mit einem FEV₁/FVC-Wert von $\leq 70\%$ gegenüber dem Ausgangswert nach Anwendung eines Bronchodilatators und einem maximalen FEV₁-Wert zwischen $< 80\%$ und $< 50\%$ vom Soll eingeschlossen. In einigen Studien wurden Patienten ausgeschlossen, die mit einem Mindestwert von FEV₁ $< 30\%$ unter einer sehr schweren COPD (Stufe IV). In einer der Studien wurden fast nur Patienten mit einer leichten COPD mit einem FEV₁-Wert von $> 60\%$ vom Soll berücksichtigt (Stufe II: 68,6; Stufe III: 3,2%). In den Fällen, in denen die Exazerbationsraten als Einschluss- oder Ausschlusskriterium galten, schlossen alle identifizierten Studien Patienten aus, die erst kürzlich eine Exazerbation hatten oder medizinische Leistungen in Anspruch genommen hatten, welche auf eine Exazerbation hindeuteten, wie z. B. stationäre Behandlung oder Einnahme von oralen Corticosteroiden. In den Studien ATTAIN, ACCORD I und ACCORD II wurden Patienten ausgeschlossen, die eine Exazerbation in den vergangenen 6 Wochen hatten oder bis zu 3 Monate vor der Randomisierung stationär behandelt wurden. Entsprechend wurden in weiteren 9 Studien Patienten ausgeschlossen, die in den vergangenen 4 oder 6 Wochen Exazerbationen hatten (SAFE, SAFE-Portugal, MISTRAL, UPLIFT, Covelli 2005, Niewoehner 2005, Verkindre 2006, Johansson 2008, Beeh 2006). Zwei Tiotropiumbromid-Studien schlossen nur Patienten ein, die im vergangenen Jahr mindestens eine Exazerbation hatten, die behandelt werden musste (SAFE, MISTRAL).

Die zugelassenen Begleitmedikamente unterschieden sich in den identifizierten Studien: In den Aclidiniumbromid-Studien wurden LABAs explizit ausgeschlossen, waren aber in acht Tiotropiumbromid-Studien zugelassen (SAFE, SAFE-Portugal, UPLIFT, SPRUCE, Niewoehner 2005, Covelli 2005, Bateman 2010b, Johansson 2008). Der Anteil der Patienten, die gleichzeitig zur Studienmedikation ICS einnahmen, war unterschiedlich. In den Studien, in denen Daten zum ICS-Hintergrund erhoben wurden, war die Einnahme eines ICS zugelassen (20 Studien). In zwei Studien fehlten Daten zum ICS-Hintergrund (Brusasco 2003, Beeh 2006) und in zwei Studien wurde nicht berichtet, dass die Einnahme eines ICS nicht zugelassen war (Bateman 2010b, Powrie 2007). Die Einnahme von Theophyllin war in den

meisten Studien zugelassen, ebenso wie die Einnahme eines SABA als Bedarfsmedikation (Salbutamol).

Charakteristika der Studienpopulation

Tabelle 4.3-47 zeigt die Charakteristika der in den placebokontrollierten Acclidiniumbromid- und Tiotropiumbromid-Studien untersuchten Patienten. Die Acclidiniumbromid-Studien geben sowohl Charakteristika der Studienarme, als auch der gesamten Studienpopulation an. Die Tiotropium-Studien machen lediglich Angaben zu den einzelnen Studienarmen. Dies kann die Vergleichbarkeit der Studien einschränken.

Der Altersdurchschnitt der Patienten lag in allen Studien bei ca. 65 Jahren. Männer waren zum Teil leicht (Verhältnis von 55-65% zu 45-35%) und zum anderen Teil stärker (Verhältnis von 70-99% zu 30-1%) überrepräsentiert. Dies dürfte aufgrund der nicht geschlechtsspezifischen Unterschiede im Krankheitsverlauf (vgl. Modul 3) keinen Einfluss auf die Vergleichbarkeit der Studien haben. Alle Studien bewerteten eine überwiegend kaukasische Studienpopulation.

Der mittlere FEV₁-Wert vom Soll gegenüber dem Ausgangswert betrug zwischen 50 % und 57 % in den Acclidiniumbromid-Studien und zwischen 35 % und 74 % in den Tiotropiumbromid-Studien. Dies stimmt mit den Unterschieden bei den Einschlusskriterien für den FEV₁-Wert vom Soll zwischen den einzelnen Studien überein (Acclidiniumbromid-Studien FEV₁-Wert von < 80 % vom Soll, Tiotropiumbromid-Studien normalerweise FEV₁-Wert von < 65-70 % vom Soll). Der Anteil der Patienten mit einer mittelschweren und schweren COPD wurde nicht von allen Tiotropiumbromid-Studien berichtet. Die Studien, die den Anteil berichteten schienen bezüglich des Schweregrades der Krankheit vergleichbar. Einzige Ausnahme bilden die Studien von Freeman (SPRUCE) und Johansson 2008. Beide Studien schlossen sehr wenige Patienten (2-4 %) mit schwerer COPD (Stufe III) ein. In den übrigen Studien konnten durchschnittlich 2/3 der Patienten der COPD-Stufe II und 1/3 der Patienten der COPD-Stufe III zugeordnet werden.

Alle Studien schlossen Patienten ein, die zu dem Zeitpunkt Raucher oder ehemalige Raucher waren, wobei die meisten Studien einen Mindestzeitraum von 10 Jahren festlegten. Die Anzahl der Packungsjahre, die in den Acclidiniumbromid-Studien bewertet wurden, lag zwischen 40 und 54. In den Tiotropiumbromid-Studien waren es 31 bis 96 Packungsjahre, wobei 4 Studien einen Mittelwert von mehr als 60 Packungsjahren angaben (Casaburi 2000 & 2002, Covelli 2005 und Niewoehner 2005). Der mittlere Zeitraum seit der COPD-Diagnose betrug bei den Patienten in den Acclidiniumbromid-Studien zwischen 6,4 und 8,9 Jahren und in den Tiotropiumbromid-Studien zwischen 4,6 und 12,2 Jahren. Auch der prozentuale Anteil der Studienabbrecher variiert aufgrund der unterschiedlichen Studiendauern zwischen den Studien.

Trotz einiger der genannten Unterschiede wurden die 24 Studien als größtenteils vergleichbar eingestuft und gelten als geeignet für einen indirekten Vergleich.

4.3.2.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-48: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel für den indirekten Vergleich

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial auf Studienebene
			Patient	Behandler			
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>
ACCORD I	ja	ja	ja	ja	nein	nein	niedrig
ACCORD II	ja	ja	ja	ja	nein	ja	niedrig
ATTAIN	ja	ja	ja	ja	nein	nein	niedrig
Bateman 2010a	unklar	unklar	ja	ja	ja	ja	niedrig
Bateman 2010b	ja	ja	ja	ja	unklar	nein	niedrig
Beeh 2006	unklar	unklar	ja	ja	unklar	nein	niedrig
Brusasco 2003	unklar	unklar	ja	ja	unklar	ja	hoch
Casaburi 2000	ja	ja	ja	ja	nein	nein	niedrig
Casaburi 2002	unklar	unklar	ja	ja	nein	nein	niedrig
(SAFE)	unklar	unklar	ja	ja	nein	nein	niedrig
Covelli 2005	unklar	unklar	ja	ja	nein	ja	niedrig
INHANCE	ja	nein	nein	nein	nein	nein	niedrig
Donohue 2002	unklar	unklar	ja	ja	nein	nein	niedrig

MISTRAL	unklar	unklar	ja	ja	nein	nein	niedrig
SPRUCE	unklar	ja	ja	ja	ja	nein	hoch
Johansson 2008	unklar	unklar	ja	ja	nein	ja	niedrig
SAFE-Portugal	unklar	unklar	ja	ja	nein	nein	niedrig
Niewoehner 2005	ja	unklar	ja	ja	nein	nein	niedrig
Powrie 2007	unklar	unklar	ja	ja	nein	nein	niedrig
UPLIFT	ja	ja	ja	ja	nein	nein	niedrig
TIPHON	ja	unklar	ja	ja	nein	nein	niedrig
Verkindre 2006	unklar	unklar	ja	ja	nein	nein	niedrig
Vogelmeier 2008	unklar	unklar	nein	nein	ja	nein	hoch
Voshaar 2008	unklar	unklar	ja	ja	nein	nein	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial wurde bei allen Studien insgesamt als gering eingestuft. Einzige Ausnahmen waren Brusasco 2003 [55], SPRUCE [21] und Vogelmeier 2008 [50]. In der Publikation von Brusasco 2003 [55] wurden indirekte Messungen verwendet, um die Ergebnisse der Endpunkte anzugeben. Außerdem fehlten klare Daten zum Patientenfluss und zur Patientenzahl, die den berichteten Ergebnissen zugrunde lagen. In der SPRUCE-Studie [21] wurde das Randomisierungsverfahren nicht eindeutig angegeben, die Subgruppenanalyse wurde nicht vorab definiert und unterschiedliche Subgruppenanalysen wurden unbegründet für die einzelnen Endpunkte vorgelegt. Dies kann als subjektive Berichterstattung angesehen werden. In der Studie von Vogelmeier 2008 [50] wurde das Verzerrungspotenzial ebenfalls als hoch eingestuft, da diese Tiotropiumbromid als unverblindete Behandlung einschloss und die Ergebnisse der Studie auf eine die Tiotropiumbromid-Behandlung bevorzugenden Art und Weise darstellte. Dies zeigte sich in der Art, in der die Ergebnisse in Bezug auf die Exazerbationen berichtet wurden. Die Beschreibung der Ergebnisse unterschied sich von den tatsächlichen Ergebnissen der statistischen Analyse und unterschied sich zudem von der Berichterstattung anderer Studien. Des Weiteren wurde die statistische Analyse des SGRQ nicht beschrieben.

Die Aclidiniumbromid-Studien ACCORD I [40], ACCORD II [19] und ATTAIN [38], die UPLIFT-Studie [45] und die Studien von Casaburi 2000 [56] und Bateman 2010b [64] waren

randomisierte, placebokontrollierte, doppelblinde klinische Studien in der alle Studiencharakteristika eindeutig beschrieben wurden. Bei diesen Studien wurde das Verzerrungspotenzial als gering angesehen.

Die ACCORD II-Studie [19] wies bei Studienbeginn ein Ungleichgewicht zwischen den Behandlungsgruppen bezüglich des FEV₁-Wertes auf (1,249 l in der Gruppe, die mit Acclidiniumbromid 400 µg behandelt wurde, 1,397 l in der Gruppe, die Acclidiniumbromid 200 µg erhielt, und 1,459 l in der Placebo-Gruppe) auf. Außerdem war der Anteil der Patienten mit einer schweren COPD im Stadium III in den Therapiearmen unterschiedlich verteilt (55 % in der Gruppe, die mit Acclidiniumbromid 400 µg behandelt wurde, 48 % in der Gruppe, die Acclidiniumbromid 200 µg erhielt, und 37 % in der Placebo-Gruppe für Stadium III). Die Unterschiede in den Gruppen deuten nicht per se auf ein höheres Verzerrungspotenzial hin, wurden aber auch im europäischen Zulassungsprozess erkannt. Im Zulassungsdossier [66] wurde dem Ungleichgewicht zwischen den Behandlungsgruppen mit Sensitivitätsanalysen begegnet, wobei das Poolen der Ergebnisse ohne die ACCORD II-Studie die primäre Analyse darstellte. Dieser von der EMA akzeptierte Ausschluss der ACCORD II-Studie führt zur konservativen Annahme eines Verzerrungspotentials auf Endpunktebene. Auf Studienebene wird das Verzerrungspotenzial aufgrund nicht erkennbarer Einflüsse auf die Gesamtstudie als gering eingestuft.

Bei mehreren Studien [42, 43, 51, 57, 58, 60, 61, 63] wurde das Randomisierungsverfahren nicht klar dargelegt; auch fehlte eine eindeutige Angabe darüber, inwiefern die Geheimhaltung der Behandlungsfolge gewährleistet war. In den Publikationen wurde angegeben, dass die Studien doppelblind durchgeführt wurden, daher bestand kein Verdacht einer selektiven Berichterstattung. Da die Patientencharakteristika bei Studienbeginn sowie die Ergebnisse die Objektivität der Berichterstattung nicht in Frage stellten, wird das Verzerrungspotenzial bei diesen Studien als gering eingestuft.

In der Publikation von Bateman 2010b [64] wird nicht eindeutig beschrieben, warum die Subgruppenanalyse nur für die LABA-Patienten und nicht für die anderen Subgruppen durchgeführt wurde. Das Verzerrungspotenzial auf Studienebene wurde jedoch insgesamt als gering eingestuft. Eine Studie von Bateman 2010a [62] beschrieb nicht die Randomisierungsmethode und die Art, wie die Geheimhaltung der Behandlungsfolge der Patienten sichergestellt wurde, obwohl angegeben wurde, dass die Studie doppelblind durchgeführt wurde. Die statistische Analyse der Behandlungseffekte war für einige Endpunkte unterschiedlich, diese Unterschiede wurden jedoch nicht begründet. Die Berichterstattung der verschiedenen Endpunkte ist daher unter Umständen subjektiv. Es wird jedoch davon ausgegangen, dass die Unterschiede bei den angewandten statistischen Methoden keine Auswirkung auf das Verzerrungspotenzial auf Studienebene hatten.

Die Publikationen von Beeh 2008 [65], Covelli 2005 [52] und Johansson 2008 [20] lieferten keine eindeutigen und vollständigen Daten zur Art und Weise, wie die Randomisierung durchgeführt wurde und in welchem Ausmaß die Geheimhaltung der Behandlungsfolge sichergestellt wurde. Beeh berichtete auch von Ergebnissen, die nicht vorab definiert wurden, die Berichterstattung schien jedoch nicht verzerrt zu sein. In der Studie von Niewoehner 2005

[53] und der TIPHON-Studie [44] wurde die Randomisierung klar beschrieben, aber es blieb unklar, wie die Randomisierung und die Zuweisung geheim gehalten wurden.

Das Verzerrungspotenzial auf Studienebene wurde bei der INHANCE-Studie [59] insgesamt als gering eingestuft, da die Berichterstattung vollständig und objektiv erfolgte. Tiotropiumbromid wurde den Studienteilnehmern unverblindet verabreicht, wodurch möglicherweise die Ergebnisse von mehreren Endpunkten beeinflusst wurden.

Da das Verzerrungspotenzial bei einem Großteil der Studien (mit Ausnahme von drei Studien) als gering eingestuft wurde, werden diese Studien basierend auf ihrem Verzerrungspotenzial auf Studienebene als verlässliche Datenquellen für die Analyse angesehen.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Im folgenden Kapitel werden die Ergebnisse des indirekten Vergleichs von Acclidiniumbromid und Tiotropiumbromid berichtet. Es werden pro Endpunkt zwei Erhebungszeitpunkte (12 und 24 Wochen) und zwei Studienpools (mit und ohne erlaubte LABA-Begleitmedikation) ausgewertet. Zuvor werden für die Acclidiniumbromid- und Tiotropiumbromid-Studien pro End- und Erhebungszeitpunkt Meta-Analysen durchgeführt. Bei der Ergebnisdarstellung werden für jeden Endpunkt folgende Schritte vorgenommen:

- Auflistung der für die NMA des jeweiligen Endpunkt zur Verfügung stehenden Studien mit und ohne LABA-Begleitmedikation sowie für die Erhebungszeitpunkte 12 und 24 Wochen inklusive Netzwerkstruktur
- Darstellung der Operationalisierung des jeweiligen Endpunktes in allen verfügbaren Studien
- Darstellung des Verzerrungspotentials des jeweiligen Endpunktes in allen verfügbaren Studien
- Tabellarische Darstellung der Endpunktergebnisse nach 12 und 24 Wochen aller zur Verfügung stehenden Studien im Studienpool
 - Soweit vorhanden²⁴ werden dabei pro relevantem Behandlungsarm Baseline-, Follow-Up- und CFB-Werte aufgelistet
 - Der Gruppenunterschied stellt die Differenz der CFB-Werte²⁵ dar und wird als Gruppendifferenz oder OR mit zugehörigem Konfidenzintervall und p-Wert angegeben.
- Durchführung der Meta-Analysen und Darstellung des Forrest-Plots für jeden Endpunkt nach 12 und 24 Wochen
- Durchführung von Sensitivitätsanalysen und Bewertung der Auswirkung auf die Homogenität des Studienpools und den Gesamteffektschätzers

²⁴ Die Publikationen der Tiotropiumbromid-Studien berichten häufig nur die CFB-Werte.

²⁵ Die Verwendung der CFB-Werte gewährleistet Konsistenz zwischen der Ergebnisdarstellung der Acclidiniumbromid- und Tiotropiumbromid-Studien. In den Acclidiniumbromid-Studien waren im Gegensatz zu den Tiotropiumbromid-Studien auch Gruppendifferenzen aus dem ANCOVA-Modell angegeben. Diese unterscheiden sich aufgrund des Designs als RCT kaum von der Differenz der CFB.

- Ermittlung des geeigneten Modells (random oder fixed effect Modell) für den indirekten Vergleich mittels DIC
- Durchführung von zwei NMAs je Endpunkt; gesamter Studienpool und Studienpool, der nur Studien ohne LABA-Begleitmedikation enthält

Die relativen Effektschätzer werden als paarweise Vergleiche tabellarisch dargestellt. Dabei befinden sich die analysierten Interventionen in den Spalten und die Vergleichssubstanzen in den Zeilen. Bei der Beschreibung der NMA-Ergebnisse im Text, werden die in der Tabelle dargestellten Werte um eine Stelle nach dem Komma gerundet.

Tabelle 4.3-49: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche nach 12 Wochen

Studie	Morbidity									Quality of Life		Unwanted Drug Effects			Mortality	
	Trough -FEV ₁	Peak-FEV ₁	nightly symptoms	morning symptoms	EXACT	TDI	TDI Responder	Exacerbations*	hospitalizations due to exacerbations	SGRQ	SGRQ Responder	AEs	SAEs	Study discontinuations	Cardiovascular morbidity and mortality	Letality, Total mortality
ACCORD I	ja	ja	ja	nein	nein	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja
ACCORD II	ja	ja	nein	nein	nein	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja
ATTAIN	ja	ja	ja	ja	ja	ja	ja	ja	nein	ja	ja	nein	nein	nein	nein	nein
Bateman 2010a	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein	nein	nein
Bateman 2010b	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein	nein	nein
Brusasco 2003	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein	nein	nein
Beeh 2006	ja	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Casaburi 2000	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein
Casaburi 2002	ja	nein	nein	nein	nein	ja	nein	ja	nein	nein	nein	nein	nein	ja	nein	nein
SAFE	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Covelli 2005	ja	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	ja
INHANCE	ja	nein	nein	nein	nein	ja	ja	ja	nein	ja	nein	nein	nein	nein	nein	nein
Donohue 2002	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
MISTRAL	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein	nein	nein
SPRUCE	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	ja	nein	ja	nein	nein
Johansson 2008	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	ja	ja	ja	nein	nein
SAFE-Portugal	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	ja	ja	ja	ja	ja

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Niewoehner 2005	ja	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein	nein	nein
Powrie 2007	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein	nein	nein
UPLIFT	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein	nein	nein
TIPHON	nein	nein	nein	nein	nein	nein	nein	ja	nein	ja	ja	nein	nein	nein	nein	nein
Verkindre 2006	ja	ja	nein	nein	nein	ja	nein	nein	nein	ja	ja	nein	nein	ja	nein	nein
Vogelmeier 2008	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Voshaar 2008	ja	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	ja	ja	ja	nein	ja

* = Exazerbationsraten werden anhand der Anzahl an Personenjahren in der Studie berechnet, nicht nach 12 oder 24 Wochen. Die Endpunktekategorie ist in dieser Tabelle, aber nicht in der Tabelle mit 24 Wochen aufgeführt.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4.3-50: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche nach 24 Wochen

Studie	Morbidität								Lebensqualität		Unerwünschte Arzneimittelwirkungen			Mortalität	
	Trough- FEV ₁	Peak- FEV ₁	nächtliche Symptome	morgendliche Symptome	EXACT	TDI	TDI Responder	Krankenhausaufenthalte wegen Exazerbationen	SGRQ	SGRQ Responder	AEs	SAEs	Studienabbrüche	Kardiovaskuläre Morbidität und Mortalität	Letalität, Gesamtmortalität
ACCORD I	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
ACCORD II	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
ATTAIN	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja
Bateman 2010a	ja	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Bateman 2010b	ja	nein	nein	nein	nein	nein	nein	nein	ja	ja	nein	nein	nein	nein	nein
Brusasco 2003	ja	nein	nein	nein	nein	ja	ja	ja	ja	ja	nein	nein	nein	nein	nein
Beeh 2006	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Casaburi 2000	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Casaburi 2002	ja	nein	nein	nein	nein	ja	nein	nein	ja	nein	nein	nein	nein	nein	nein
SAFE	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Covelli 2005	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
INHANCE	ja	ja	nein	nein	nein	ja	ja	nein	ja	ja	ja	ja	ja	nein	ja
Donohue 2002	ja	ja	nein	nein	nein	ja	ja	nein	ja	ja	nein	nein	nein	nein	nein
MISTRAL	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
SPRUCE	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Johansson 2008	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
SAFE-	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Portugal															
Niewoehner 2005	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	ja	nein	nein	nein
Powrie 2007	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
UPLIFT	ja	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein
TIPHON	nein	nein	nein	nein	nein	nein	nein	nein	ja	nein	nein	nein	nein	nein	nein
Verkindre 2006	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein
Vogelmeier 2008	nein	nein	nein	nein	nein	nein	nein	ja	ja	nein	ja	nein	ja	nein	ja
Voshaar 2008	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein	nein

4.3.2.1.3.1 CFB Trough-FEV₁ – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-51: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Trough-FEV₁ – 12 Wochen – begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acidinium-bromid 400 µg BID	Tiotropium-bromid 5µg; OD	Tiotropium-bromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
5	Beeh 2006			•	•
	Casaburi 2000			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	Verkindre 2006			•	•
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Neun Studien gaben den CFB Trough-FEV₁ (begleitende Behandlung mit LABAs nicht gestattet) nach 12 Wochen an; alle drei Acidiniumbromid-Studien, fünf Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 6 stellt das Netzwerk dar, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 6: Netzwerkdiagramm für die Analyse des CFB Trough-FEV₁ nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

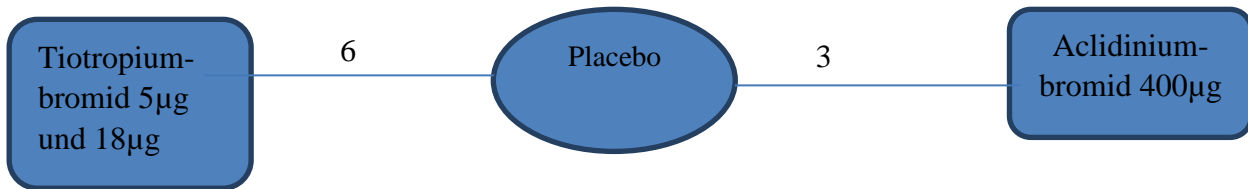


Tabelle 4.3-52: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Trough-FEV₁ – 12 Wochen – alle Studien]

Anzahl Studien	Referenzstudien	Acclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
11	Beeh 2006			•	•
	Casaburi 2000			•	•
	Casaburi 2002			•	•
	SAFE			•	•
	Covelli 2005			•	•
	INHANCE			•	•
	SPRUCE			•	•
	Johansson 2008			•	•
	SAFE Portugal			•	•
	Niewoehner 2005			•	•
	Verkindre 2006			•	•
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

15 Studien gaben den CFB Trough-FEV₁ nach 12 Wochen an; alle drei Acridiniumbromid-Studien, 11 Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 7 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 7: Netzwerkdiagramm für die Analyse des CFB Trough-FEV₁ nach 12 Wochen (alle Studien)

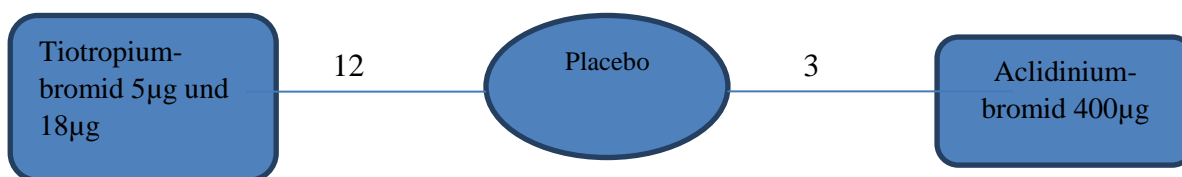


Tabelle 4.3-53: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Trough-FEV₁ - 24 Wochen- begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
4	Brusasco 2003			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	Donohue 2002			•	•
1	Bateman 2010a		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sechs Studien gaben den CFB Trough-FEV₁ (begleitende Behandlung mit LABAs nicht gestattet) nach 24 Wochen an; eine Acridiniumbromid-Studie, vier Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Ein indirekter Vergleich konnte für alle Behandlungen innerhalb des Netzwerkes durchgeführt werden. Abbildung 8 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 8: Netzwerkdiagramm für die Analyse des CFB Trough-FEV₁ nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

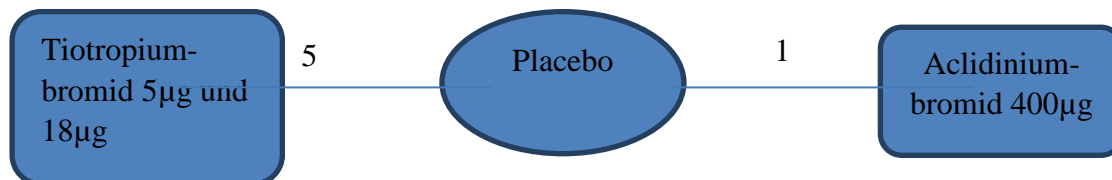


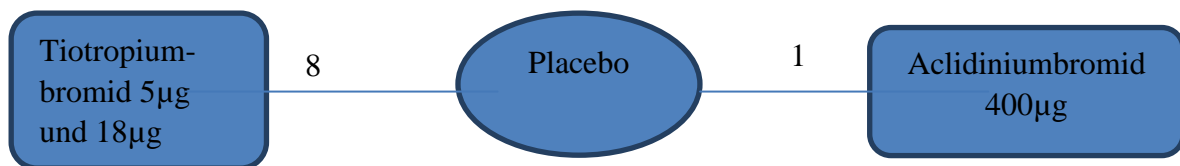
Tabelle 4.3-54: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Trough-FEV₁ - 24 Wochen – alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
6	Brusasco 2003			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	Donohue 2002			•	•
	Niewoehner 2005			•	•
	UPLIFT			•	•
2	Bateman 2010a		•		•
	Bateman 2010b		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Neun Studien gaben den CFB Trough-FEV₁ nach 24 Wochen an; eine Acclidiniumbromid-Studie, sechs Tiotropiumbromid-18µg-Studien und zwei Tiotropiumbromid-5µg-Studien. Abbildung 9 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann. Ein indirekter Vergleich konnte für alle Behandlungen innerhalb des Netzwerkes durchgeführt werden.

Abbildung 9: Netzwerkdiagramm für die Analyse des CFB Trough-FEV₁ nach 24 Wochen (alle Studien)



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-55: Operationalisierung von CFB Trough-FEV₁

Studie	Operationalisierung
ACCORD I	Der Mittelwert der zwei höchsten bei jeder Beobachtungsvisite, 11 und 12 Stunden nach der vorabendlichen Gabe des Studienmedikaments, erhaltenen FEV ₁ -Werte (einschließlich Woche 12 für den Primärwert).
ACCORD II	Der Mittelwert der zwei höchsten bei jeder Beobachtungsvisite, 11 Stunden und 11 Stunden 50 Minuten nach der vorabendlichen Gabe des Prüfmedikaments, erhaltenen FEV ₁ -Werte.
ATTAIN	Standardisierte FEV ₁ -Messungen wurden vor der morgendlichen Gabe durchgeführt.
Beeh 2006	Morgendlicher Trough-FEV ₁ - 24 Stunden nach der vorhergehenden Inhalation der Prüfmedikation. Es werden arithmetische Mittelwerte mit Standardabweichung angegeben. Zur Ermittlung des CFB wird die Lungenfunktion (FEV ₁) zur Baselineuntersuchung 30 und 10 Minuten vor der Inhalation der Prüfmedikation gemessen.
Brusasco 2003	Der Mittelwert der beiden FEV ₁ -Messungen wurde nach ATS-Richtlinien, 60 und 10 Minuten vor der Dosierung (d.h. 23-24 Stunden nach der letzten Tiotropiumbromid-Gabe) bestimmt.
Casaburi 2000	Der Trough-FEV ₁ einer beliebigen Visite wurde als der Mittelwert der beiden FEV ₁ -Werte, gemessen am Ende des Dosierungsintervalls etwa 23 bis 24 h nach der Studienmedikation, definiert. Die trough-FEV ₁ -Response wurde als Veränderung vom Ausgangswert im trough-FEV ₁ definiert und wurde als der primäre Endpunkt angesehen, um die 24-h Wirkdauer von Tiotropiumbromid festzulegen.
Casaburi 2002	Der FEV ₁ wurde 1 Stunde vor und unmittelbar vor der Gabe des Prüfmedikaments gemessen. Spirometrische Untersuchungen wurden dreifach durchgeführt und der höchste FEV ₁ -Wert in nachfolgenden Analysen verwendet.
SAFE	Der FEV ₁ wurde 10 Minuten vor der Dosierung (trough) aufgezeichnet. Die trough-FEV ₁ -Response wurde als Pre-Dosis-Wert bei jeder Klinikvisite nach Randomisierung, d. h. 23 bis 24 Stunden nach der letzten Gabe der Versuchsmedikation bestimmt.
Covelli 2005	Der FEV ₁ wurde vor der Verabreichung der Prüfmedikation gemessen. Das Spirometer wurde jeden Morgen vor der Durchführung der Tests kalibriert. Die Tests wurden dreifach durchgeführt und der beste der drei Versuche wurde als höchster bei der Ausatmung erhaltener FEV ₁ definiert (nicht unbedingt aus der gleichen Kurve). Dieses Verfahren entspricht den Kriterien der American Thoracic Society.
INHANCE	24 Stunden nach der Dosierung gemessener FEV ₁ (Mittelwert 23h 10min und 23h 45min nach der Dosierung)
Donohue 2002	Der FEV ₁ wurde vor Beginn der Therapie 60min und 10min vor der Gabe des Prüfmedikamentes gemessen. Die Messungen wurden nach den Kriterien der American Thoracic Society durchgeführt. Die höchsten FEV ₁ -Messwerte wurden beibehalten. Fehlende Werte wurden anhand von anderen Werten für den Patienten an diesem Testtag geschätzt. Es wurde eine lineare Interpolation zwischen zwei benachbarten Messungen verwendet, um die fehlenden Spirometriewerte in der Mitte des Profils abzuschätzen. Für fehlende Werte am Ende der Profile, die durch Notfallmedikation verursacht wurden, wurde der minimale beobachtete FEV ₁ -Wert an diesem Testtag als Schätzwert verwendet. Der letzte verfügbare Wert wurde als Schätzwert für Daten verwendet, die aus administrativen Gründen nicht aus der Patienten-Response auf die Behandlung gewonnen werden konnten.
SPRUCE	Die trough-FEV ₁ -Response wurde als Veränderung vom Ausgangswert am Ende des 24-Stunden-Dosierungs-Intervalls (z. B. 10 Minuten vor der Verabreichung des Medikaments) definiert. Der Ausgangs-FEV ₁ war der vor der Behandlung bei Visite 2

	<p>10 Minuten vor der Verabreichung der ersten Dosis der Studienmedikation gemessene FEV₁.</p> <p>Die Patienten wurden gebeten, ihre morgendlichen Atem-Medikamente nicht vor dem morgendlichen LFP (PFT - pulmonary function test) einzunehmen (entsprechend einer vorher festgelegten Auswaschphase). Der höchste FEV₁-Wert von drei technisch reliablen Untersuchungen wurde aufgezeichnet.</p>
Johansson 2008	<p>Die FEV₁-trough-Response war der Vordosis-Wert an den Tagen 15 und 85. Spirometrie wurde beim Screening und an den Tagen 1, 15 und 85 durchgeführt. Der FEV₁ wurde 10 Minuten vor der Dosierung aufgezeichnet. Alle spirometrischen Tests wurden dreifach durchgeführt, und die höchsten Messungen wurden für die Analyse verwendet. Die Ausrüstung und die verwendeten Methoden erfüllt ATS (1994) oder die Standards der European Respiratory Society (1993).</p>
SAFE Portugal	<p>FEV₁ wurde an Test-Tagen 10 min (\pm 5 min) vor der Verabreichung der Studienmedikation (trough) gemessen. Die Messungen wurden unter Verwendung eines Datospir 120C-Spirometers (Sibelmed, Barcelona, Spanien) gemäß den Kriterien der American Thoracic Society dreifach durchgeführt</p> <p>Der trough-FEV₁ wurde 24 Stunden nach der vorhergehenden Dosierung des Studienmedikaments am Tag 1 (Randomisierung) wie auch nach 6 und 12 Behandlungswochen gemessen.</p>
Niewoehner 2005	<p>Die Teilnehmer führten eine Spirometrie vor Inhalation des Studienmedikaments als Ausgangswert und nochmals bei der 3-Monats- und 6-Monats-Visite durch. Die Studienzentren führten die Spirometrie unter Verwendung eines üblichen prädiktiven Nomogramms mit Geräten und Methoden durch, die mit den Empfehlungen der American Thoracic Society übereinstimmen.</p>
UPLIFT	<p>Der trough-FEV₁ ist der etwa 24 Stunden nach der letzten Gabe der Studienmedikation gemessene Vordosis-Wert.</p>
Verkindre 2006	<p>Alle spirometrischen Daten wurden 30min vor der Dosierung aufgezeichnet. Die FEV₁-trough-Response wurde als Vordosis-Wert bei jeder Klinikvisite nach Randomisierung, d.h. ca. 23 bis 24 Stunden nach der letzten Gabe der Versuchsmedikation bestimmt.</p>
Bateman 2010a	<p>FEV₁ trough-Response in Woche 48 (24-Stunden-Nachverabreichungs-FEV₁ wurde als Veränderung des Baseline-Studien-Vordosis-FEV₁ ausgedrückt).</p>
Bateman 2010b	<p>trough-FEV₁ = Vordosis-FEV₁</p>
Voshaar 2008	<p>Primärer Endpunkt war die Veränderung beim trough-(d.h. morgendlichen Vordosis) FEV₁ nach 12 Behandlungswochen. Spirometrische Messungen wurden in Übereinstimmung mit den Kriterien der American Thoracic Society bei 10 (\pm 5)min Vordosis durchgeführt.</p> <p>Der trough-FEV₁ ist eine wichtige Messgröße, da sie die minimale durch den Patienten ausgedrückte Verbesserung im Laufe der Dosierungszeit darstellt und, für Tiotropiumbromid, die ganztägigen Auswirkungen einer vorhergehenden Dosierung zeigt.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-56: Bewertung des Verzerrungspotenzials für CFB Trough-FEV₁ in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Beeh 2006	Niedrig	Ja	Unklar	Unklar	Nein	Niedrig
Brusasco 2003	Niedrig	Ja	Nein	Unklar	Nein	Hoch
Casaburi 2000	Niedrig	Ja	Nein	Nein	Nein	Hoch
Casaburi 2002	Niedrig	Ja	Ja	Nein	Nein	Niedrig
SAFE	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Covelli 2005	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
INHANCE	Niedrig	Nein	Nein	Nein	Nein	Hoch
Donohue 2002	Niedrig	Ja	Nein	Nein	Nein	Hoch
SPRUCE	Hoch	Ja	Nein	Nein	Nein	Hoch
Johansson 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
SAFE Portugal	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Niewoehner 2005	Niedrig	Ja	Ja	Nein	Nein	Niedrig
UPLIFT	Niedrig	Ja	Nein	Nein	Nein	Hoch
Verkindre 2006	Niedrig	Ja	Nein	Nein	Nein	Hoch
Bateman 2010a	Niedrig	Ja	Nein	Nein	Ja	Hoch
Bateman 2010b	Niedrig	Ja	Nein	Unklar	Nein	Hoch
Voshaar 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Tabelle 4.3-56 zeigt, dass sich für alle Studien mit Ausnahme der SPRUCE-Studie ein geringes Verzerrungspotenzial auf Studienebene ergab. Unter Berücksichtigung des spezifischen Verzerrungspotenzials für den CFB Trough-FEV₁-Wert ergab sich bei 10 von 20 Studien ein hohes Verzerrungspotenzial aufgrund der schlechten Vergleichbarkeit der Baseline-Charakteristika der randomisierten Patienten (ACCORD II) oder der Tatsache, dass das ITT-Prinzip nicht für die Endpunkt-Analyse angewendet wurde (Bruscano 2003, Casaburi

2000, Donohue 2002, INHANCE, SPRUCE, UPLIFT, Verkindre 2006, Bateman 2010b, Bateman 2010a).

Stellen Sie die Ergebnisse für den CFB Trough-FEV₁ für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-57: Ergebnisse für CFB Trough-FEV₁ (ml) aus RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mittelwert (SD)	Mittelwert (SD)	CFB (SD)	Gruppendifferenz (SE) KI [] p-Wert
ACCORD I	Aclidiniumbromid	190	1332 (493)	1433 (510)	102 (207)	128 (20,4) [86,4; 169,6] p < 0,001
	Placebo	185	1376 (570)	1349 (585)	-26 (204)	
ACCORD II	Aclidiniumbromid	177	1249 (519)	1323 (511,8)	73 (194)	86 (19,8) [45,6; 126,4] p = 0,001
	Placebo	182	1459 (519)	1446 (522)	-13 (197)	
ATTAIN	Aclidiniumbromid	269	1508 (494)	1573 (246)	65 (262)	101 (21,1) [56,7; 145,3] p < 0,001
	Placebo	273	1500 (497)	1464 (248)	-36 (264)	
Beeh 2006	Tiotropiumbromid 18µg	1236	1320 (480)	(-)	(-)	79 (17) [45,7; 112,3] p < 0,001
	Placebo	403	1360 (500)	(-)	(-)	
Casaburi 2000	Tiotropiumbromid 18µg	276	1040 (420)	1150 (166,1)	110 (-)	150 (14,1) [122,3; 177,7] p < 0,001
	Placebo	188	1000 (430)	960 (137,1)	-40 (-)	
Casaburi 2002	Tiotropiumbromid 18µg	507	1040 (410)	1143 (-)	103 (-)	114 (15,0) [84,6; 143,5] p < 0,001
	Placebo	332	1000 (440)	989 (-)	-11 (-)	
SAFE (Chan 2007)	Tiotropiumbromid 18µg	587	970 (390)	1079 (157,5)	109 (-)	98 (13,3) [72; 124] p < 0,001
	Placebo	287	960 (380)	981 (84,7)	21 (-)	
Covelli 2005	Tiotropiumbromid 18µg	94	1060 (370)	(-)	185 (-)	184 (37) [111,5; 256,5] p < 0,001
	Placebo	84	990 (420)	(-)	1 (-)	
INHANCE (Donohue 2010)	Tiotropiumbromid 18µg	415	1450 (505)	1420 (-)	(-)	151 (14,9) [121,3; 180,7] p < 0,001
	Placebo	418	1510 (490)	1280 (-)	(-)	
SPRUCE (Freeman 2007)	Tiotropiumbromid 18µg	191	1250 (420)	(-)	90 (-)	60 (19,3) [21,4; 98,6] p = 0,002
	Placebo	183	1320 (440)	(-)	30 (-)	

Johansson 2008	Tiotropiumbromid 18µg	107	2150 (590)	(-)	(-)	118 (31) [57,2; 178,8]
	Placebo	117	2010 (550)	(-)	(-)	p < 0,001
SAFE Portugal (Moita 2008)	Tiotropiumbromid 18µg	144	(-)	(-)	94 (-)	102 (30,8) [40,5; 163,5]
	Placebo	160	(-)	(-)	-23 (-)	p = 0,001
Niewoehner 2005	Tiotropiumbromid 18µg	914	1040 (400)	(-)	(-)	100 (10,2) [80; 120]
	Placebo	915	1040 (400)	(-)	(-)	p < 0,001
Verkindre 2006	Tiotropiumbromid 18µg	46	1050 (350)	1170 (203,5)	(-)	110 (40) [-0,87; 220,9]
	Placebo	54	1080 (310)	1060 (220,5)	(-)	p = 0,05
Voshaar 2008	Tiotropiumbromid 5µg	180	(-)	(-)	102 (-)	118 (23) [72,9; 163,1]
	Placebo	181	(-)	(-)	-16 (-)	p < 0,001

N=Anzahl der Patienten, CFB=Veränderung gegenüber dem Ausgangswert, SE=Standardfehler
SD=Standardabweichung, ITT=Intention-to-treat-Population, KI=Konfidenzintervall

Tabelle 4.3-57 zeigt die einzelnen Studienergebnisse für den 12 Wochen nach Studienbeginn ermittelten CFB Trough-FEV₁. In allen Studien zeigten die mit Acclidiniumbromid oder Tiotropiumbromid behandelten Patienten verbesserte Spirometrie-Ergebnisse.

Tabelle 4.3-58: Ergebnisse für CFB Trough-FEV₁ (ml) aus RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten	Baseline	Follow-up	Veränderung	
		(ITT)	Mittelwert (SD)	Mittelwert (SD)	CFB (SD)	Gruppendifferenz (SE) [KI] p-Wert
ATTAIN	Acidiniumbromid	269	1508 (494)	1573 (246)	66 (274)	128 (22) [85; 170] p < 0,001
	Placebo	273	1500 (497)	1442 (248)	-58 (244)	
Bateman 2010a	Tiotropiumbromid 5µg	670	(-)	(-)	112 (-)	123 (14,8) [94,0; 152,0] p < 0,001
	Placebo	653	(-)	(-)	-11 (-)	
Bateman 2010b	Tiotropiumbromid 5µg	1989	(-)	(-)	121 (268)	103 (8,5) [86,3; 119,7] p < 0,001
	Placebo	2002	(-)	(-)	18 (358)	
Brusasco 2003	Tiotropiumbromid 18µg	386	1120 (390)	(-)	(-)	120 (19,0) [82,7; 157,3] p < 0,001
	Placebo	362	1090 (400)	(-)	(-)	
Casaburi 2002	Tiotropiumbromid 18µg	507	1040 (410)	1130 (-)	(-)	123 (18,1) [87,5; 158,5] p < 0,001
	Placebo	332	1000 (440)	967 (-)	(-)	
Donohue 2002	Tiotropiumbromid 18µg	209	1110 (390)	(-)	(-)	137 (20) [97,8; 176,2] p < 0,001
	Placebo	201	1060 (360)	(-)	(-)	
INHANCE (Donohue 2010)	Tiotropiumbromid 18µg	415	1450 (505)	(-)	250 (-)	140 (20) [100; 180] p < 0,001
	Placebo	418	1510 (490)	(-)	70 (-)	
Niewoehner 2005	Tiotropiumbromid 18µg	914	1040 (400)	(-)	(-)	100 (12,8) [75; 125] p < 0,001
	Placebo	915	1040 (400)	(-)	(-)	
UPLIFT (Tashkin 2010)	Tiotropiumbromid 18µg	2987	1100 (400)	1226 (136,6)	(-)	90 (7,0) [76,3; 103,7] p < 0,001
	Placebo	3006	1090 (400)	1126 (109,7)	(-)	

N=Anzahl der Patienten, CFB=Veränderung gegenüber dem Ausgangswert, SE=Standardfehler SD=Standardabweichung, ITT=Intention-to-treat-Population, KI=Konfidenzintervall

Die Tabelle 4.3-58 stellt die einzelnen Studienresultate zum Endpunkt CFB Trough-FEV₁ 24 Wochen nach Studienbeginn dar. In allen Studien verbesserten Acidiniumbromid bzw. Tiotropiumbromid den Trough-FEV₁. Da nur eine Acidiniumbromidstudie den Endpunkt CFB Trough-FEV₁ nach 24 Wochen untersuchte, wurde für diesen Endpunkt mit Acidiniumbromid keine Metaanalyse durchgeführt.

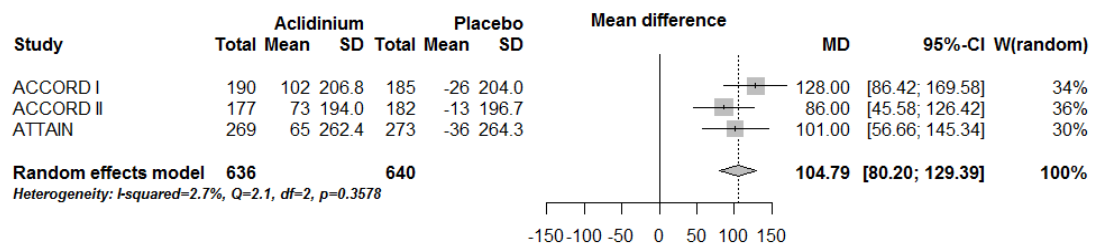


Abbildung 10: Meta-Analyse für CFB Trough-FEV₁ nach 12 Wochen (ml) aus RCT; Acridiniumbromid im Vergleich zu Placebo

Die Meta-Analyse der Acridiniumbromidstudien in Abbildung 10 zeigt die überlegene Wirksamkeit von Acridiniumbromid auf den Endpunkt CFB Trough-FEV₁ nach 12 Wochen im Vergleich zu Placebo (mittlere Differenz: 104,79 ml [95 % CI: 80,2; 129,4]). Der Heterogenitätstest bestätigt die Homogenität der drei Studien (Q = 2,1, p = 0.36, I² = 2.7 %).

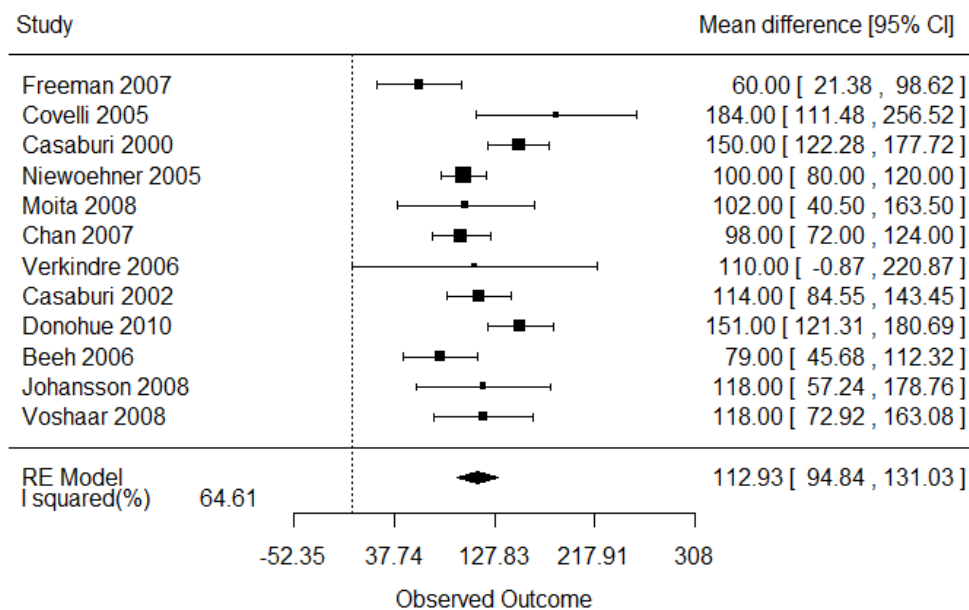


Abbildung 11: Meta-Analyse für CFB Trough-FEV₁ nach 12 Wochen (ml) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.

Die Metaanalyse der Tiotropiumstudien in Abbildung 11 zeigt ebenfalls eine höhere Effektivität von Tiotropiumbromid im Vergleich zu Placebo beim Endpunkt CFB Trough-

FEV₁ nach 12 Wochen (mittlere Differenz 112,9 ml [95 % CI: 94,8; 131,0]). Der Heterogenität-Test ergab jedoch eine signifikante Heterogenität der Studien ($Q = 31,1$, $p = 0,001$, $I^2 = 64,6$ %). Da aber Tiotropiumbromid mit Ausnahme einer kleineren Studie (Verkindre 2006) in allen Studien den CFB Trough-FEV₁ signifikant im Vergleich zu Placebo verbesserte, kann dieser Gesamteffektschätzer dennoch verwendet werden.

Mögliche Gründe für die Heterogenität sind Unterschiede bei der Krankheitsschwere in den untersuchten Populationen (die Studien von Freeman 2007 (SPRUCE) und Johansson 2008 schlossen fast nur Patienten mit mittelschwerer COPD ein) oder beim verwendeten Inhalator (HandiHaler[®]/ Respimat[®]) Dosis (18 µg oder 5 µg). Um Gründe für die Unterschiede zu identifizieren, wurde eine Sensitivitätsanalyse durchgeführt.

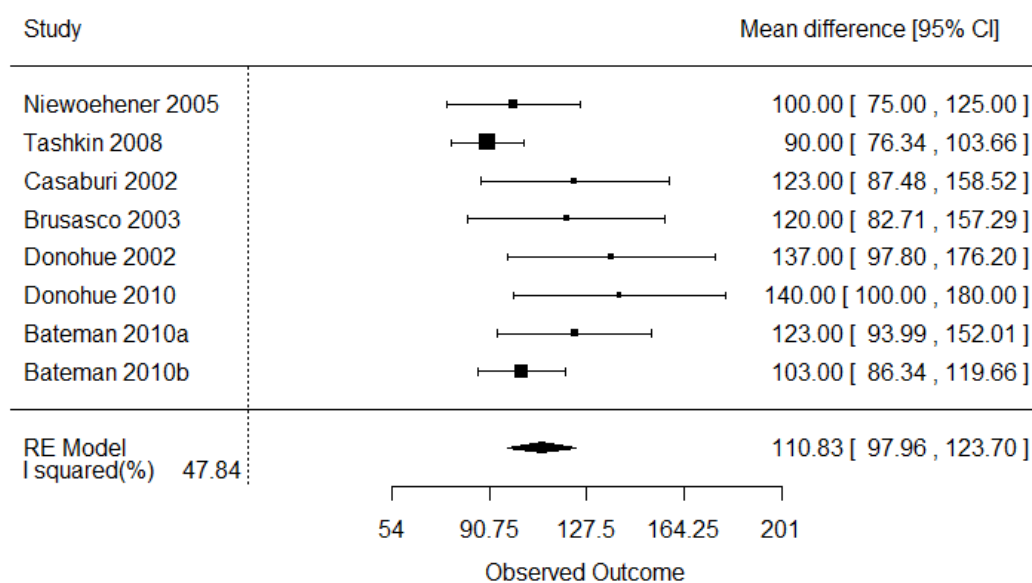


Abbildung 12: Meta-Analyse für CFB Trough-FEV₁ nach 24 Wochen (ml) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.

Die Metaanalyse (Abbildung 12) zeigt die überlegene Wirksamkeit von Tiotropiumbromid auf den Endpunkt CFB Trough-FEV₁ nach 24 Wochen im Vergleich zu Placebo (mittlere Differenz 110,8 ml [95 % CI 98,0; 123,7]). Der Heterogenitätstest ergab eine Heterogenität der Studien ($Q = 13,4$, $p = 0,06$, $I^2 = 57,4$ %). Zur Identifizierung der möglichen Ursache der Heterogenität wird eine Sensitivitätsanalyse durchgeführt.

Sensitivitätsanalyse

Die ACCORD-II-Studie wies hinsichtlich der Baseline-Charakteristika ein Ungleichgewicht zwischen den Behandlungsgruppen auf. Um die Auswirkungen des Ausschlusses dieser Studie von der Meta-Analyse zu erforschen, wurde eine Sensitivitätsanalyse durchgeführt.

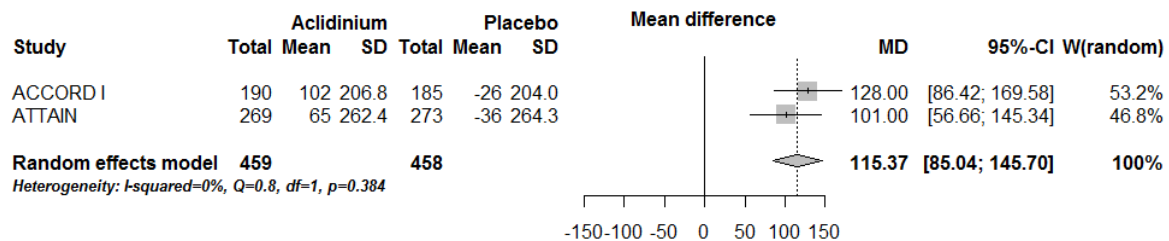


Abbildung 13: Sensitivitätsanalyse (Verzerrungspotential) für CFB Trough-FEV₁ nach 12 Wochen (ml) aus RCT; Acidiniumbromid im Vergleich zu Placebo

Die Sensitivitätsanalyse (Abbildung 13) bestätigt, die überlegene Wirksamkeit von Acidiniumbromid auf den Endpunkt CFB Trough-FEV₁ nach 12 Wochen im Vergleich zu Placebo (mittlere Differenz 115,3 ml [CI: 95% 85,0; 145,7]). Der Ausschluss der ACCORD II-Studie hatte keinen signifikanten Effekt auf die kombinierten Ergebnisse, wie das Konfidenzintervall der Metaanalyse zeigt. Der Heterogenitätstest ergab keine Hinweise auf Heterogenität ($Q = 0,8$, $p < 0,38$, $I^2 = 0,0 \%$).

Um den Einfluss des COPD-Schweregrades auf den Behandlungseffekt zu analysieren, wurden die Studien, die fast nur Patienten mit mittelschwerer COPD untersuchten, ausgeschlossen (Freeman 2007 (SPUCE) und Johansson 2008). Das Ergebnis dieser Analyse zeigt die Abbildung 14.

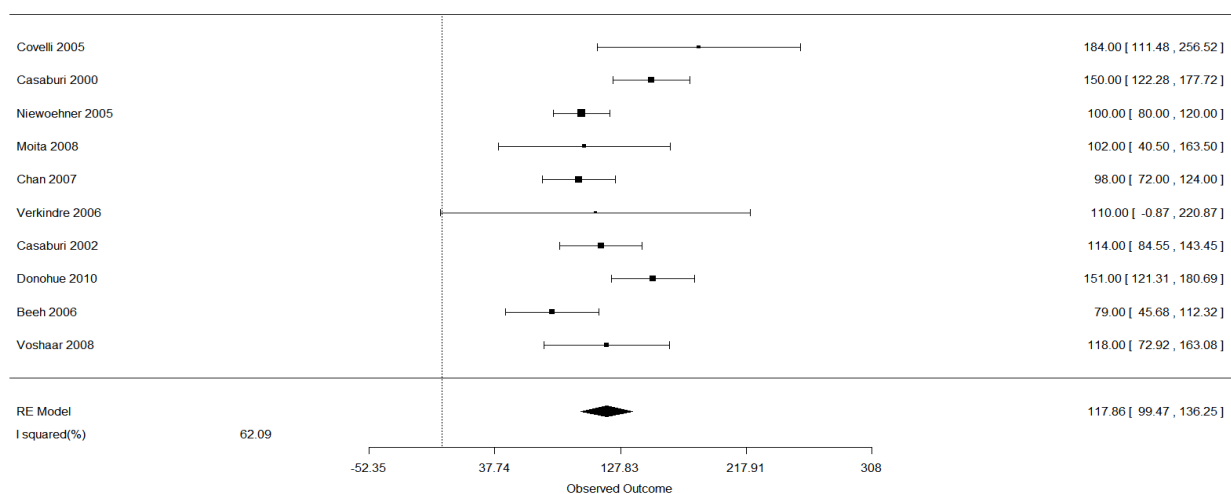


Abbildung 14: Sensitivitätsanalyse (COPD-Schweregrad) für CFB Trough-FEV₁ nach 12 Wochen (ml) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo.

Abbildung 14 zeigt, dass Tiotropiumbromid nach 12 Wochen im Vergleich zu Placebo wirksamer ist (mittlere Differenz 117,9 [KI 95% 99,5; 136,3]). Das Ergebnis ist mit dem Ergebnis der Meta-Analyse vergleichbar, welche SPRUCE und Johansson 2008 einschloss (Abbildung 11). Somit beeinflusst der Schweregrad der Erkrankung nicht notwendigerweise den Behandlungseffekt in Bezug auf den CFB Trough-FEV₁. Der Heterogenitätstest ($Q=23.7$; $p=0.005$; $I^2=62.09$) zeigt weiterhin eine signifikante Heterogenität der Studien.

Um den Einfluss des Verzerrungspotenzials auf den gepoolten Effekt zu untersuchen, wurden Tiotropiumbromid-Studien mit einem hohen Verzerrungspotenzial des nach einem Behandlungszeitraum von 12 Wochen ermittelten CFB Trough-FEV₁ ausgeschlossen. Die Ergebnisse dieser Sensitivitätsanalyse sind in Abbildung 15 dargestellt.

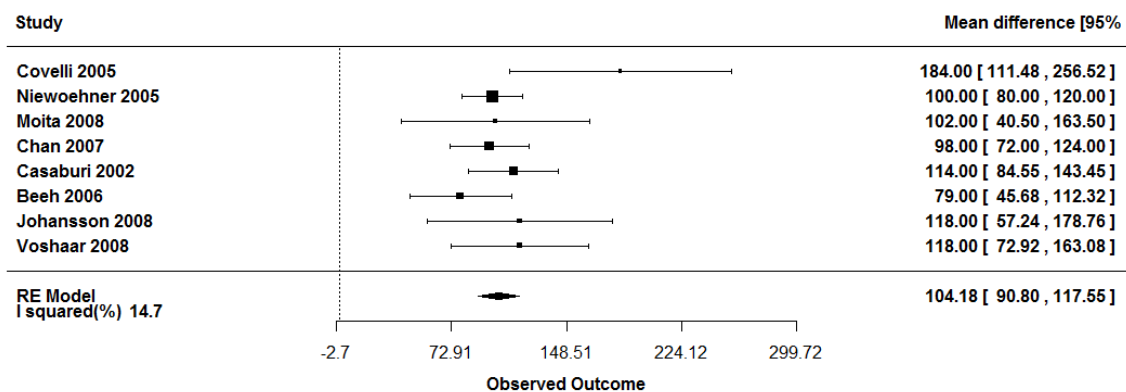


Abbildung 15: Sensitivitätsanalyse (Verzerrungspotential) für CFB Trough-FEV₁ nach 12 Wochen (ml) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo.

Der Ausschluss der Studien mit einem hohen Verzerrungsrisiko führt zu einer Homogenität des Studienpools ($Q = 8,2$, $p < 0,31$, $I^2 = 14,7$ %). Der Ausschluss der verzerrten Studien hatte kaum einen Einfluss auf den Gesamteffektschätzer, wie das Konfidenzintervall der Metaanalyse zeigt (mittlere Differenz 104,2 [95 % CI: 90,1; 117,6]).

Um den Einfluss des Verzerrungspotenzials auf den gepoolten Effekt von Tiotropiumbromid auf den CFB Trough-FEV₁ nach 24 Wochen zu untersuchen, wurde eine weitere Sensitivitätsanalyse, die Studien mit einem hohen Verzerrungspotenzial des nach einem Behandlungszeitraum von 24 Wochen gemessenen CFB Trough-FEV₁ ausschloss, durchgeführt. Die Mehrheit der Studien, die über dieses Outcome berichteten, hatte ein hohes Verzerrungspotenzial, sodass nur zwei Studien verblieben. Die Ergebnisse sind in Abbildung 16 dargestellt.

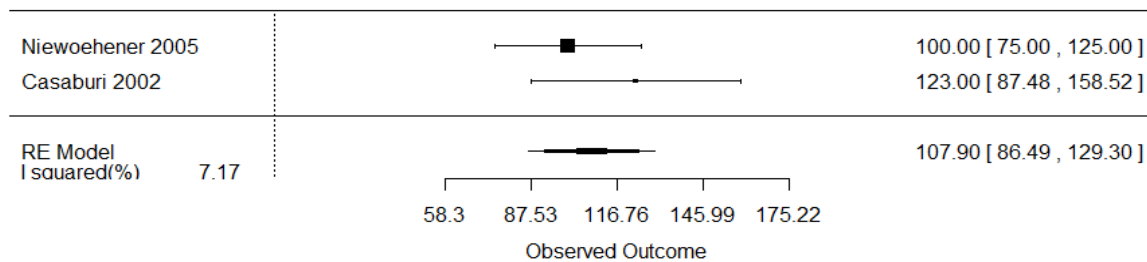


Abbildung 16: Sensitivitätsanalyse (Verzerrungspotential) für CFB Trough-FEV₁ nach 24 Wochen (ml) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo.

Abbildung 16 zeigt, dass der Ausschluss der Studien mit einem hohen Verzerrungsrisiko zur Homogenität des Studienpools führt ($Q=1,1$, $p=0,30$, $I^2=7,2\%$). Der Ausschluss der verzerrten Studien hatte kaum einen Einfluss auf den Gesamteffektschätzer, wie das Konfidenzintervall der Metaanalyse zeigt (mittlere Differenz 107,9 [95%-KI 86,5; 129,3]).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Die Tabelle 4.3-59 zeigt die DIC-Schätzwerte der durchgeführten Analysen. Die DIC bestimmt die Qualität der Anpassung der Analysen an die Daten durch Berechnung der Rest-Abweichung, definiert als die Differenz zwischen der Abweichung der angepassten Analyse und der Abweichung der genügenden Analyse. Die beste Analyse ist die Analyse mit der niedrigsten DIC bei Verwendung der selben Ausgangsdaten.

Für den Endpunkt CFB Trough-FEV₁ nach 12 Wochen war die Random-Effekt-Analyse basierend auf den DIC-Ergebnissen die beste Analyse sowohl für Szenario 1 als auch für Szenario 2 (Tabelle 4.3-59). Die Ergebnisse der Random-Effekt-Analyse sprechen dafür, dass die beobachteten Differenzen der Behandlungseffekte in den Studien nicht nur durch verschiedene Behandlungsvergleiche verursacht wurden, sondern dass es auch eine Heterogenität der relativen Effekte eines bestimmten Vergleichstyps gab. Das war mit den Ergebnissen der Tiotropium-Metaanalyse konsistent, wo der Q-Test für eine signifikante Heterogenität sprach (Abbildung 14).

Für den Endpunkt CFB Trough-FEV₁ nach 24 Wochen sind basierend auf den DIC-Ergebnissen die Fixed-Effekt-Analyse für Szenario 1 und die Random-Effekt-Analyse für Szenario 2 die besser geeigneten Analysen (Tabelle 4.3-59). Dies stimmt mit den Ergebnissen der Tiotropiumbromid-Meta-Analyse überein, in der eine Heterogenität der Studien gezeigt wurde ($p = 0,06$), der Gesamteffektschätzer aber signifikant war.

Tabelle 4.3-59: DIC für CFB Trough-FEV₁ (ml) – indirekter Vergleich aus RCT

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung für die NMA-Modelle
Woche 12			
Ohne LABA-Begleitmedikation	99,809	89.849	Zufällige Effekte
Alle Studien	164,872	153.956	Zufällige Effekte
Woche 24			
Ohne LABA-Begleitmedikation	51.198	60,748	Feste Effekte
Alle Studien	91,954	81.669	Zufällige Effekte

Die Ergebnisse der Netzwerk-Metaanalyse (NMA) für den Endpunkt CFB Trough-FEV₁ nach 12 Wochen und nach 24 Wochen werden in Tabelle 4.3-60, Tabelle 4.3-61, Tabelle 4.3-62 und Tabelle 4.3-63 präsentiert.

Tabelle 4.3-60: CFB Trough-FEV₁ (ml) Indirekter Vergleich aus RCT (12 Wochen); Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit zufälligen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	125.20	
	95% KI	111.50	139.60
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	105.10	-20.14
	95% KI	81.22	128.50
	P(Überlegenheit)	>99%	8%

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-61: CFB Trough-FEV₁ (ml) Indirekter Vergleich aus RCT (12 Wochen);
Szenario 2: Alle Studien; Modell mit zufälligen Effekten

Intervention		Kontrolle		
		Pbo		TIO 5 und 18
TIO 5 und 18	Effektschätzer	111.60		
	95% KI	102.10	121.40	
	P(Überlegenheit)	>99%		
AB 400	Effektschätzer	104.20		-7.41
	95% KI	81.37	127.60	-32.99 17.45
	P(Überlegenheit)	>99%		29%

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Acclidiniumbromid steigerte den Trough-FEV₁ nach 12 Wochen um 105 ml versus Placebo im Szenario 1 (Tabelle 4.3-60; Studien ohne LABA-Begleittherapie). Da das 95 % Konfidenzintervall die Null nicht umfasste (95 % KI-Grenzwerte: 81, 129), kann geschlussfolgert werden, dass Acclidiniumbromid effektiver als Placebo ist. (Wahrscheinlichkeit der Überlegenheit gegenüber Placebo > 99 %). Ähnliche Ergebnisse ergab die Analyse von Szenario 2 (alle Studien Tabelle 4.3-61).

Die Wirksamkeit von Acclidiniumbromid war mit Tiotropiumbromid vergleichbar; es zeigte sich eine mittlere Differenz des CFB Trough-FEV₁ von -20 ml [95 %-KI: -48; 7] in Szenario 1 (Tabelle 4.3-60). In Szenario 2 können ähnliche Schlussfolgerungen gezogen werden. Da das 95 % Konfidenzintervall in beiden Fällen die Null einschloss, kann geschlussfolgert werden, dass Acclidiniumbromid und Tiotropiumbromid eine gleichwertige Effektivität bezüglich des Endpunkts CFB Trough-FEV₁ nach 12 Wochen (Tabelle 4.3-61) haben.

Tabelle 4.3-62: CFB Trough-FEV₁ (ml) Indirekter Vergleich aus RCT (24 Wochen); Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	127.50	
	95% KI	111.70	143.30
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	127.90	
	95% KI	84.78	171.10
	P(Überlegenheit)	>99%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Aclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-63: CFB Trough-FEV₁ (ml) Indirekter Vergleich aus RCT (24 Wochen); Szenario 2: Alle Studien; Modell mit zufälligen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	105.40	
	95% KI	96.71	114.00
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	128.50	
	95% KI	84.05	173.20
	P(Überlegenheit)	>99%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Aclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Die Ergebnisse nach 24 Wochen bestätigen die Beobachtungen nach 12 Wochen: Aclidiniumbromid steigerte den Trough-FEV₁ nach 24 Wochen um 128 ml versus Placebo in beiden Szenarien (Tabelle 4.3-62 und Tabelle 4.3-63). Die Gleichwertigkeit von Aclidiniumbromid und Tiotropiumbromid konnte ebenfalls bestätigt werden. Dies war in Szenario 2 besonders deutlich: Es zeigte sich eine mittlere Differenz des CFB Trough-FEV₁ von 23 ml [95 %-KI: -21; 68] (Tabelle 4.3-63). Da das 95 % Konfidenzintervall in beiden Fällen die Null einschloss, kann geschlussfolgert werden, dass Aclidiniumbromid und Tiotropiumbromid eine gleichwertige Effektivität bezüglich des Endpunkts CFB Trough-FEV₁ nach 24 Wochen (Tabelle 4.3-62 und Tabelle 4.3-63) haben. Die Ergebnisse der NMA sind konsistent zu den Ergebnissen der Metaanalysen: Beide Therapien sind gegenüber Placebo überlegen.

Beide Therapien, Aclidiniumbromid und Tiotropiumbromid, verbesserten den Trough-FEV₁ nach 12 und 24 Wochen statistisch signifikant stärker als Placebo. Der indirekte Vergleich ergab eine gleichwertige Wirkung von Aclidiniumbromid und Tiotropiumbromid bezüglich des Endpunktes CFB Trough-FEV₁. Sowohl nach 12 als auch nach 24 Wochen schloss das 95 %-Konfidenzintervall die Null ein, d.h. bei der Wirkung auf den CFB Trough-FEV₁ ergaben sich keine statistisch signifikanten Unterschiede. Daraus lässt sich schlussfolgern, dass der Nutzen von Aclidiniumbromid und Tiotropiumbromid auf den Endpunkt CFB Trough-FEV₁ gleichwertig ist. Der gleichwertige Nutzen zwischen beiden Therapien bestand sowohl bei Analyse der Patienten ohne Kombinationstherapie mit LABAs als auch bei Analyse aller Studien.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.2 CFB Peak-FEV₁ – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

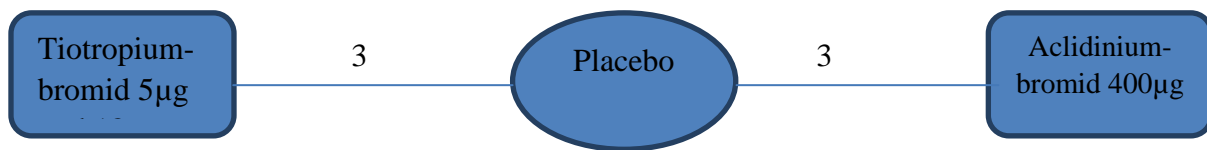
Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-64: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Peak-FEV₁ - 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
2	Beeh 2006			•	•
	Verkindre 2006			•	•
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sechs Studien berichteten über den Peak-FEV₁ nach 12 Wochen (Begleitmedikation mit LABAs nicht gestattet); alle drei Acclidiniumbromid-Studien, zwei Tiotropium-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 17 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 17: Netzwerk-Diagramm für die Analyse des CFB Peak-FEV₁ nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)Tabelle 4.3-65: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Peak-FEV₁ - 12 Wochen - alle Studien]

Anzahl Studien	Referenzstudien	Acclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
3	Beeh 2006			•	•
	Covelli 2005			•	•
	Verkindre 2006			•	•
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sieben Studien berichteten über den Peak-FEV₁ nach 12 Wochen; alle drei Acclidiniumbromid-Studien, drei Tiotropium-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 18 veranschaulicht das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

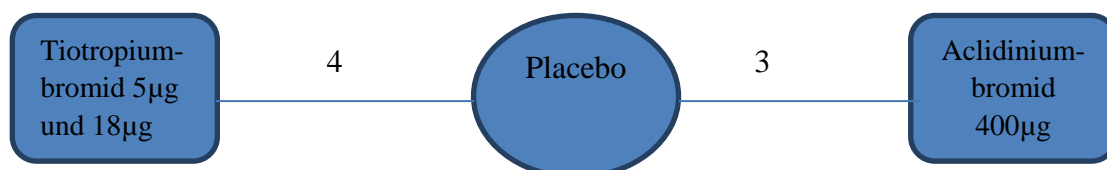
Abbildung 18: Netzwerkdiagramm für die CFB Peak-FEV₁-Analyse nach 12 Wochen

Tabelle 4.3-66: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Peak-FEV₁ - 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
2	INHANCE			•	•
	Donohue 2002			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Drei Studien berichteten über den FEV₁-Peak nach 24 Wochen (Begleitmedikation mit LABAs nicht gestattet); eine Acridiniumbromid-Studie (ATTAIN) und zwei Tiotropium-18µg-Studien (Donohue 2002 und INHANCE). Keine Tiotropiumbromid-5µ-Studie berichtete über diesen Endpunkt. Abbildung 19 veranschaulicht das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 19: Netzwerk-Diagramm für die Analyse des CFB Peak-FEV₁ nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

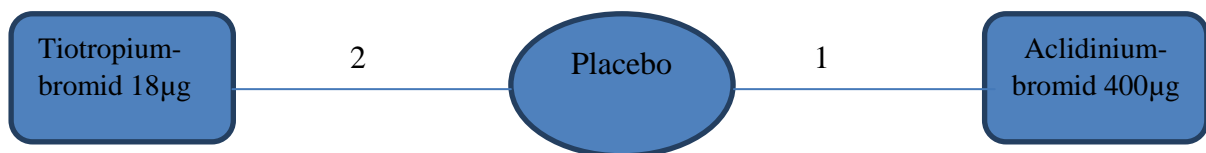


Tabelle 4.3-67: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Peak-FEV₁ - 24 Wochen - alle Studien]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
2	INHANCE			•	•
	Donohue 2002			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Drei Studien berichteten über den Peak-FEV₁ nach 24 Wochen; eine Acclidiniumbromid-Studie (ATTAIN) und zwei Tiotropiumbromid-18µg-Studien (Donohue, 2002 & 2010). Keine Tiotropiumbromid-5µg-Studie berichtete über diesen Endpunkt. Abbildung 20 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 20: Netzwerkdiagramm für die CFB Peak-FEV₁-Analyse nach 24 Wochen



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-68: Operationalisierung des CFB Peak-FEV₁

Studie	Operationalisierung
ACCORD I	Der FEV ₁ -Peak war der maximale FEV ₁ -Wert der FEV ₁ -Bestimmungen vom Zeitpunkt 0 bis 3 Stunden nach der morgendlichen Gabe des Prüfmedikaments bei jeder Beobachtungsuntersuchung (einschließlich Woche 12 als Zweitwert).
ACCORD II	Der FEV ₁ -Peak war der maximale FEV ₁ -Wert der FEV ₁ -Bestimmungen vom Zeitpunkt 0 bis 3 Stunden nach der morgendlichen Dosierung des Prüfmedikaments bei jeder Beobachtungsuntersuchung.
ATTAIN	Die maximalen FEV ₁ -Werte aus den FEV ₁ Betrachtungen vom Zeitpunkt 0 bis 3 Stunden nach der morgendlichen Gabe bei jeder Beobachtungsuntersuchung. Mittels einer Kovarianzanalyse (ANCOVA)-Modell wurden die Variablen Behandlungsgruppe und Geschlecht als Faktoren und Ausgangswert und Alter als Kovariaten analysiert.
Covelli 2005	Die FEV ₁ Peak-CFB-Werte wurden mit Hilfe eines ANCOVA-Modells mit den Faktoren Behandlungsgruppe und Zentren und der Kovariate Baseline-FEV ₁ -Werte analysiert.
INHANCE (Donohue 2010)	FEV ₁ 5 Minuten nach der Einnahme. Die Daten werden als kleinste Quadrat-Mittelwerte (least square (LS) means) mit einem Standardfehler (SE) oder 95%-Konfidenzintervall (CI) angegeben.
Beeh 2006	Der FEV ₁ -Peak wurde 2 Stunden nach der Dosierung gemessen. Die Spirometrie wurde nach den Richtlinien der American Thoracic Society durchgeführt, die Wash-out-Behandlungs-Perioden vor der Messung gestatten.
Donohue 2002	Die Spirometrie für Peak-FEV ₁ wurde 30 min, 60vmin und 2, 3, 4, 6, 8, 10 und 12 Stunden nach der Gabe des Prüfmedikaments unter Verwendung eines Ko-Ko Spirometers (Pulmonary Data Services, Boulder, CO) durchgeführt. Die Spirometrie wurde mit den gleichen Zeitintervallen nach 2, 8, 16 und 24 Therapiewochen wiederholt. Die Messungen wurden nach den Kriterien der American Thoracic Society durchgeführt. Die höchsten FEV ₁ -Werte und die FVC-Messwerte wurden beibehalten. Kurzwirksame Theophylline wurden mindestens 24 Stunden ausgesetzt, langwirksame Theophylline mindestens 48 Stunden und kurz wirkende B2-Agonisten 8 Stunden vor der Spirometrie. Die Patienten erhielten ihre morgendliche Gabe des Studienmedikaments bei Kliniktagen nach der ersten Vordosis-Spirometrie. Die Kovarianzanalyse wurde mit dem Ausgangswert als Kovariate durchgeführt.
Verkindre 2006	Die Peak-Response wurde als Postdosis-Wert (2 Stunden nach der Medikamenten-Verabreichung) bei jeder Klinikuntersuchung nach Randomisierung definiert. Alle spirometrischen Tests wurden dreifach durchgeführt, und die höchsten Messungen wurden für die nachfolgenden Analysen verwendet. Um an den Testtagen standardisierte Bedingungen zu gewährleisten wurden Salbutamol und kurz- und langwirkende Theophylline mindestens 12, 24 bzw. 48 Stunden vor dem Test ausgesetzt. Für alle Analysen wurde eine Intent-to-treat-Methode mit einer Fortschreibung anhand der letzten durchgeführten Beobachtung verwendet.
Voshaar 2008	Spirometrische Messungen wurden in Übereinstimmung mit den Kriterien der American Thoracic Society bis zu 6 Stunden nach der Gabe des Prüfmedikamentes durchgeführt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-69: Bewertung des Verzerrungspotenzials für CFB Peak-FEV₁ in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Covelli 2005	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
INHANCE	Niedrig	Nein	Nein	Nein	Nein	Hoch
Beeh 2006	Niedrig	Ja	Unklar	Unklar	Nein	Niedrig
Donohue 2002	Niedrig	Ja	Nein	Nein	Nein	Hoch
Verkindre 2006	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Voshaar 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial bezüglich des CFB Peak-FEV₁ wurde für die ACCORD II-Studie als hoch betrachtet (Tabelle 4.3-69). Die Behandlungsgruppen waren zum Studienbeginn nicht vollständig vergleichbar, was die Ergebnisse der CFB Peak-FEV₁-Veränderung gegenüber dem Ausgangswert beeinträchtigt haben könnte. Alle Tiotropiumbromid-Studien hatten ein geringes Verzerrungspotenzial auf der Studienebene, aber 2 von 6 Tiotropiumbromid-Studien wiesen ein hohes Verzerrungspotenzial hinsichtlich des Endpunktes CFB Peak-FEV₁ auf (Donohue 2002 und INHANCE). Das ist der Fall, weil das ITT-Prinzip nicht für die Endpunkt-Analyse eingesetzt wurde und die CFB Peak-FEV₁-Messung für die Behandlungsallokation nicht verblindet war (siehe Anhang 4-G).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-70: Ergebnisse für CFB Peak-FEV₁ (ml) aus RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mittelwert (SD)	Mittelwert (SD)	CFB (SD)	Gruppendifferenz (SE) [KI] p-Wert
ACCORD I	Aclidiniumbromid	190	1332 (493)	1593 (551)	262 (248)	194 (21,8) [144, 244] p<0,001
	Placebo	185	1376 (570)	1463 (626)	68 (231)	
ACCORD II	Aclidiniumbromid	177	1249 (519)	1466 (557)	216,8 (246)	130 (25,7) [77; 183] p<0,001
	Placebo	182	1459 (518)	1546,4 (553)	87,2 (273)	
ATTAIN	Aclidiniumbromid	269	1508 (525)	1735 (262)	227 (262)	190 (21) [146; 234] p<0,0001
	Placebo	273	1500 (489)	1544 (264)	37 (264)	
Verkindre 2006	Tiotropiumbromid 18µg	46	1050 (350)	1290 (203)	(-)	210 (40) [132; 288] p<0,001
	Placebo	54	1080 (310)	1080 (220)	(-)	
Covelli 2005	Tiotropiumbromid 18µg	94	1060 (370)	(-)	275 (28)	266 (40) [187; 343] p<0,0001
	Placebo	84	990 (420)	(-)	10 (29)	
Voshaar 2008	Tiotropiumbromid 5µg	180	(-)	(-)	(-)	193 (28) [138; 248] p<0,0001
	Placebo	181	(-)	(-)	(-)	
Beeh 2006	Tiotropiumbromid 18µg	1236	(-)	(-)	(-)	128 (19) [91; 165] p<0,001
	Placebo	403	(-)	(-)	(-)	

N=Anzahl der Patienten, CFB=Veränderung gegenüber dem Ausgangswert, SE=Standardfehler SD=Standardabweichung, ITT=Intention-to-treat-Population, KI=Konfidenzintervall

Tabelle 4.3-70 zeigt den CFB Peak-FEV₁ nach 12 Wochen für Aclidiniumbromid- und Tiotropiumbromid-5µg- sowie -18µg-Studien, die diesen Endpunkt berichten. In allen Studien zeigten mit Aclidiniumbromid oder mit Tiotropiumbromid behandelte Patienten eine Verbesserung des CFB Peak-FEV₁.

Tabelle 4.3-71: Ergebnisse für CFB Peak-FEV₁ (ml) aus RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten	Baseline	Follow-up	Veränderung	
		(ITT)				
		N	Mittelwert (SD)	Mittelwert (SD)	CFB (SD)	Gruppendifferenz (SE) [KI] p-Wert
ATTAIN	Acidiniumbromid	269	1508 (525)	1739 (279)	231 (279)	209 (24) [187; 211] p<,0001
	Placebo	273	1500 (489)	1529 (280)	22 (280)	
Donohue 2002	Tiotropiumbromid 18µg	209	(-)	(-)	(-)	244 (24) [197,0; 291,0] p<,0,0001
	Placebo	201	(-)	(-)	(-)	
INHANCE (Donohue 2010)	Tiotropiumbromid 18µg	415	1450 (505)	(-)	250 (-)	180 (38) [105,5;254,5] p<,0,001
	Placebo	418	1510 (490)	(-)	70 (-)	

N=Anzahl der Patienten, CFB=Veränderung gegenüber dem Ausgangswert, SE=Standardfehler SD=Standardabweichung, ITT=Intention-to-treat-Population, KI=Konfidenzintervall

Tabelle 4.3-71 zeigt den CFB Peak-FEV₁ nach 24 Wochen für alle Acidiniumbromid- und Tiotropiumbromid-18µg-Studien, die diesen Endpunkt berichten. Keine Tiotropiumbromid-5µg-Studien berichteten diesen Endpunkt. In allen Studien zeigten die mit Acidiniumbromid und mit Tiotropiumbromid behandelten Patienten eine Verbesserung des CFB Peak-FEV₁. Nur eine Acidiniumbromid-Studie berichtete den CFB Peak-FEV₁ nach 24 Wochen, daher war für diesen Zeitpunkt keine Meta-Analyse möglich.

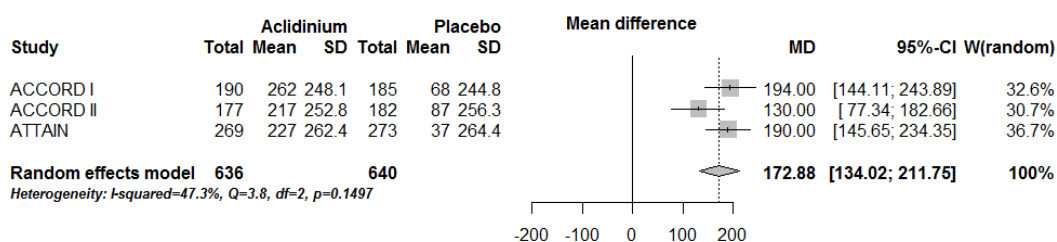
Abbildung 21: Meta-Analyse für CFB Peak-FEV₁ nach 12 Wochen (ml) aus RCT; Acidiniumbromid im Vergleich zu Placebo

Abbildung 21 zeigt, dass Acidiniumbromid wirksamer als Placebo ist (172,9 [95%-KI 134,0; 211,8]). Der Heterogenitätstest (Q=3,8; p=0,15; I²=47,3%) weist auf eine signifikante Studienheterogenität hin. Die Effektrichtung stimmt jedoch in allen drei Studien überein. Eine Sensitivitätsanalyse wird diesen Sachverhalt näher untersuchen.

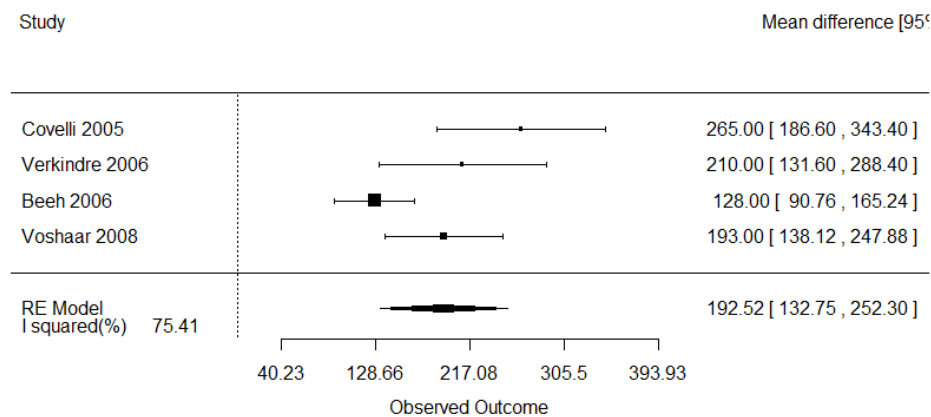


Abbildung 22: Meta-Analyse für CFB Peak-FEV₁ nach 12 Wochen (ml) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo

Abbildung 22 zeigt, dass Tiotropiumbromid nach 12 Wochen wirksamer als Placebo ist (192,5 [95%-KI 132,8; 252,3]). Der Heterogenitätstest ($Q=11,3$; $p=0,007$; $I^2=75,4\%$) lässt auf eine signifikante Heterogenität schließen. Die Effektrichtung stimmt jedoch in allen Studien überein. Eine Sensitivitätsanalyse wird diesen Sachverhalt näher untersuchen.

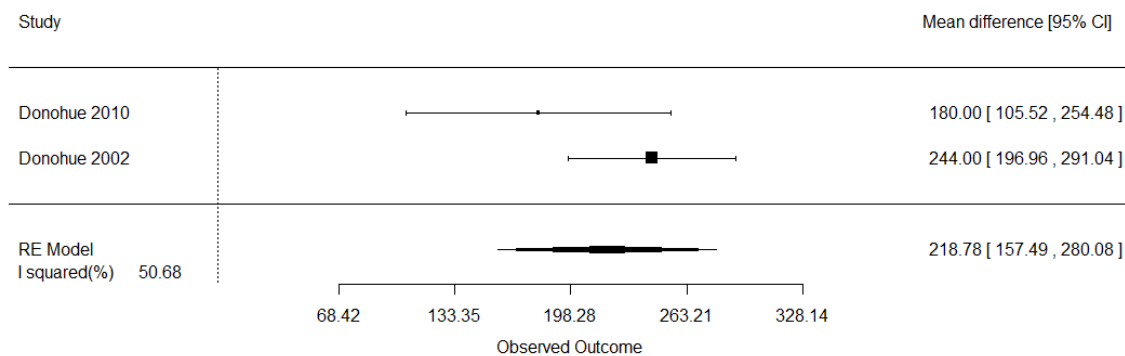


Abbildung 23: Meta-Analyse für CFB Peak-FEV₁ nach 24 Wochen (ml) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo

Abbildung 23 zeigt, dass Tiotropiumbromid wirksamer ist als Placebo (218,8 [95%-KI 157,5; 280,1]). Beide Studien wiesen jedoch ein hohes Verzerrungspotenzial auf und der Heterogenitätstest spricht für eine signifikante Heterogenität ($Q=2,0$; $p < 0,15$; $I^2=50,7\%$). Die Effektrichtung stimmt jedoch in allen Studien überein. Eine Sensitivitätsanalyse kann nicht durchgeführt werden, da nur zwei Tiotropiumbromid-Studien diesen Endpunkt berichteten.

Sensitivitätsanalyse

Die ACCORD-II-Studie wies hinsichtlich der Baseline-Patientencharakteristika ein Ungleichgewicht zwischen den Behandlungsgruppen auf. Um die Auswirkungen des Ausschlusses dieser Studie aus der Meta-Analyse zu erforschen, wurde eine Sensitivitätsanalyse durchgeführt (Abbildung 24).

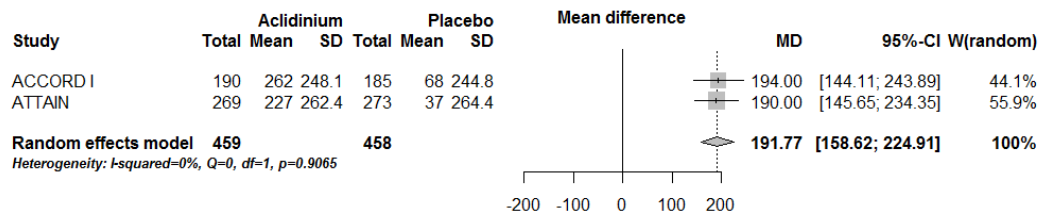


Abbildung 24: Sensitivitätsanalyse (Verzerrungspotential) für CFB Peak-FEV₁ (ml) nach 12 Wochen aus RCT; Acclidiniumbromid im Vergleich zu Placebo.

Abbildung 24 zeigt, dass Acclidiniumbromid eine höhere Wirksamkeit hinsichtlich der Verbesserung des nach 12 Wochen gemessenen CFB Peak-FEV₁ aufwies als Placebo (mittlere Differenz 191,8 [95%-KI 158,6; 224,9]). Wie aus den Konfidenzintervallen der Meta-Analyse ersichtlich ist, hatte der Ausschluss der ACCORD-II-Studie keine signifikanten Auswirkungen auf den gepoolten Effektschätzer. Der Heterogenitätstest (Q=0,0; p=0,91; I²=0,0%) ergab nun keine Hinweise mehr auf signifikante Studienheterogenität.

Um die Ursache für die Heterogenität der Tiotropiumbromid-Studien zu untersuchen, wird eine Meta-Analyse unter Ausschluss der Tiotropiumbromid-5µg-Studie durchgeführt. Der Ausschluss von verzerrten Studien ist nicht möglich, da dies den Studienpool auf unter zwei Studien schrumpfen lassen würde. Die Ergebnisse des Ausschlusses von Vosharr 2008 sind in Abbildung 25 dargestellt.

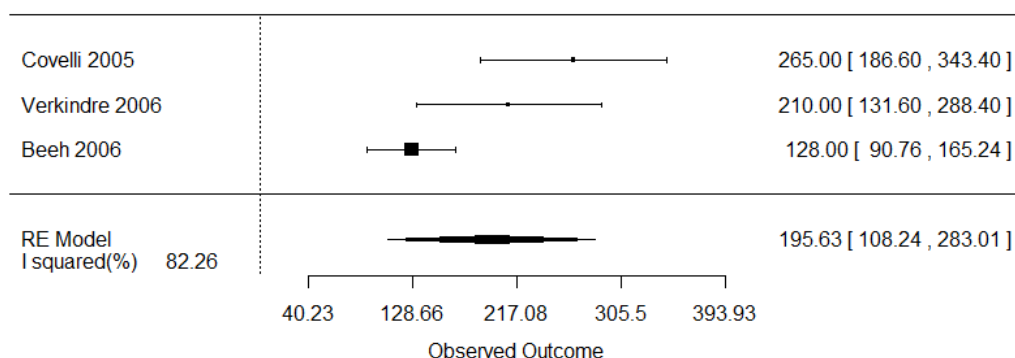


Abbildung 25: Sensitivitätsanalyse (Dosierung) für CFB Peak-FEV₁ (ml) nach 12 Wochen aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.

Abbildung 25 zeigt, dass Tiotropiumbromid eine höhere Wirksamkeit hinsichtlich der Verbesserung des nach 12 Wochen gemessenen CFB Peak-FEV₁ aufwies als Placebo (mittlere Differenz 195,6 [95%-KI 108,2; 283,0]). Der Ausschluss der Tiotropiumbromid-5µg-Studie (Voshaar 2008) wirkte sich nicht signifikant auf den gepoolten Effektschätzer aus, wie die Konfidenzintervalle der Meta-Analyse zeigen. Der Heterogenitätstest ($Q=11,3$; $p=0,004$; $I^2=82,3\%$) lässt weiterhin auf eine signifikante Studienheterogenität schließen. Die unterschiedliche Dosierung von Tiotropiumbromid stellt hier keinen beeinflussenden Faktor dar. Der Gesamteffektschätzer der Meta-Analyse wird trotz vorliegender Heterogenität berücksichtigt, da Tiotropiumbromid in allen Studien den Peak.FEV₁ signifikant verbessert (gleiche Effektrichtung).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-72 zeigt die DIC-Schätzwerte für die geplanten NMAs. Nach 12 Wochen war das beste durch DIC-Ergebnisse ermittelte Modell ein Modell mit zufälligen Effekten sowohl in Szenario 1 als auch in Szenario 2 (Tabelle 4.3-72). Die Verwendung eines Modells mit zufälligen Effekten ist konsistent mit den Befunden der Meta-Analyse von Tiotropiumbromid 18µg (Abbildung 22), wobei der Heterogenitätstest eine signifikante Heterogenität ergab.

Für die Analyse der 24 Wochen-Daten wurde wegen des niedrigeren DIC-Wertes ein Random-Effekt-Modell gewählt (Tabelle 4.3-72). Das ist mit der Tiotropiumbromid 18µg-Metaanalyse, die eine signifikante Heterogenität zeigte, konsistent ($p < 0,15$; Abbildung 23). Das Szenario 2 wurde nach 24 Wochen nicht durchgeführt, da die Patienten aller Studien, die Daten zu FEV₁-Peak nach 24 Wochen berichteten, keine Begleitmedikation mit LABA erhalten hatten.

Tabelle 4.3-72: DIC für CFB (ml) Peak-FEV₁ - Indirekter Vergleich aus RCT

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung zum Modell
Woche 12			
Ohne LABA-Begleitmedikation	73,151	63.352	Zufällige Effekte
Alle Studien	88,830	63.352	Zufällige Effekte
Woche 24			
Ohne LABA-Begleitmedikation	41,0	31.452	Zufällige Effekte
Alle Studien	-	-	-

Die Ergebnisse der NMA werden in Tabelle 4.3-73, Tabelle 4.3-74 und Tabelle 4.3-75 dargestellt.

Tabelle 4.3-73: CFB (ml) Peak-FEV₁ (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit zufälligen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	156.40	
	95% KI	128.10	184.80
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	175.60	
	95% KI	151.30	199.60
	P(Überlegenheit)	>99%	
			19.23
		-17.88	57.38
		84%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-74: CFB (ml) Peak-FEV₁ (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: alle Studien; Modell mit zufälligen Effekten

Intervention		Kontrolle			
		Pbo		TIO 5 und 18	
TIO 5 und 18	Effektschätzer	169.00			
	95% KI	141.70	195.30		
	P(Überlegenheit)	>99%			
AB 400	Effektschätzer	175.70		6.74	
	95% KI	149.90	201.60	-31.64	45.58
	P(Überlegenheit)	>99%		63%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 12 Wochen war Acclidiniumbromid wirksamer als Placebo [Unterschied 176ml (95% KI: 151, 200)] und genauso wirksam wie Tiotropiumbromid [Unterschied 19ml (95% KI: -18, 57)] (Tabelle 4.3-73). Ähnliche Ergebnisse wurden mit dem Szenario 2 (Tabelle 4.3-74) erzielt. Die aus der NMA für den Vergleich von Acclidiniumbromid, Tiotropiumbromid 5µg und Tiotropiumbromid 18µg mit Placebo erhaltenen Ergebnisse stimmten mit den Ergebnissen der Meta-Analysen überein. Acclidiniumbromid und Tiotropiumbromid ergaben nach 12 Wochen eine Verbesserung des Peak-FEV₁ im Vergleich zu Placebo.

Tabelle 4.3-75: CFB (ml) Peak-FEV₁ (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit zufälligen Effekten

Intervention		Kontrolle			
		Pbo		TIO 18	
TIO 18	Effektschätzer	224.20			
	95% KI	185.60	264.00		
	P(Überlegenheit)	>99%			
AB 400	Effektschätzer	206.00		-18.17	
	95% KI	158.40	252.80	-80.52	42.23
	P(Überlegenheit)	>99%		29%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen war Acclidiniumbromid effektiver als Placebo [Unterschied 206ml (95% KI: 158, 253)] und genauso wirksam wie Tiotropiumbromid [Unterschied -18ml (95% KI: -81, 42)] (Tabelle 4.3-75). In beiden Analysen (NMA und Metaanalyse) verbesserte Acclidiniumbromid den Peak-FEV₁ nach 24 Wochen im Vergleich zu Placebo signifikant.

Beide Therapien, Acclidiniumbromid und Tiotropiumbromid, verbesserten den Endpunkt CFB Peak-FEV₁ nach 12 und 24 Wochen statistisch signifikant stärker als Placebo. Der indirekte Vergleich ergab eine gleichwertige Wirkung von Acclidiniumbromid und Tiotropiumbromid auf die CFB Peak-FEV₁. Sowohl nach 12 als auch nach 24 Wochen schloss das 95 %-Konfidenzintervall die Null ein, d.h. bei der Wirkung auf den CFB Peak-FEV₁ ergaben sich keine statistisch signifikanten Unterschiede. Daraus lässt sich schlussfolgern, dass der Nutzen von Acclidiniumbromid und Tiotropiumbromid bezüglich des Endpunktes CFB Peak-FEV₁ gleichwertig ist.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.3 TDI Focal Score – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-76: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI Focal Score – 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
3	Casaburi 2002			•	•
	INHANCE			•	•
	Verkindre 2006			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sechs Studien berichteten über die Veränderung des TDI Focal Score (begleitende Behandlung mit LABAs nicht gestattet) nach 12 Wochen; alle drei Aclidiniumbromid-Studien und drei Tiotropiumbromid-18µg-Studien. In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 26 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 26: Netzwerkdiagramm für die Analyse des TDI Focal Score nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

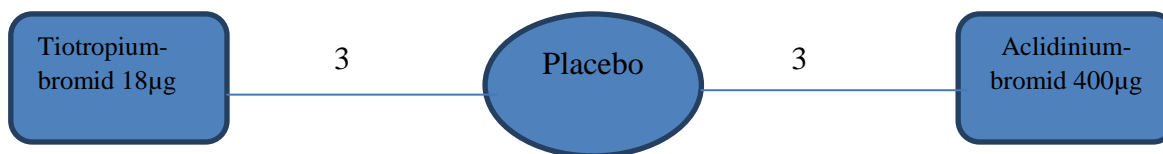


Tabelle 4.3-77: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI Focal Score – 12 Wochen - alle Studien]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
3	Casaburi 2002			•	•
	INHANCE			•	•
	Verkindre 2006			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sechs Studien berichteten die Veränderung des TDI Focal Score nach 12 Wochen; alle drei Acridiniumbromid-Studien und drei Tiotropiumbromid-18µg-Studien. In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 27 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 27: Netzwerkdiagramm für die Analyse des TDI Focal Score nach 12 Wochen

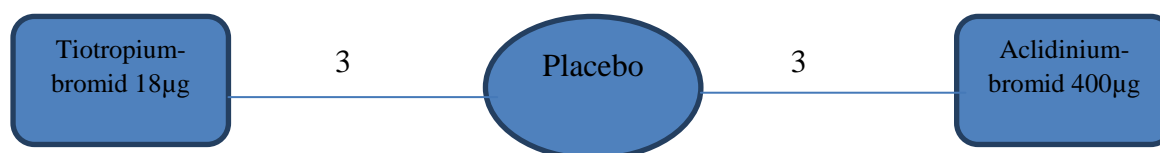


Tabelle 4.3-78: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI Focal Score – 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
4	Brusasco 2003			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	Donohue 2002			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Fünf Studien berichteten über die Veränderung des TDI Focal Score (begleitende Behandlung mit LABAs nicht gestattet) nach 24 Wochen; eine Acridiniumbromid-Studie (ATTAIN) und vier Tiotropiumbromid-18µg-Studien (Brusasco 2003, Casaburi 2002, Donohue 2002, INHANCE). In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 28 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 28: Netzwerkdiagramm für die Analyse des TDI Focal Score nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)



Tabelle 4.3-79: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI Focal Score – 24 Wochen - alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
4	Brusasco 2003			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	Donohue 2002			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Fünf Studien berichteten über die Veränderung des TDI Focal Score nach 24 Wochen; eine Aclidiniumbromid-Studie (ATTAIN) und vier Tiotropiumbromid-18µg-Studien (Brusasco 2003, Casaburi 2002, Donohue 2002, INHANCE). In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 29 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 29: Netzwerkdiagramm für die Analyse des TDI Focal Score nach 24 Wochen



Beschreiben Sie die Operationalisierung von TDI Focal Score für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-80: Operationalisierung des TDI Focal Score

Studie	Operationalisierung
ACCORD I	Der BDI und der TDI messen die Schwere der Atemnot bei symptomatischen Patienten. Mit dem BDI wurde die Schwere der Dyspnoe zu Beginn der Studie gemessen und mit dem TDI wurden Änderungen dieser Ausgangs-Messung bewertet. Der TDI wurde vor der Gabe des Medikaments gemessen. Beide Indizes enthalten eine Bewertung für 3 Kategorien, die wichtige Faktoren für die Entwicklung der Dyspnoe darstellen. Das interessierende Outcome war die Veränderung des Dyspnoestatus der Baselineerhebung bis zu Woche 12 durch den gemessenen TDI (focal score und 3-Dimensionen-Score).
ACCORD II	Der BDI / TDI stellen ein geprüfetes Instrument dar, das die Schwere der Atemnot bei symptomatischen Patienten misst. Der BDI wurde während der Visite 2 (vor der Verabreichung des Prüfmedikamentes) beurteilt; Der TDI wurde während Visite 6 (12 Wochen) beurteilt. Das Outcome war die Veränderung des Dyspnoestatus der Baseline bis zu Woche 12 durch den gemessenen TDI (focal score und 3-Dimensionen-Score). Das interessierende Outcome war die Veränderung des Dyspnoestatus der Baseline bis zu Woche 12 durch den gemessenen TDI (focal score und 3-Dimensionen-Score).
ATTAIN	Das Ziel des BDI und TDI war es, den Schweregrad der Atemnot bei symptomatischen Patienten zu messen. Mit dem BDI wurde die Schwere der Dyspnoe zu Beginn der Studie gemessen und mit dem TDI wurden Änderungen dieser Baseline (Ausgangs-Messung) bewertet. Für diese Studie wurde der BDI während Visite 2 und der TDI bei Visite 6 (Woche 12) und Visite 8 (Woche 24) angewandt. Das interessierende Outcome war die Veränderung des Dyspnoestatus der Baseline bis zu Woche 12 durch den gemessenen TDI (focal score und 3-Dimensionen-Score).
Brusasco 2003	Die Dyspnoe wurde anhand des Baseline-Dyspnoe-Index (BDI) und des Transition-Dyspnoe-Index (TDI) bestimmt. BDI und TDI bestehen aus drei Achsen (funktionale Beeinträchtigung, Ausmaß der Aufgabe, und Ausmaß des Einflusses), die zur Schaffung eines focal scores summiert werden. Der BDI wurde am Ende der Baseline-Untersuchung angewandt und der TDI wurde bei den Folgeuntersuchungen in der 8., 16. und 24. Woche angewandt. Fehlende TDI-Fragen wurden unter Verwendung der Fortschreibung anhand der letzten durchgeführten Beobachtung berechnet, es sei denn, dass der Patient aufgrund einer Verschlechterung der COPD abbrach, in diesem Fall wurde die ungünstigste Beobachtung übertragen.
Casaburi 2002	Die Dyspnoe zu Beginn wurde mit dem Basis-Dyspnoe-Index (BDI) bewertet. Dieses Instrument hat drei Bereiche (funktionelle Beeinträchtigung, Ausmaß der Aufgabe und Ausmaß der Anstrengung), die Summe der Werte ergibt einen kombinierten focal score. Änderungen der Beeinträchtigung gegenüber dem Ausgangswert (Baseline) wurden mit dem Transition-Dyspnoe-Index (TDI) an den Tagen 50 und 90 und danach alle 3 Monate gemessen.
INHANCE	Die Dyspnoe wurde mit dem TDI beurteilt. Der TDI wurde in den Wochen 4, 8, 12 und 26 gemessen. Fehlende Werte für den TDI-Gesamtwert wurden durch Werte aus der Fortschreibung der letzten durchgeführten Beobachtung ersetzt, aber nicht länger als 14 Wochen. Ein Unterschied von 1 Punkt im TDI-Wert wurde als klinisch relevant angesehen.
Donohue 2002	Die Dyspnoe wurde anhand des Baseline-Dyspnoe-Index (BDI) und des Transition-Dyspnoe-Index (TDI) bewertet. Der BDI und der TDI bestehen aus drei Achsen (Funktionsbeeinträchtigung, Ausmaß der Aufgabe, und Ausmaß der Anstrengung), die zur Schaffung eines focal scores summiert werden. Der BDI wurde am Ende des Baseline-Zeitraumes angewandt und der TDI wurde bei den oben angegebenen Folgeuntersuchungen angewandt. Ein höherer TDI-Wert bedeutet eine Verbesserung. Eine Änderung von 1 Einheit wird als klinisch signifikant angesehen.
Verkindre 2006	Die Dyspnoe bei Aktivitäten des täglichen Lebens wurde am Tag 1 mit dem Baseline-Dyspnoe-Index (BDI) bewertet. Änderungen gegenüber dem Ausgangswert (der Baseline) wurden mit dem Transition-Dyspnoe-Index (TDI) an den Tagen 42 und 84 gemessen. Der BDI und der TDI

setzen sich aus drei Bereichen (funktionelle Beeinträchtigung, Ausmaß der Aufgabe und Ausmaß der Anstrengung) zusammen. Für den BDI reicht jede Komponenten-Skala von 0 (schwere Dyspnoe) bis 4 (keine Dyspnoe) und für den TDI von -3 (starke Verschlechterung) bis +3 (wesentliche Verbesserung). Die Bewertungen für jede der drei Kategorien von BDI und TDI wurden zur Bildung der Ausgangswerte (Baseline 0-12) und der Transitional-focal-scores (-9 bis +9) hinzugefügt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-81: Bewertung des Verzerrungspotenzials für TDI Focal Score in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Brusasco 2003	Niedrig	Ja	Unklar	Unklar	Nein	Niedrig
Casaburi 2002	Niedrig	Ja	Ja	Nein	Nein	Niedrig
INHANCE	Niedrig	Nein	Nein	Nein	Nein	Hoch
Donohue 2002	Niedrig	Ja	Nein	Nein	Nein	Hoch
Verkindre 2006	Niedrig	Ja	Nein	Nein	Nein	Hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

In Bezug auf den TDI Focal Score Score wurde das Verzerrungspotenzial der ACCORD II-Studie aufgrund der nicht vollständig gegebenen Vergleichbarkeit der Behandlungsgruppen zur Baseline als hoch eingestuft. Alle Tiotropiumbromid-Studien wiesen ein geringes Verzerrungspotenzial auf Studienebene auf. Das Verzerrungspotenzial wurde bei drei von fünf Tiotropiumbromid-Studien jedoch besonders in Bezug auf den TDI Focal Score als hoch eingestuft. Dies ist darauf zurückzuführen, dass das ITT-Prinzip nicht für die Endpunkt-Analyse angewendet wurde (detaillierte Erklärung im Anhang 4-G).

Stellen Sie die Ergebnisse für TDI Focal Score für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-82: Ergebnisse für TDI Focal Score aus RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Baseline	Veränderung	Gruppendifferenz (SE) [KI] p-Wert
		N	BDI, Mittelwert (SD)	TDI, Mittelwert (SD)	
ACCORD I	Aclidiniumbromid	190	6,2 (2,1)	1,5 (2,8)	1,00 (0,28) [0,45; 1,55] p = 0,002
	Placebo	185	6,5 (2,2)	0,5 (2,7)	
ACCORD II	Aclidiniumbromid	177	6,0 (1,9)	1,3 (2,7)	1,00 (0,28) [0,45; 1,55] p = 0,005
	Placebo	182	6,2 (2,2)	0,3 (2,7)	
ATTAIN	Aclidiniumbromid	269	6,7 (2,1)	1,7 (3,1)	0,88 (0,28) [0,35; 1,41] p = 0,001
	Placebo	273	6,7 (2,0)	0,9 (3,3)	
Casaburi 2002	Tiotropiumbromid 18µg	507	(-)	1,3 (2,9)	0,95 (0,18) [0,59; 1,31] p<0,001
	Placebo	332	(-)	0,3 (2,4)	
INHANCE (Donohue 2010)	Tiotropiumbromid 18µg	415	(-)	1,9 (2,2)	0,72 (0,15) [0,42; 1,02] p<0,001
	Placebo	418	(-)	1,2 (2,3)	
Verkindre 2006	Tiotropiumbromid 18µg	46	5,7 (2,0)	0,6 (5,2)	1,28 (1,07) [-0,82; 3,38] p=0,232
	Placebo	54	6,1 (1,9)	-0,7 (5,6)	

N=Anzahl der Patienten, CFB=Veränderung gegenüber dem Ausgangswert, SE=Standardfehler SD=Standardabweichung, ITT=Intention-to-treat-Population, KI=Konfidenzintervall, TDI= Transition Dyspnea Index. BDI= Baseline Dyspnea Index.

Tabelle 4.3-82 zeigt die TDI Focal Score nach 12 Wochen für die Aclidiniumbromid- und Tiotropiumbromid-Studien, welche diese Ergebnisse berichten. In allen Studien zeigten die mit Aclidiniumbromid oder Tiotropiumbromid behandelten Patienten einen verbesserten TDI-Focal Score gegenüber der Baseline (BDI) und Placebo.

Tabelle 4.3-83: Ergebnisse für den TDI Focal Score aus RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten	Baseline	Veränderung	Gruppendifferenz (SE) [KI] p-Wert
		(ITT)	BDI, Mittelwert (SD)	TDI, Mittelwert (SD)	
		N			
ATTAIN	Aclidiniumbromid	269	6,7 (2,1)	1,94 (3,4)	1,00 (0,29) [0,43; 1,57] p = 0,001
	Placebo	273	6,7 (2,0)	0,94 (3,5)	
Brusasco 2003	Tiotropiumbromid 18µg	386	(-)	(-)	1,10 (0,30) [0,51; 1,69] p<0,001
	Placebo	362	(-)	(-)	
Casaburi 2002	Tiotropiumbromid 18µg	507	(-)	0,95	0,84 (0,17) [0,50; 1,18] p<0,001
	Placebo	332	(-)	0,11	
Donohue 2002	Tiotropiumbromid 18µg	209	(-)	0,38	1,02 (0,40) [0,23; 1,81] p=0,011
	Placebo	201	(-)	-0,63	
Donohue 2010	Tiotropiumbromid 18µg	415	(-)	2,23	0,87 (0,23) [0,42; 1,32] p<0,001
	Placebo	418	(-)	1,38	

N= Anzahl der Patienten SE= Standardfehler (engl.: standard error) SD=Standardabweichung (eng.: standard deviation) (-) = nicht angegeben ITT= Intention-to-treat-Population [KI]= Konfidenzintervall TDI= Transition Dyspnea Index. BDI= Baseline Dyspnea Index.

Tabelle 4.3-83 zeigt den TDI Focal Score nach 24 Wochen für alle Aclidiniumbromid- und Tiotropiumbromid-5µg- sowie -18µg-Studien, die dieses Outcome berichten. In allen Studien zeigten mit Aclidiniumbromid behandelte Patienten verbesserte TDI-Scores gegenüber der Baseline (BDI) und Placebo. Nur eine Aclidiniumbromid-Studie berichtete dieses Outcome, daher war eine Meta-Analyse nicht möglich.

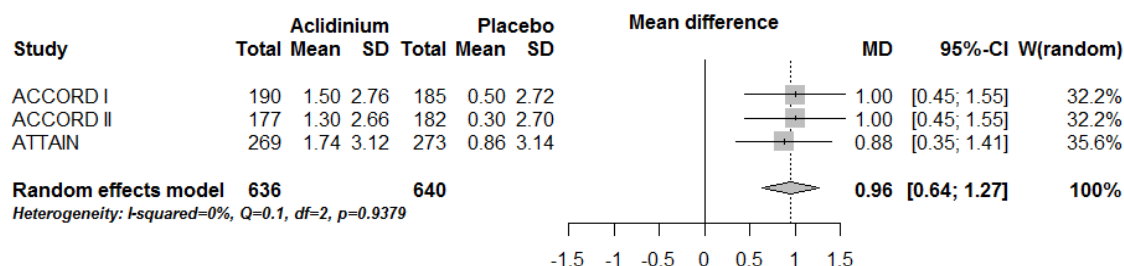


Abbildung 30: Meta-Analyse für TDI Focal Score (12 Wochen) aus RCT; Aclidiniumbromid 400µg im Vergleich zu Placebo

Abbildung 30 zeigt, dass Acclidiniumbromid nach einem Behandlungszeitraum von 12 Wochen eine höhere Wirksamkeit hinsichtlich des TDI Focal Score aufwies als Placebo (gepoolte mittlere Differenz 0,96 [95%-KI 0,64; 1,27]). Es bestand keine Heterogenität zwischen den Studien.

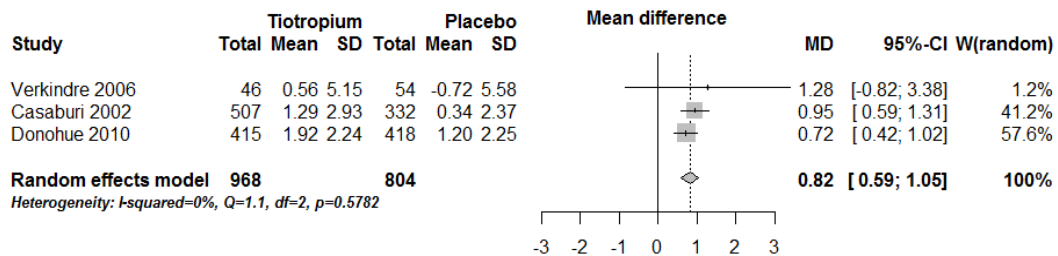


Abbildung 31: Meta-Analyse für TDI Focal Score (12 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.

Abbildung 31 zeigt, dass Tiotropiumbromid nach einem Behandlungszeitraum von 12 Wochen eine höhere Wirksamkeit hinsichtlich des TDI Focal Scores aufwies als Placebo (mittlere Differenz 0,82 [95%-KI 0,59; 1,05]). Die Tests ($Q=1,1$; $p=0,58$; $I^2=0\%$) ergaben keine Hinweise auf Heterogenität zwischen den Studien.

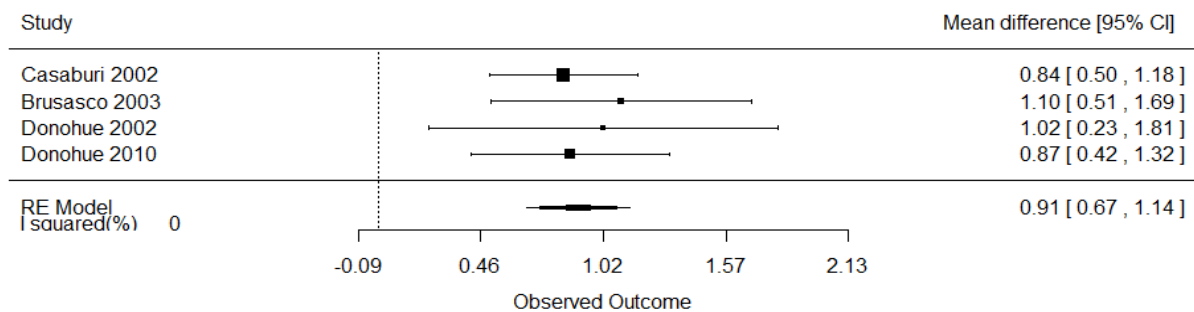


Abbildung 32: Meta-Analyse für TDI Focal Score (24 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.

Abbildung 32 zeigt, dass Tiotropiumbromid nach einem Behandlungszeitraum von 24 Wochen eine höhere Wirksamkeit hinsichtlich des TDI Focal Scores aufwies als Placebo (mittlere Differenz 0,91 [95%-KI 0,67; 1,14]). Die Tests ($Q=0,67$; $p=0,88$; $I^2=0\%$) ergaben keine Hinweise auf Heterogenität zwischen den Studien..

Sensitivitätsanalyse

Die ACCORD-II-Studie wies hinsichtlich der Baseline-Patientencharakteristika ein Ungleichgewicht zwischen den Behandlungsgruppen auf. Um die Auswirkungen des Ausschlusses dieser Studie aus der Meta-Analyse zu ermitteln, wurde eine Sensitivitätsanalyse durchgeführt (Abbildung 33).

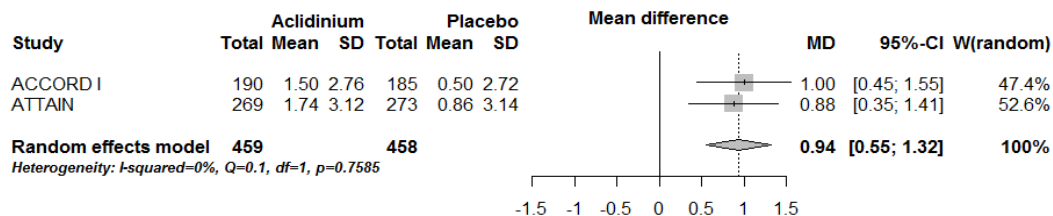


Abbildung 33: Sensitivitätsanalyse (Verzerrungspotential) für den TDI Focal Score nach 12 Wochen (ml) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.

Abbildung 33 zeigt, dass Aclidiniumbromid nach einem Behandlungszeitraum von 12 Wochen eine höhere Wirksamkeit hinsichtlich des TDI Focal Scores aufwies als Placebo (Differenz 0,94 [95%-KI 0,55; 1,32]). Wie aus den Konfidenzintervallen der Meta-Analyse ersichtlich ist, hatte der Ausschluss der ACCORD-II-Studie keine signifikanten Auswirkungen auf den gepoolten Effektschätzer. Der Heterogenitätstest ($Q=0,1$; $p<0,76$; $I^2=0,0\%$) ergab keine Hinweise auf signifikante Studienheterogenität.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-84 zeigt die DIC-Schätzwerte für die geplanten NMAs. Nach 12 Wochen ergab sich aus den DIC-Ergebnissen als bestes Modell ein Modell mit festen Effekten. (Tabelle 4.3-84). Dies stimmte auch mit den Ergebnissen der Meta-Analyse überein; hier ließ der Heterogenitätstest auf keine signifikante Heterogenität schließen, weder für die Meta-Analyse der Aclidiniumbromid-400µg-Studien noch für die Meta-Analyse der Tiotropiumbromid-18µg-Studien. Nach 24 Wochen ergab sich aus den DIC-Ergebnissen als bestes Modell ein Modell mit festen Effekten (Tabelle 4.3-84). Dies stimmte auch mit den Ergebnissen der Meta-Analyse für Tiotropiumbromid 18µg überein; hier ließ der Heterogenitätstest auf keine

signifikante Heterogenität schließen. Szenario 2 wurde nicht durchgeführt, da alle Patienten in den Studien, die Daten zu dem TDI Focal Score berichteten, keine Begleitmedikation mit LABA einsetzten.

Tabelle 4.3-84: DIC des indirekten Therapievergleichs zwischen Acridiniumbromid und Tiotropiumbromid 18µg; Differenz des TDI Focal Scores

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung für das NMA-Modell
Woche 12			
Ohne LABA-Begleitmedikation	2,2	3,8	Feste Effekte
Alle Studien	-	-	-
Woche 24			
Ohne LABA-Begleitmedikation	1,4	3,2	Feste Effekte
Alle Studien	-	-	-

Die Ergebnisse der Netzwerk-Meta-Analyse sind in Tabelle 4.3-85 und Tabelle 4.3-86 aufgeführt.

Tabelle 4.3-85: TDI Focal Score (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle		
		Pbo	TIO 18	
TIO 18	Effektschätzer	0,82		
	95% KI	0,59	1,04	
	P(Überlegenheit)	>99%		
AB 400	Effektschätzer	0,96		
	95% KI	0,64	1,28	
	P(Überlegenheit)	>99%		
			0,14	
			-0,24	0,53
			76%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acridiniumbromid 400 µg;

Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 12 Wochen erzielte Acridiniumbromid im Vergleich zu Placebo eine Verbesserung des TDI Focal Score [Differenz 0,96 Punkt (95%-KI: 0,6, 1,3)]. Acridiniumbromid 400µg war mit Tiotropiumbromid gleichwertig [Differenz 0,1 Punkt (95%-KI: -0,2, 0,5)] (Tabelle 4.3-85). Die Ergebnisse der Netzwerk-Meta-Analyse für den Vergleich zwischen Acridiniumbromid 400µg, Tiotropiumbromid und Placebo stimmten mit den Ergebnissen der Meta-Analysen überein; nach 12 Wochen wurde mit Acridiniumbromid 400µg und Tiotropiumbromid eine Verbesserung des TDI Focal Score im Vergleich zu Placebo beobachtet.

Tabelle 4.3-86: TDI Focal Score (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 18
TIO 18	Effektschätzer	0,90	
	95% KI	0,67	1,14
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	1,00	
	95% KI	0,43	1,57
	P(Überlegenheit)	>99%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg;

Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen zeigte sich Acclidiniumbromid 400µg wirksamer als Placebo und gleichwertig mit Tiotropiumbromid, mit ähnlichen mittleren Differenzwerten wie nach 12 Wochen (Tabelle 4.3-86). Die bestätigt die Ergebnisse der Meta-Analyse, die eine bessere Wirksamkeit von Acclidiniumbromid gegenüber Placebo zeigte.

Acclidiniumbromid und Tiotropiumbromid verbesserten beide den TDI focal Score sowohl nach 12 als auch nach 24 Wochen signifikant im Vergleich zu Placebo. Der indirekte Vergleich beider Therapien ergab einen gleichwertigen Effekt beider Substanzen auf den Endpunkt TDI Focal Score. Sowohl nach 12 als auch nach 24 Wochen schloss das 95 %-Konfidenzintervall die Null mit ein, d.h. es ergaben sich keine statistisch signifikanten Unterschiede auf diesen Endpunkt, der Nutzen beider Therapien ist demnach gleichwertig.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.4 TDI -Responder – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-87: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI-Responder (%) - 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
1	INHANCE			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Vier Studien berichteten über die Anzahl der TDI-Responder nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet; TDI-Response: Verbesserung des TDI-Gesamtwerts ≥ 1 Punkt); alle drei Acridiniumbromid-Studien und eine Tiotropiumbromid-18µg-Studie. In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 34 zeigt das Netzwerk, das basierend auf diesen drei Studien erzeugt werden kann.

Abbildung 34: Netzwerkdiagramm für die Analyse der TDI-Responder nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)



Tabelle 4.3-88: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI-Responder (%) - 12 Wochen - alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
1	INHANCE			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Vier Studien berichteten über die Anzahl der TDI-Responder nach 12 Wochen (Patienten erreichten eine Verbesserung des TDI-Gesamtwerts von ≥ 1 Punkten); alle drei Aclidiniumbromid-Studien und eine Tiotropiumbromid-18µg-Studie. In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 35 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 35: Netzwerkdiagramm für die Analyse der TDI-Responder nach 12 Wochen

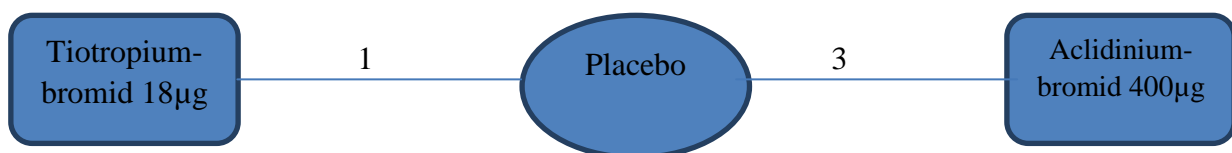


Tabelle 4.3-89: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI-Responder (%) – 24 Wochen – begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
3	Brusasco 2003			•	•
	INHANCE			•	•
	Donohue 2002			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Vier Studien berichteten über die Anzahl der SGRQ-Responder nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet; Verbesserung des TDI-Gesamtwerts ≥ 1 Punkt); eine Acridiniumbromid-Studie (ATTAIN) und drei Tiotropiumbromid-18µg-Studien (Brusasco 2003, Donohue 2002 & 2010). In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 36 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 36: Netzwerkdiagramm für die Analyse der TDI-Responder nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)



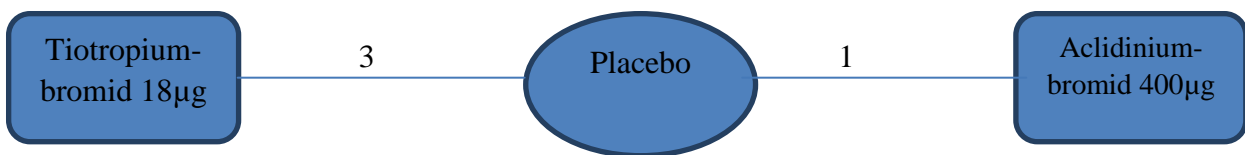
Tabelle 4.3-90: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [TDI-Responder (%) – 24 Wochen – alle Studien]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
3	Brusasco 2003			•	•
	INHANCE			•	•
	Donohue 2002			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Vier Studien berichteten über die Anzahl der TDI-Responder nach 24 Wochen (Patienten erreichten eine Verbesserung des TDI-Gesamtwerts von ≥ 1 Punkten); eine Acridiniumbromid-Studie (ATTAIN) und drei Tiotropiumbromid-18 μ g-Studien (Brusco 2003, Donohue 2002 und INHANCE). In keiner der Tiotropiumbromid-5 μ g-Studien wurde dieses Outcome berichtet. Abbildung 37 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 37: Netzwerkdiagramm für die Analyse der TDI-Responder nach 24 Wochen



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-91: Operationalisierung von Prozentsatz der TDI-Responder

Studie	Operationalisierung
ACCORD I	<p>Der BDI und der TDI messen die Schwere der Atemnot bei symptomatischen Patienten. Mit dem BDI wurde die Schwere der Dyspnoe zu Beginn der Studie gemessen und mit dem TDI wurden Änderungen dieser Baseline-Messung bewertet. Der TDI wurde vor der Gabe des Prüfmedikamentes gemessen. Beide Indizes enthalten eine Bewertung für 3 Kategorien, die wichtige Faktoren für die Entwicklung der Dyspnoe darstellen. Das interessierende Outcome war die Veränderung des Dyspnoestatus vom Ausgangswert bis zu Woche 12 durch den gemessenen TDI (focal score und 3-Dimensionen-Score).</p> <p>Prozentualer Anteil der TDI-Responder: Das Maß des BDI und TDI-Schweregrades der Atemnot bei symptomatischen Patienten. Mit dem BDI wurde die Schwere der Dyspnoe zu Beginn der Studie gemessen und mit dem TDI wurden Änderungen dieser Ausgangs-Messung bewertet. Beide Indizes enthalten eine Bewertung für 3 Kategorien, die wichtige Faktoren für die Entwicklung der Dyspnoe darstellen. Das interessierende Outcome war die Anzahl (%) der Patienten, die einen klinisch aussagefähigen Unterschied im TDI-focal score (Verbesserung von ≥ 1 Einheit) in Woche 12 erreichten.</p>
ACCORD II	<p>Der BDI / TDI stellen ein geprüfetes Instrument dar, das die Schwere der Atemnot bei symptomatischen Patienten misst. Der BDI wurde während der Visite 2 (vor der Gabe des Prüfmedikamentes) beurteilt; Der TDI wurde während Visite 6 (12 Wochen) beurteilt. Das interessierende Outcome war die Veränderung des Dyspnoestatus der vom Ausgangswert bis zu Woche 12 durch den gemessenen TDI (focal score und 3-Dimensionen-Score). Das interessierende Outcome war die Veränderung des Dyspnoestatus vom Ausgangswert bis zu Woche 12 durch den gemessenen TDI (focal score und 3-Dimensionen-Score).</p> <p>Prozentualer Anteil der TDI-Responder: Eine Änderung von 1 Einheit beim TDI wird als Kriterium für die minimale aussagefähige Verbesserung verwendet. TDI-Responder sind solche mit einem TDI ≥ 1 Einheit. Patienten mit einem Anstieg von weniger als 1 Einheit werden als Non-Responder eingestuft. Das interessierende Outcome war die Anzahl (%) der Patienten, die einen klinisch aussagefähigen Unterschied im TDI- focal score (Verbesserung von ≥ 1 Einheit) innerhalb von 12 Wochen erreichten.</p>
ATTAIN	<p>Das Ziel des BDI und TDI war es, den Schweregrad der Atemnot bei symptomatischen Patienten zu messen. Mit dem BDI wurde die Schwere der Dyspnoe zu Beginn der Studie gemessen und mit dem TDI wurden Änderungen dieser Ausgangs-Messung bewertet. Für diese Studie wurde der BDI während Visite 2 und der TDI bei Visite 6 (Woche 12) und Visite 8 (Woche 24) angewandt. Das interessierende Outcome war die Veränderung des Dyspnoestatus vom Ausgangswert bis zu Woche 12 durch den gemessenen TDI (focal score und 3-Dimensionen-Score).</p> <p>Prozentualer Anteil der TDI-Responder: Eine Änderung von 1 Einheit in TDI wurde - als Kriterium für eine minimale aussagefähige Verbesserung verwendet. Transition Dyspnoea Index Responder waren solche mit einem TDI ≥ 1 Einheit. Patienten mit einer geringeren Erhöhung wurden als Non-Responder eingestuft. Das interessierende Outcome war die Anzahl (%) der Patienten, die einen klinisch aussagefähigen Unterschied im TDI- focal score (Verbesserung von ≥ 1 Einheit) in Woche 12 erreichten.</p>
Brusasco 2003	<p>Die Dyspnoe wurde anhand des Basis-Dyspnoe-Index (BDI) und des Transitions-Dyspnoe-Index (TDI) bestimmt. BDI und TDI bestehen aus drei Achsen (funktionale Beeinträchtigung, Ausmaß der Aufgabe, und Ausmaß des Einflusses), die zur Berechnung eines focal scores summiert werden. Der BDI wurde am Ende des Baseline-Zeitraumes angewandt und der TDI wurde bei Folgevisiten in der 8., 16. und 24. Woche angewandt.</p> <p>Ein minimaler klinisch relevanter Unterschied im TDI wurde als eine Verbesserung von mindestens 1 Einheit im TDI-focal score festgelegt.</p>
INHANCE	<p>Die Dyspnoe wurde mit dem TDI beurteilt.</p> <p>Der TDI wurde in den Wochen 4, 8, 12 und 26 gemessen.</p>
Donohue	<p>Die Dyspnoe wurde anhand des Baseline-Dyspnoe-Index (BDI) und des Transitions-Dyspnoe-</p>

2002	Index (TDI) bewertet. Der BDI und der TDI bestehen aus drei Achsen (Funktionsbeeinträchtigung, Ausmaß der Aufgabe, und Ausmaß der Anstrengung), die zur Berechnung eines focal scores summiert werden. Der BDI wurde am Ende des Baseline-Zeitraumes angewandt und der TDI wurde bei den angegebenen Folgeuntersuchungen (Woche 2, 8, 16, 24) angewandt. Ein höherer TDI-Wert bedeutet eine Verbesserung. Eine Änderung von 1 Einheit wird als klinisch signifikant angesehen.
------	--

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-92: Bewertung des Verzerrungspotenzials für TDI-Responder in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Brusasco 2003	Niedrig	Ja	Unklar	Unklar	Nein	Niedrig
INHANCE	Niedrig	Nein	Nein	Nein	Nein	Hoch
Donohue 2002	Niedrig	Ja	Nein	Nein	Nein	Hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Von den Aclidiniumbromid-Studien wurde das Verzerrungspotenzial nur bei ACCORD II für TDI-Responder als hoch eingestuft (Tabelle 4-16). Zwei von drei der Tiotropiumbromid-Studien wiesen besonders bezüglich der ergebnisbezogenen TDI-Responder ein hohes Verzerrungspotenzial auf, weil das ITT-Prinzip nicht für die Endpunkt-Analyse verwendet wurde (detaillierte Erklärung im Anhang 4-G).

Stellen Sie die Ergebnisse für TDI-Responder für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-93: Ergebnisse für TDI-Responder aus RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N Responder (%)	Gruppendifferenz (SE) [KI] p-Wert
ACCORD I	AB	190	91 (47,7%)	1,82 (0,21) [1,20; 2,77] p=0,005
	Placebo	185	62 (32,9%)	
ACCORD II	AB	177	90 (50,7%)	1,95 (0,22) [1,28; 2,99] p=0,003
	Placebo	182	63 (34,5%)	
ATTAIN	AB	269	160 (59,5%)	1,99 (0,18) [1,41; 2,8] p<0,001
	Placebo	273	116 (42,4%)	
Donohue 2010	Tiotropiumbromid 18µg	415	228 (55%)	1,68 (0,14) [1,27; 2,21] p<0,001
	Placebo	418	176 (42%)	

N=Anzahl der Patienten, Responder = Anzahl der Patienten mit einer Reduzierung des TDI Wertes von mindestens einer Einheit im Vergleich zum Ausgangswert, SE=Standardfehler, ITT=Intention-to-treat-Population, KI=Konfidenzintervall, AB=Aclidiniumbromid

Tabelle 4.3-93 zeigt die Ergebnisse der Aclidiniumbromid-400µg- und Tiotropiumbromid-18µg-Studien hinsichtlich des Prozentanteils der TDI-Responder (d. h. Patienten mit einem TDI-Score von ≥ 1 Einheit) nach einem Behandlungszeitraum von 12 Wochen. Nur eine Tiotropiumbromid-Studie berichtete dieses Outcome, daher war eine Meta-Analyse der Tiotropiumbromid-Studien nicht möglich.

Tabelle 4.3-94: Ergebnisse für TDI-Responder aus RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N Responder (%)	Gruppendifferenz (SE) [KI] p-Wert
ATTAIN	AB	269	149 (56,9%)	1,66 (0,17) [1,17; 2,32] p= 0,003
	Placebo	273	117 (45,5%)	
Brusasco 2003	Tiotropiumbromid 18µg	386	173 (44,8%)	1,78 (0,15) [1,33; 2,39] p<0,001
	Placebo	362	119 (32,9%)	
Donohue 2002	Tiotropiumbromid 18µg	209	88 (42,1%)	2,08 (0,21) [1,37; 3,17] p=0,001
	Placebo	201	52 (25,9%)	
Donohue 2010	Tiotropiumbromid 18µg	415	179 (43,1%)	1,56 (0,16) [1,13; 2,14] p=0,006
	Placebo	418	145 (34,7%)	

N=Anzahl der Patienten, Responder = Anzahl der Patienten mit einer Reduzierung des TDI Wertes von mindestens einer Einheit im Vergleich zum Ausgangswert, SE=Standardfehler, ITT=Intention-to-treat-Population, KI=Konfidenzintervall, AB=Aclidiniumbromid

Tabelle 4.3-94 zeigt die Ergebnisse der Aclidiniumbromid- und Tiotropiumbromid-18µg-Studien hinsichtlich des Prozentsatzes der TDI-Responder nach einem Behandlungszeitraum von 24 Wochen. Nur eine Aclidiniumbromid-Studie berichtete dieses Outcome, daher war eine Meta-Analyse der Aclidiniumbromid-Studien nicht möglich.

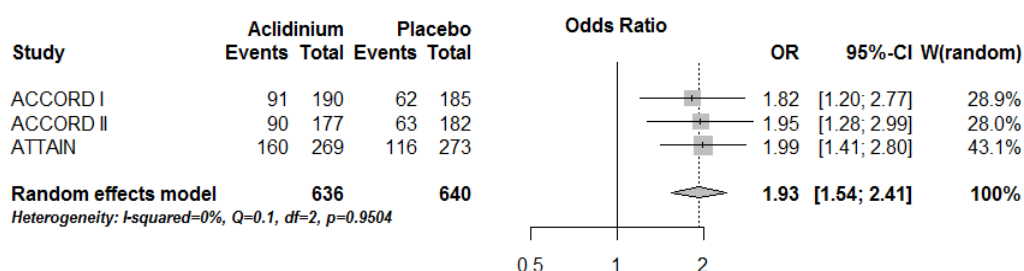


Abbildung 38: Meta-Analyse für den Prozentsatz der TDI-Responder (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo.

Abbildung 38 zeigt, dass Aclidiniumbromid wirksamer ist als Placebo (gepooltes OR 1,93 [95%-KI 1,54; 2,41]). Für diesen Endpunkt bestand keine Studienheterogenität (Q=0,1; p=0,95; I²=0%).

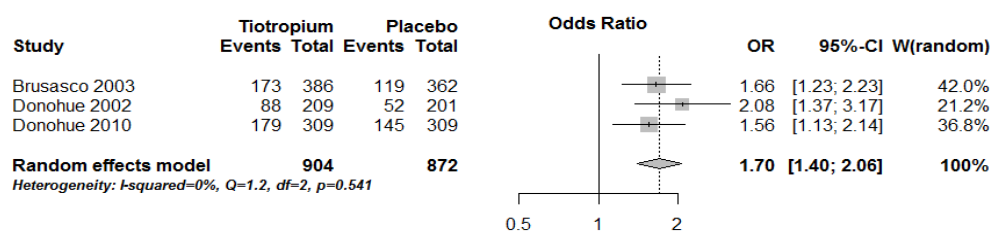


Abbildung 39: Meta-Analyse für den Prozentsatz der TDI-Responder (24 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.

Abbildung 39 zeigt, dass Tiotropiumbromid nach einem Behandlungszeitraum von 24 Wochen wirksamer war als Placebo (gepooltes OR 1,70 [95%-KI 1,40; 2,06]). Für dieses Outcome bestand keine Studienheterogenität ($Q=1,2$; $p=0,54$; $I^2=0,0\%$).

Sensitivitätsanalyse

Die ACCORD-II-Studie wies hinsichtlich der Baseline-Patientencharakteristika ein leichtes Ungleichgewicht zwischen den Behandlungsgruppen auf. Um die Auswirkungen des Ausschlusses dieser Studie aus der Meta-Analyse zu ermitteln, wurde eine Sensitivitätsanalyse durchgeführt (Abbildung 40).

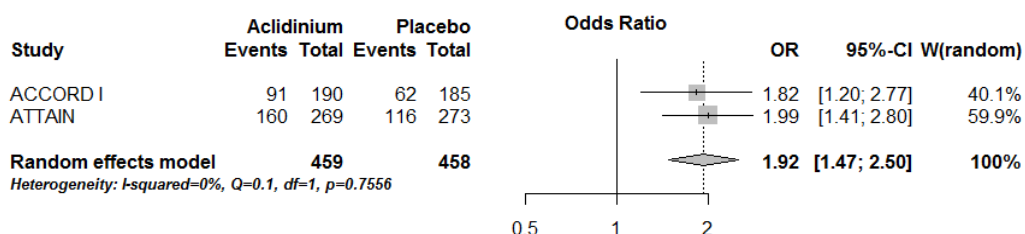


Abbildung 40: Sensitivitätsanalyse (Verzerrungspotential) für TDI-Responder nach 12 Wochen aus RCT; Aclidiniumbromid im Vergleich zu Placebo.

Abbildung 40 zeigt, dass im Vergleich zu Placebo durch die Behandlung mit Aclidiniumbromid (12 Wochen) die TDI-Responsewahrscheinlichkeit erhöht wird (OR 1,92 [95%-KI 1,47; 2,50]). Wie aus den Konfidenzintervallen der Meta-Analyse ersichtlich ist, hatte der Ausschluss der ACCORD-II-Studie keine signifikanten Auswirkungen auf den gepoolten Effektschätzer. Der Heterogenitätstest ($Q=0,1$; $p<0,76$; $I^2=0,0\%$) ergab keine Hinweise auf signifikante Studienheterogenität.

Für die Tiotropiumbromid-Studien konnte keine Sensitivitätsanalyse bezüglich des Verzerrungspotentials durchgeführt werden, da zwei von drei Studien verzerrt waren.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-95 zeigt die DIC-Schätzwerte für die geplanten NMAs. Nach 12 Wochen ergab sich aus den DIC-Ergebnissen als bestes Modell ein Modell mit festen Effekten. Dies stimmt auch mit den Ergebnissen der Meta-Analyse für Acridiniumbromid überein; hier ließ der Heterogenitätstest auf keine signifikante Heterogenität schließen. Nach 24 Wochen ergab sich aus den DIC-Ergebnissen als bestes Modell ein Modell mit festen Effekten. (Tabelle 4.3-95). Dies stimmte auch mit den Ergebnissen der Meta-Analyse für Tiotropiumbromid 18µg überein; hier ließ der Heterogenitätstest auf keine signifikante Heterogenität schließen.

Tabelle 4.3-95: DIC des indirekten Therapievergleichs zwischen Acridiniumbromid und Tiotropiumbromid 18µg; Anteil der TDI-Responder

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung für das NMA-Modell
Woche 12			
Ohne LABA-Begleitmedikation	59,6	61,4	Feste Effekte
Alle Studien	-	-	-
Woche 24			
Ohne LABA-Begleitmedikation	63,0	64,2	Feste Effekte
Alle Studien	-	-	-

Die Ergebnisse der Netzwerk-Meta-Analyse sind in Tabelle 4.3-96 und Tabelle 4.3-97 aufgeführt. Szenario 2 wurde nicht durchgeführt, da alle Patienten in den Studien, die Daten zu dem Endpunkt TDI-Responder nach 12 und 24 Wochen berichteten, keine Begleitmedikation mit LABA einsetzten.

Tabelle 4.3-96: TDI-Responder (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 18
TIO 18	Effektschätzer	1,69	
	95% KI	1,28	2,21
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	1,94	
	95% KI	1,54	2,42
	P(Überlegenheit)	>99%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Aclidiniumbromid 400 µg;

Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 12 Wochen erhöhte Aclidiniumbromid die Anzahl der TDI-Responder (Anzahl der Patienten mit einer Verbesserung des TDI Focal Score ≥ 1 Punkt) im Vergleich zu Placebo [OR 1,9 (95%-KI: 1,5, 2,4)]. Die von der Netzwerk-Meta-Analyse ermittelten Ergebnisse für den Vergleich von Aclidiniumbromid 400µg, Tiotropiumbromid und Placebo stimmten mit den Ergebnissen der Meta-Analysen überein. Beide Analysen ergaben, dass Aclidiniumbromid 400µg die Anzahl der TDI-Responder im Vergleich zu Placebo erhöht. Der Effekt bei Aclidiniumbromid 400µg war dem bei Tiotropiumbromid gleichwertig [OR 1,2 (95%-KI: 0,8, 1,6)] (Tabelle 4.3-96).

Tabelle 4.3-97: TDI-Responder (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 18
TIO 18	Effektschätzer	1,64	
	95% KI	1,35	1,96
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	1,68	
	95% KI	1,18	2,33
	P(Überlegenheit)	>99%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Aclidiniumbromid 400 µg;

Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen zeigte sich Aclidiniumbromid 400µg wirksamer als Placebo [OR 1,7 (95%-KI: 1,2, 2,3)]. Dies stimmt auch mit den Ergebnissen der ATTAIN-Studie überein. Im Vergleich zu Tiotropiumbromid ist Aclidiniumbromid 400µg als gleichwertig zu betrachten [OR 1,0 (95%-KI: 0,7, 1,5)] (Tabelle 4.3-97).

Sowohl mit Acclidiniumbromid als auch mit Tiotropiumbromid war der Anteil der TDI-Responder nach 12 und 24 Wochen signifikant höher als bei Behandlung mit Placebo. Sowohl nach 12 Wochen als auch nach 24 Wochen war der Effekt beider Behandlungen auf die TDI-Responderrate gleichwertig, das 95 %-Konfidenzintervall des indirekten Vergleich schloss zu beiden Zeitpunkten die Eins mit ein, d. h. es ergaben sich keine statistisch signifikanten Unterschiede auf diesen Endpunkt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.5 CFB EXACT-Respirationssymptome – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-98: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB EXACT-Respirationssymptome – 12 Wochen]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•

Tabelle 4.3-99: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB EXACT-Respirationssymptome – 24 Wochen]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Für den Endpunkt CFB EXACT-Respirationssymptome berichtete nur eine Aclidiniumbromid-Studie Ergebnisse. Es konnte kein indirekter Vergleich mit Tiotropiumbromid durchgeführt werden.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-100: Operationalisierung von CFB EXACT-Respirationssymptomen

Studie	Operationalisierung
ATTAIN	Das Messinstrument für Exazerbationen der chronischen Lungenerkrankung (Exacerbations of Chronic Pulmonary Disease Tool) – der Patientenfeedback-Fragebögen (EXACT-PRO), ist in verschiedenen Sprachen validiert. EXACT ist ein einzelnes von Patienten berichtetes Outcome (Patient Reported Outcome - PRO), um die Auswirkungen pharmakologischer Behandlungen (präventiv und kurativ) auf akute Exazerbationen von COPD, einschließlich der Präsenz /Häufigkeit chronischer Bronchitis, Schwere, Dauer und Auflösung zu bewerten. Die Veränderung zur Baseline (CFB) wird durch die Gegenüberstellung der Baseline- und Follow-Up Werte ermittelt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-101: Bewertung des Verzerrungspotenzials für CFB EXACT-Respirationssymptome in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die ATTAIN-Studie hatte ein geringes Verzerrungspotenzial auf Studienebene wie auch ein geringes Endpunkt-spezifisches Verzerrungspotenzial für CFB EXACT-Respirationssymptome (Tabelle 4-16; Details sind in Anhang 4-G beschrieben).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-102: Ergebnisse für CFB EXACT-Respirationssymptome aus RCT für indirekte Vergleiche (CFB nach 12 Wochen)

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mittelwert (SD)	Mittelwert (SD)	CFB (SD)	Gruppendifferenz (SE) [KI] p-Wert
ATTAIN	AB	269	14,1 (6,4)	11,7 (6,7)	-2,5 (0,29)	-1,98 (0,4) [-2,76; -1,19] p < 0,001
	Placebo	273	13,6 (6,6)	13,3 (6,8)	-0,5 (0,29)	

N=Anzahl der Patienten, SD=Standardabweichung CFB= Veränderung gegenüber dem Ausgangswert, SE= Standardfehler ITT= Intention-to-treat-Population KI=Konfidenzintervall, AB= Acridiniumbromid

Tabelle 4.3-103: Ergebnisse für CFB EXACT-Respirationssymptome aus RCT für indirekte Vergleiche (CFB nach 24 Wochen)

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mittelwert (SD)	Mittelwert (SD)	CFB (SD)	Gruppendifferenz (SE) [KI] p-Wert
ATTAIN	AB	269	14,1 (6,4)	11,6 (6,4)	-2,5 (4,2)	-2,31 (0,43) [-3,15; -1,46] p < 0,001
	Placebo	273	13,6 (6,6)	13,2 (6,6)	-0,4 (4,2)	

N=Anzahl der Patienten, SD=Standardabweichung CFB= Veränderung gegenüber dem Ausgangswert, SE= Standardfehler ITT= Intention-to-treat-Population KI=Konfidenzintervall, AB= Acridiniumbromid

Tabelle 4.3-102 zeigt die Ergebnisse der ATTAIN-Studie nach 12 Wochen und Tabelle 4.3-103 zeigt die Ergebnisse derselben Studie nach 24 Wochen. Die Acridiniumbromid-Behandlung ergab deutlich weniger Symptome. Die Durchführung einer Meta-Analyse war nicht möglich.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Für die beiden Zeitpunkte 12 und 24 Wochen war die ATTAIN-Studie die einzige Studie, die Daten über CFB EXACT-Respirationssymptome berichtete. Daher wurde kein indirekter Vergleich durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.6 CFB Nächtliche COPD-Symptome – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-104: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Nächtliche COPD-Symptome – 12 Wochen]

Anzahl Studien	Referenzen Studien	Aclidiniumbromid Bromide 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ACCORD I	•			•

Tabelle 4.3-105: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Nächtliche COPD-Symptome – 24 Wochen]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar

Für die beiden Zeitpunkte 12 und 24 Wochen berichtete jeweils nur eine Aclidiniumbromid-Studie über die Veränderung der nächtlichen COPD-Symptome. Daher konnten keine indirekten Vergleiche mit Tiotropiumbromid durchgeführt werden. Die Ergebnisse der Aclidiniumbromid-Studien werden der Vollständigkeit halber im Folgenden dargestellt.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-106: Operationalisierung von CFB nächtliche COPD-Symptome

Studie	Operationalisierung
ACCORD I	Die Fragebögen zu nächtlichen COPD-Symptomen wurden vom Patienten täglich in das eDiary eingegeben. Der Fragebogen enthält insgesamt 9 Items die rückblickend auf einen Zeitraum von 24 Stunden bewertet werden. Der Fragebogen wurde so konzipiert, dass er morgens vom Patienten selbstständig ausgefüllt werden kann. Die ersten 5 Items wurden unter Verwendung von Antwortmöglichkeiten aus dem bestehenden COPD-Symptom-Fragebogen nach Welte et al. [35] erarbeitet. Diese Items beurteilen die Wirkung von COPD-Symptomen auf den Schlaf, die produzierte Sputummenge während des Schlafes, die Höhe der Sputum-Produktion in den letzten 24 Stunden und Einfluss von Atemnot auf Aktivitäten in den letzten 12 Stunden. Vom Sponsor wurden zusätzliche Items entwickelt, um die Häufigkeit des Auftretens in der vergangenen Nacht für folgende COPD-Symptome zu beurteilen: Gefühl von Atemlosigkeit, Keuchen beim Atmen, Husten und Auswurf (4 Items), und Verwendung von Notfall-Medikamenten (5 Items). Die Veränderung zur Baseline (CFB) wird durch die Gegenüberstellung der Baseline- und Follow-Up Werte ermittelt.
ATTAIN	Die nächtlichen Symptome der COPD wurden täglich morgens elektronisch beurteilt. Es handelt sich dabei um einen 6-Item-Fragebogen für die Frage nach dem Auftreten von Atemnot, Husten, Auswurf, Brustenge oder Kongestion und Keuchen in der Nacht. Der Fragebogen beurteilt, wie diese Symptome den Schlaf. Die Fragebögen konnten durch die Patienten mittels eines PDA (elektronisches Tagebuch) selbst ausgefüllt werden. Die Veränderung zur Baseline (CFB) wird durch die Gegenüberstellung der Baseline- und Follow-Up Werte ermittelt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-107: Bewertung des Verzerrungspotenzials für CFB nächtliche COPD-Symptome in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Beide Aclidiniumbromid-Studien hatten über die gesamte Studienebene ein geringes Verzerrungspotenzial (Tabelle 4.3-107). Das Endpunkt-spezifische Verzerrungspotenzial für

die Veränderung der nächtlichen COPD-Symptome wurde für die beiden Studien als gering angesehen (Details sind in Anhang 4-G beschrieben).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-108: Ergebnisse für CFB nächtliche COPD-Symptome (12 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten	Baseline	Follow-up	Veränderung	
		(ITT)	Mean (SD)	Mean (SD)	CFB (SD)	Gruppen-differenz ¹ (SE) [KI] p-Wert
N						
ACCORD I						
Häufigkeit von nächtlichem Husten	Aclidinium-bromid	190	1,92 (1,61)	-	-0,36 (1,29)	-0,46 (0,13) [-0,72; -0,20]
	Placebo	185	2,06 (1,50)	-	0,10 (1,36)	p<0,001
Häufigkeit von nächtlichem Auswurf	Aclidinium-bromid	190	1,35 (1,55)	-	-0,37 (0,92)	-0,42 (0,09) [-0,60; -0,24]
	Placebo	185	1,33 (1,43)	-	0,05 (0,98)	p<0,001
Häufigkeit der nächtlichen Atemnot	Aclidinium-bromid	190	1,41 (1,27)	-	-0,44 (1,12)	-0,31 (0,11) [-0,53; -0,09]
	Placebo	185	1,41 (1,20)	-	-0,13 (0,92)	p=0,005
Häufigkeit von nächtlichem Keuchen	Aclidinium-bromid	190	1,25 (1,47)	-	-0,53 (1,27)	-0,53 (0,13) [-0,79; -0,27]
	Placebo	185	1,33 (1,47)	-	0 (1,15)	p<0,001
Schwere und Grad der Einschränkung durch Atemprobleme im Schlaf	Aclidinium-bromid	190	0,85 (0,77)	-	-0,24 (0,57)	-0,18 (0,06) [-0,30; -0,06]
	Placebo	185	0,82 (0,72)	-	-0,06 (0,59)	p=0,003
Schwere und Grad der Einschränkung durch Atemnot	Aclidinium-bromid	190	1,72 (0,92)	-	-0,44 (0,86)	-0,25 (0,09) [-0,42; -0,08]
	Placebo	185	1,82 (0,92)	-	-0,19 (0,70)	p=0,005
Schwere und Grad der Einschränkung durch Husten	Aclidinium-bromid	190	1,39 (0,97)	-	-0,24 (0,76)	-0,14 (0,08) [-0,29; 0,01]
	Placebo	185	1,49 (0,88)	-	-0,10 (0,78)	p=0,08

¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, N = Anzahl der Patienten, CFB= Veränderung gegenüber dem Ausgangswert, SD = Standardabweichung, SE = Standardfehler, ITT = intention-to-treat population, KI = Konfidenzintervall

Tabelle 4.3-108 zeigt die Ergebnisse der Aclidiniumbromid-Studie die über die Veränderung der nächtlichen COPD-Symptome nach 12 Wochen berichtet. Die Behandlung mit

Acclidiniumbromid führte nach 12 Wochen zu wesentlich weniger nächtlichen COPD-Symptomen.

Tabelle 4.3-109: Ergebnisse für CFB nächtliche COPD-Symptome (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mean (SD)	Mean (SD)	CFB (SD)	Gruppen- differenz ¹ (SE) [KI] p-Wert
ATTAIN						
Auswurf	AB	269	57,7 (37,1)	46,4 (38,7)	-11,3 (53,6)	-7,9 (4,62)
	Placebo	273	56,0 (37,1)	52,6 (39,6)	-3,4 (54,3)	[-16,96; 1,16] p=0,087
Brustenge oder Kongestion	AB	269	27,9 (31,1)	23,7 (32,0)	-4,2 (44,6)	-5,0 (3,85)
	Placebo	273	26,6 (31,2)	27,4 (34,9)	0,8 (46,8)	[-12,54; 2,54] p=0,087
Husten	AB	269	63,2 (37,5)	51,0 (38,9)	-12,2 (54,0)	-6,0 (4,66)
	Placebo	273	65,3 (35,3)	59,1 (38,7)	-6,2 (52,4)	[-15,13; 3,13] p=0,198
Gefühl von Atemlosigkeit	AB	269	35,8 (34,3)	30,2 (36,6)	-5,6 (50,2)	-7,5 (8,48)
	Placebo	273	34,2 (34,5)	36,1 (38,4)	1,9 (51,6)	[-15,98; 0,98] p=0,376
Keuchen	AB	269	37,0 (38,4)	28,2 (37,1)	-8,8 (53,4)	-4,2 (4,60)
	Placebo	273	31,4 (35,7)	26,8 (36,0)	-4,6 (50,7)	[-13,22; 4,82] p=0,361
Jegliches Symptome	AB	269	84,0 (27,3)	74,1 (33,7)	-9,9 (43,4)	-6,6 (3,74)
	Placebo	273	84,2 (26,4)	80,9 (30,8)	-3,3 (40,6)	[-13,93; 0,73] p=0,078

¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, N = Anzahl der Patienten, CFB= Veränderung gegenüber dem Ausgangswert, SD = Standardabweichung, SE = Standardfehler, ITT = intention-to-treat population, KI = Konfidenzintervall

Tabelle 4.3-109 zeigt die Ergebnisse der ATTAIN-Studie, die als einzige Studie über die Veränderung der nächtlichen COPD-Symptome nach 24 Wochen berichtete. Die Acclidiniumbromid-Behandlung führte im Vergleich zu Placebo zu weniger nächtlichen Symptomen.

Für beide Zeitpunkte, 12 und 24 Wochen, berichtet jeweils nur eine Acclidiniumbromid-Studie über die Veränderung der nächtlichen COPD-Symptome. Eine Meta-Analyse konnte nicht durchgeführt werden.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Für beide Zeitpunkte, 12 und 24 Wochen, berichteten nur die Aclidiniumbromid-Studien über die Veränderung der nächtlichen COPD-Symptome. Daher wurde kein indirekter Vergleich durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.7 CFB Morgendliche COPD-Symptome – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-110: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Morgendliche COPD-Symptome – 12 Wochen]

Anzahl Studien	Referenzen Studien	Aclidiniumbromid Bromide 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ACCORD I	•			•

Tabelle 4.3-111: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [CFB Morgendliche COPD-Symptome – 24 Wochen]

Anzahl Studien	Referenzstudien	Acclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Für die beiden Zeitpunkte 12 und 24 Wochen berichtete jeweils nur eine Acclidiniumbromid-Studie über die Veränderung der morgendlichen COPD-Symptome. Daher konnten keine indirekten Vergleiche mit Tiotropiumbromid durchgeführt werden. Die Ergebnisse der Acclidiniumbromid-Studien werden der Vollständigkeit halber im Folgenden dargestellt.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-112: Operationalisierung von CFB morgendliche COPD-Symptome

Studie	Operationalisierung
ACCORD I	Fragen zu morgendlichen Symptomen mussten die Patienten täglich morgens mit Hilfe eines eDiarys beantworten. Dabei bewerteten die Patienten rückblickend die morgendliche Aktivitätseinschränkung wegen Atemnot und den Grad der Atemnot in der ersten Stunde nach dem Aufstehen. Diese beiden Fragen sind Teil des COPD-Symptom-Fragebogen von Welte et al. [35]. Die Veränderung zur Baseline (CFB) wird durch die Gegenüberstellung der Baseline- und Follow-Up Werte ermittelt.
ATTAIN	Die morgendlichen Symptome der COPD wurden täglich morgens elektronisch beurteilt. Es handelt sich dabei um einen 6-Item-Fragebogen für die Frage nach dem Auftreten von Atemnot, Husten, Auswurf, Brustenge oder Kongestion und Keuchen am Morgen. Der Fragebogen beurteilt, wie diese Symptome die Patienten in ihren morgendlichen Tätigkeiten einschränken. Der Fragebogen konnte durch die Patienten mittels eines PDA (elektronisches Tagebuch) selbst ausgefüllt werden. Die Veränderung zur Baseline (CFB) wird durch die Gegenüberstellung der Baseline- und Follow-Up Werte ermittelt.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-113: Bewertung des Verzerrungspotenzials für CFB morgendliche COPD-Symptome in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Beide Studien hatten ein geringes Verzerrungspotenzial auf Studienebene wie auch ein geringes Endpunkt-spezifisches Verzerrungspotenzial für den Endpunkt morgendliche COPD-Symptome (Tabelle 4.3-113; Details siehe Anhang 4-G).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-114: Ergebnisse für CFB morgendliche COPD-Symptome (12 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) [KI] p-Wert
ACCORD I						
Grad der Einschränkung bei morgendlichen Aktivitäten	AB	190	1,39 (0,88)	-	-0,28 (0,76)	-0,25 (0,08) [-0,40; -0,10] p=0,002
	Placebo	185	1,42 (0,87)	-	-0,03 (0,56)	
Schwere der Atemnot in der ersten Stunde nach dem Aufstehen am Morgen	AB	190	1,51 (0,92)	-	-0,32 (0,79)	-0,23 (0,08) [-0,39; -0,07] p=0,004
	Placebo	185	1,57 (0,92)	-	-0,09 (0,61)	

¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, N = Anzahl der Patienten, CFB= Veränderung gegenüber dem Ausgangswert, SD = Standardabweichung, SE = Standardfehler, ITT = intention-to-treat population, KI = Konfidenzintervall

Tabelle 4.3-114 zeigt die Ergebnisse der ACCORD I-Studie, die als einzige Studie über den Endpunkt morgendlicher COPD-Symptome nach 12 Wochen berichtete. Die Acclidiniumbromid-Behandlung ergab deutlich weniger morgendliche Symptome.

Tabelle 4.3-115: Ergebnisse für CFB morgendliche COPD-Symptome (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studien-arm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) [KI] p-Wert
ATTAIN						
Auswurf	AB	269	64,9 (37,7)	60,3 (37,9)	-4,6 (53,5)	-8,6 (4,61) [-17,64; 0,44] p=0,062
	Placebo	273	63,9 (36,6)	67,9 (36,6)	4,0 (51,8)	
Brustenge oder Kongestion	AB	269	29,4 (32,9)	26,7 (34,2)	-2,7 (47,5)	-6,8 (4,09) [-14,82; 1,22] p=0,096
	Placebo	273	26,1 (31,5)	30,2 (36,8)	4,1 (48,4)	
Husten	AB	269	68,9 (37,0)	62,8 (38,1)	-6,1 (53,1)	-4,6 (4,56) [-13,57; 4,37] p=0,313
	Placebo	273	70,8 (34,2)	69,3 (35,8)	-1,5 (49,5)	
Gefühl von Atemlosigkeit oder Atemnot	AB	269	41,9 (37,4)	38,0 (39,2)	-3,9 (54,2)	-7,2 (4,67) [-16,36; 1,96] p=0,123
	Placebo	273	41,2 (36,8)	44,5 (40,0)	3,3 (54,4)	
Keuchen	AB	269	37,8 (38,3)	30,6 (38,3)	-7,2 (54,2)	-5,6 (4,67) [-14,76; 3,56] p=0,230
	Placebo	273	31,7 (36,3)	30,1 (37,8)	-1,6 (52,4)	
Jegliches Symptom	AB	269	89,5 (23,2)	85,0 (26,0)	-4,5 (34,8)	-5,2 (3,0) [-11,08; 0,68] p=0,083
	Placebo	273	89,3 (22,6)	90,0 (22,9)	0,7 (32,2)	

¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, N = Anzahl der Patienten, CFB= Veränderung gegenüber dem Ausgangswert, SD = Standardabweichung, SE = Standardfehler, ITT = intention-to-treat population, KI = Konfidenzintervall

Tabelle 4.3-115 zeigt die Ergebnisse der ATTAIN-Studie, die als einzige Studie über die Veränderung der morgendlichen COPD-Symptome nach 24 Wochen berichtete. Über alle Punkte wurde separat berichtet. Nach 24 Wochen führte die Behandlung mit Acclidiniumbromid zu wesentlich weniger morgendlichen COPD-Symptomen.

Für beide Zeitpunkte, 12 und 24 Wochen, berichtet jeweils nur eine Acclidiniumbromid-Studie über die Veränderung der nächtlichen COPD-Symptome. Eine Meta-Analyse konnte nicht durchgeführt werden.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Für beide Zeitpunkte, nach 12 und 24 Wochen, berichteten nur die Acclidiniumbromid-Studien über die Veränderung der morgendlichen COPD-Symptome. Daher war kein indirekter Vergleich mit Tiotropiumbromid möglich.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.8 Exazerbationsrate – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-116: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Exazerbationsrate; begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
6	Brusasco 2003			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	MISTRAL			•	•
	Powrie 2007			•	•
	TIPHON			•	•
1	Bateman 2010a		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

In 10 Studien wurden Daten zu den Exazerbationsraten der teilnehmenden Patienten erhoben (begleitende Behandlung mit LABAs nicht gestattet). Abbildung 41 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 41: Netzwerkdiagramm für die Analyse der Exazerbationsrate (begleitende Behandlung mit LABAs nicht gestattet)



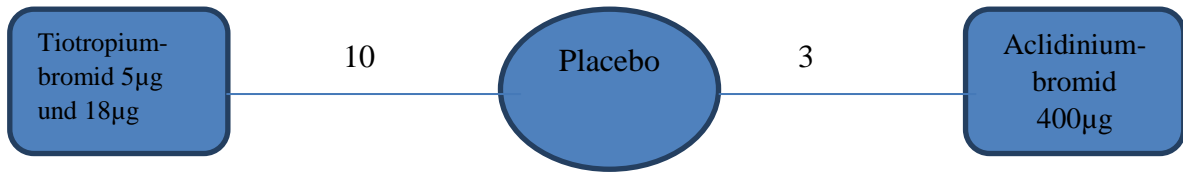
Tabelle 4.3-117: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Exazerbationsrate; alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
8	Brusasco 2003			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	MISTRAL			•	•
	Niewoehner 2005			•	•
	Powrie 2007			•	•
	UPLIFT			•	•
	TIPHON			•	•
2	Bateman 2010a		•		•
	Bateman 2010b		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

In 13 Studien wurden Daten zu den Exazerbationsraten der teilnehmenden Patienten erhoben. Abbildung 42 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 42: Netzwerkdiagramm für die Analyse der Exazerbationsrate



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-118: Operationalisierung der Exazerbationsraten

Studie	Operationalisierung
ACCORD I	<p>COPD-Exazerbationen wurden vom Prüfarzt bei jeder Untersuchung auf der Grundlage von Patienteninformationen ausgewertet. Eine COPD-Exazerbation wurde definiert als eine Zunahme der COPD-Symptome (z. B. Atemnot, Husten, Auswurfvolumen, eitriger Auswurf) an mindestens 2 aufeinander folgenden Tagen mit einem wie folgt gekennzeichneten Schweregrad:</p> <ul style="list-style-type: none"> • Leicht: die vom Patienten zu Hause durch eine Erhöhung kurzwirksamer Bronchodilatoren und / oder Anwendung von ICS [inhalative Corticosteroide] selbst zu behandeln • Mittel: führten nicht zu Krankenhausaufenthalten, wurde aber mit Antibiotika und/oder systemischen Corticosteroiden behandelt • Schwer: führte zur Hospitalisierung (Übernachtung im Krankenhaus oder in der Notaufnahme)
ACCORD II	<p>COPD-Exazerbationen wurden vom Prüfarzt bei jeder Untersuchung auf der Grundlage von Patienteninformationen ausgewertet. Die Schwere der COPD-Exazerbation wurde nach folgender Skala bewertet:</p> <ul style="list-style-type: none"> • Leicht: Erhöhung der COPD-Symptome (z.B. Atemnot, Husten, Auswurfvolumen, eitriger Auswurf) an mindestens 2 aufeinander folgenden Tagen, die vom Patienten zu Hause durch die Erhöhung kurzwirksamer Bronchodilatoren und/oder Einsatz inhalativer Corticosteroide behandelt werden • Mäßig: Zunahme der COPD-Symptome an mindestens 2 aufeinanderfolgenden Tagen, die nicht zu Krankenhausaufenthalten führt, aber mit Antibiotika und/oder systemischen Corticosteroiden behandelt wird • Schwer Zunahme der COPD-Symptome an mindestens 2 aufeinanderfolgenden Tagen, die zur Hospitalisierung (Übernachtung im Krankenhaus oder in der Notaufnahme) führt
ATTAIN	COPD-Exazerbationen wurden mit dem EXACT-PRO-Fragebogen bewertet.
Brusasco 2003	Die Exazerbationen wurden als ein Komplex respiratorischer Symptome (Neuaufreten oder eine Erhöhung von mindestens einem der Symptome Husten, Auswurf, Atemnot, Keuchen, Engegefühl in der Brust) innerhalb von mindestens 3 Tagen und in der Regel verbunden mit einem therapeutischen Eingriff.
Casaburi 2002	Ein Komplex respiratorischer Ereignisse (z. B. Husten, Keuchen, Atemnot oder Auswurf) mit einer Dauer von mehr als 3 Tagen. Diese wurden in der Regel mit Antibiotika und/oder oralen Steroiden behandelt.
MISTRAL (Dusser 2006)	Eine Exazerbation wurde definiert als das Auftreten mindestens eines klinischen Anzeichens (Verschlechterung von Dyspnoe, Husten oder Auswurf; Auftreten von eitrigem Auswurf, Fieber (> 38 °C); Auftreten neuer Anomalien bei Röntgenaufnahme des Thorax) mit einer Dauer von ≥ 2 Tagen, die eine neue Verschreibung oder eine Dosiserhöhung von Beta-2-Agonisten, Antibiotika, Corticosteroiden oder Bronchodilatoren. Eine Post-hoc-Analyse wurde durchgeführt, die eine allgemeinere Klassifizierung der Schwere der Exazerbation auf Basis von HRU und auf Behandlungsebene verwendete. Eine schwere Exazerbation wurde als eine solche eingestuft, die stationäre Behandlung erfordert. Eine mäßige Verschlechterung wurde definiert als eine, die eine Behandlung mit systemischen Steroiden und/oder Antibiotika erfordert. Alle restlichen Ereignisse wurden als leichte Exazerbationen eingestuft.
INHANCE (Donohue 2010)	COPD-Exazerbationen wurden als Auftreten oder eine Verschlechterung von einem oder mehreren respiratorischen Symptomen (Atemnot, Husten, eitriger Auswurf/Auswurfvolumen oder Keuchen) für 3 oder mehr aufeinanderfolgende Tage definiert mit einer zusätzlichen intensivierten Therapie (z. B. systemische Steroide, Antibiotika, Sauerstoff) und/oder Krankenhaus- oder Notaufnahmehaufenthalt.
Niewoehner 2005	Ein Komplex respiratorischer Symptome (Erhöhung oder Neuaufreten) von mehr als einem der Folgenden: Husten, Auswurf, Keuchen, Atemnot oder Engegefühl in der Brust mit einer Dauer von mindestens 3 Tagen, der eine Behandlung mit Antibiotika oder systemischen

	Steroiden, einen Krankenhausaufenthalt oder beides erfordert.
Powrie 2007	Eine Exazerbation wurde definiert als das Auftreten einer Erhöhung von zwei Hauptsymptomen (Atemnot, eitriger Auswurf und Auswurfvolumen) oder einem Haupt- und einem Nebensymptom (Keuchen, Halsschmerzen, Husten und allgemeine Erkältungssymptome) an ≥ 2 aufeinanderfolgenden Tagen.
UPLIFT (Tashkin 2008)	Eine Erhöhung oder Neuauftreten von mehr als einem respiratorischen Symptom (Husten, Auswurf, eitriger Auswurf, Keuchen, Atemnot) mit einer Dauer von 3 Tagen oder mehr, die eine Behandlung mit einem Antibiotikum oder einem systemischen Corticosteroid erfordern
TIPHON (Tonnel 2008)	Eine akute Exazerbation ist definiert als anhaltende Verschlechterung der COPD-Patienten (aus dem stabilen Zustand und über die normalen Tagesschwankungen hinausgehend), die akut auftraten und eine Änderung der regelmäßigen Medikation erforderten
Bateman 2010b	Die Exazerbationen wurden als ein Komplex von respiratorischen Ereignissen oder Symptomen mit einer Dauer von ≥ 3 Tagen definiert, die eine Behandlung mit Antibiotika und/oder systemischen Corticosteroiden erfordern oder die Untersuchungsperson veranlasste, die übliche Atemwegsmedikation des Patienten zu ändern.
Bateman 2010a	COPD-Exazerbationen pro Patient und Jahr wurden definiert als respiratorische Nebenwirkungen mit einer Dauer von 3 Tagen, die eine Behandlung mit Antibiotika und/oder oralen Corticosteroiden und/oder einer wesentlichen Veränderung der respiratorischen vorgeschriebenen Medikamente einschließlich inhalativer Bronchodilatoren erfordern.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-119: Bewertung des Verzerrungspotenzials für Exazerbationsraten in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Brusasco 2003	Niedrig	Ja	Ja	Unklar	Nein	Niedrig
Casaburi 2002	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
MISTRAL	Niedrig	Ja	Ja	Nein	Nein	Niedrig
INHANCE	Niedrig	Nein	Ja	Nein	Nein	Hoch
Niewoehner 2005	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Powrie 2007	Niedrig	Ja	Ja	Nein	Nein	Niedrig
UPLIFT	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
TIPHON	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Bateman 2010b	Niedrig	Ja	Ja	Unklar	Nein	Niedrig
Bateman 2010a	Niedrig	Ja	Unklar	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die ACCORD II-Studie wurde das Verzerrungspotenzial hinsichtlich der Exazerbationsraten als hoch eingestuft. Zwei der fünf Tiotropiumbromid-Studien wiesen ein hohes Verzerrungspotenzial auf Studienebene auf (SPRUCE; Vogelmeier 2008). Selbst unter Berücksichtigung des spezifischen Verzerrungspotenzials für die ergebnisbezogenen Exazerbationen ergab sich bei einer von 10 Tiotropiumbromid-Studien ein hohes Verzerrungspotenzial (INHANCE), weil das ITT-Prinzip nicht für die Endpunkt-Analyse angewendet wurde (detaillierte Erklärung im Anhang 4-G).

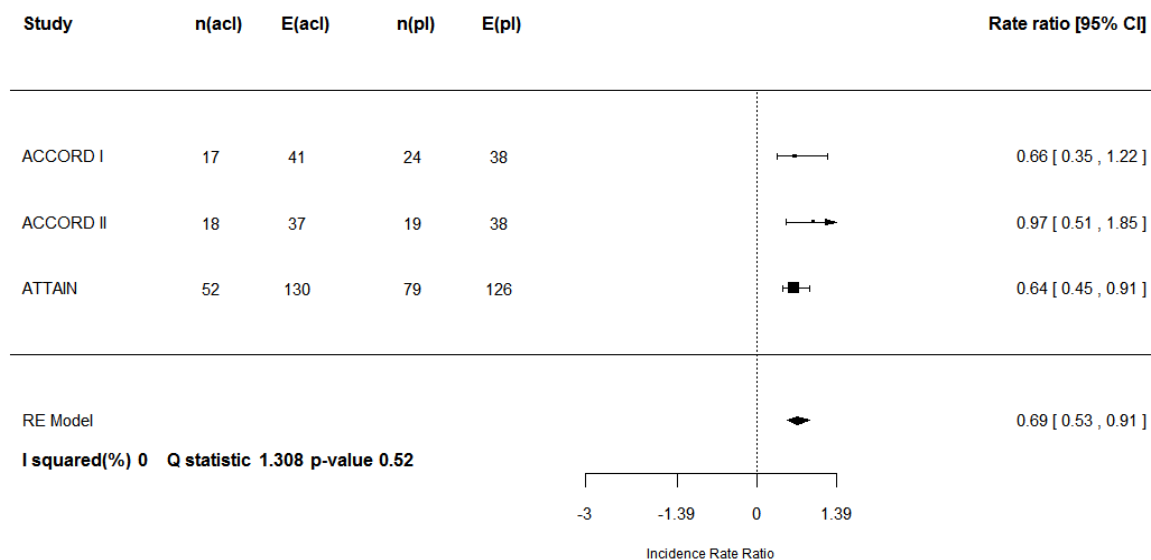
Stellen Sie die Ergebnisse für Exazerbationsraten für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-120: Ergebnisse für die Exazerbationsrate aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (N in ITT)	Anzahl Per- sonenjahre	Anzahl Exa- zerbationen	Rate	Rate ratio (SE) [KI] p- Wert
ACCORD I	Aclidiniumbromid	190	41	17	0,42	0,67 (0,30) [0,36; 1,25] p= 0,180
	Placebo	185	38	24	0,63	
ACCORD II	Aclidiniumbromid	177	37	18	0,48	0,97 (0,33) [0,51; 1,85] p= 0,990
	Placebo	182	38	19	0,50	
ATTAIN	Aclidiniumbromid	269	130	52	0,40	0,64 (0,12) [0,45; 0,91] p= 0,020
	Placebo	273	126	79	0,60	
Bateman 2010a	Tiotropiumbromid 5µg	670	618	575	0,91	0,49 (0,05) [0,44; 0,54] p<0,001
	Placebo	653	603	1152	1,91	
Bateman 2010b	Tiotropiumbromid 5µg	1989	1672	1154	0,69	0,79 (0,04) [0,73; 0,86] p< 0,001
	Placebo	2002	1659	1443	0,87	
Brusasco 2003	Tiotropiumbromid 18µg	386	186	198	1,06	0,71 (0,10) [0,59; 0,86] p<0,001
	Placebo	362	174	260	1,49	
Casaburi 2002	Tiotropiumbromid 18µg	507	537	408	0,76	0,80 (0,08) [0,69; 0,92] p=0,003
	Placebo	332	344	327	0,95	
INHANCE (Donohue 2010)	Tiotropiumbromid 18µg	415	180	95	0,53	0,74 (0,08) [0,56; 0,97] p=0,028
	Placebo	418	165	118	0,72	
MISTRAL (Dusser 2006)	Tiotropiumbromid 18µg	500	437	686	1,57	0,65 (0,05) [0,59; 0,72] p<0,001
	Placebo	510	442	1065	2,41	
Niewoehner 2005	Tiotropiumbromid 18µg	914	416	354	0,85	0,81 (0,08) [0,70;0,94] p=0,004
	Placebo	915	380	399	1,05	
Powrie 2007	Tiotropiumbromid 18µg	69	59	68	1,15	0,47 (0,15) [0,35; 0,63] p<0,001
	Placebo	73	63	154	2,44	
TIPHON (Tonnel 2008)	Tiotropiumbromid 18µg	247	185	194	1,05	0,57 (0,09) [0,48; 0,69] p<0,001
	Placebo	245	188	344	1,83	
UPLIFT (Tashkin 2010)	Tiotropiumbromid 18µg	2987	9784	7116	0,73	0,86 (0,01) [0,83; 0,88] p<0,001
	Placebo	3006	9308	7912	0,85	

N= Anzahl der Patienten ITT= Intention-to-treat-Population [KI]= Konfidenzintervall SE= Standardfehler

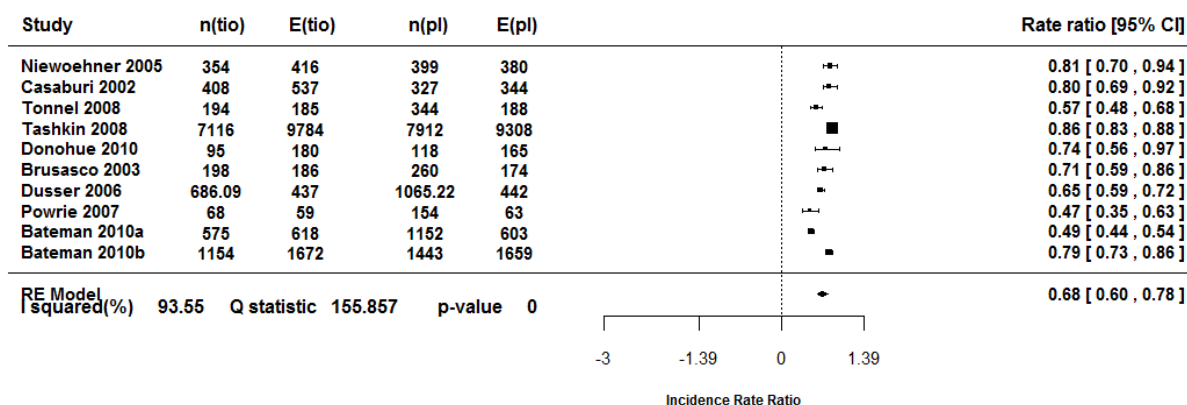
Tabelle 4.3-97 zeigt die berichteten Exazerbationsdaten der Studien mit Acclidiniumbromid 400µg sowie Tiotropiumbromid 5µg und 18µg gegen Placebo. Die Exazerbationsraten der mit Acclidiniumbromid behandelten Patienten sind geringer als die der Placebo-Patienten. In der ATTAIN-Studie war dieser Unterschied signifikant ($p=0,02$). Tiotropiumbromid (5µg und 18µg) war bezüglich der Reduzierung der Exazerbationen in allen Studien gegenüber Placebo überlegen.



n: Anzahl der Ereignisse, E: Personenjahre, (acl): Acclidiniumbromid, (pl): Placebo

Abbildung 43: Meta-Analyse für Exazerbationsraten aus RCT; Acclidiniumbromid im Vergleich zu Placebo.

Abbildung 43 zeigt, dass Acclidiniumbromid nach 12 Wochen eine bessere Wirksamkeit als Placebo bei der Reduzierung der Exazerbationsrate aufwies (gepooltes Hazard Ratio: 0,69 [95%-KI 0,53; 0,91]). Der Heterogenitätstest ($Q=1,31$; $p=0,52$; $I^2=0$) zeigt, dass es keine Heterogenität unter den Studien gibt.



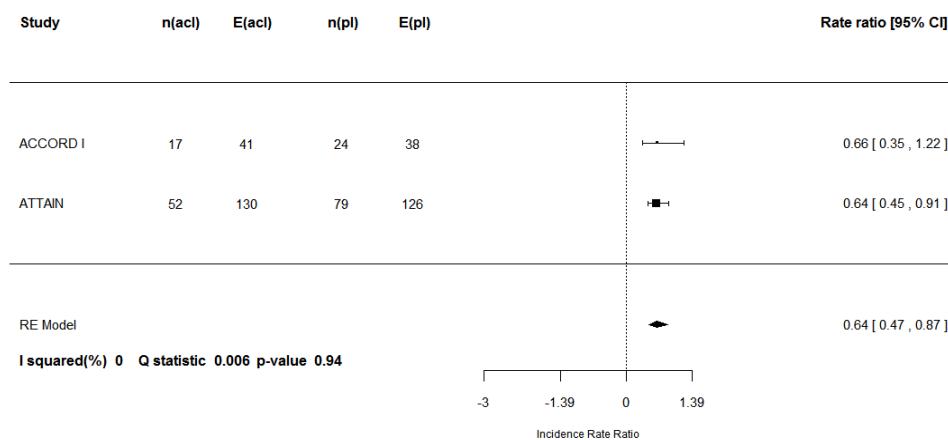
n: Anzahl der Ereignisse, E: Personenjahre, (tio): Tiotropiumbromid, (pl): Placebo

Abbildung 44: Meta-Analyse für Exazerbationsraten aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.

Abbildung 44 zeigt, dass Tiotropiumbromid eine bessere Wirksamkeit als Placebo bei der Reduzierung der Exazerbationsrate aufwies (gepooltes Ratenverhältnis 0,68 [95%-KI 0,60; 0,78]). Der Heterogenitätstest ($Q=155,9$; $p<0,001$; $I^2=93,55$) zeigt, dass eine signifikante Heterogenität zwischen den Studien existiert. Dies könnte ggf. auf die unterschiedliche Dosierung von Tiotropiumbromid (5µg oder 18µg) oder auf die unterschiedliche Definition der Exazerbationen in den Studien zurückzuführen sein. Eine Sensitivitätsanalyse wird diesen Sachverhalt näher untersuchen. Der Einfluss des Verzerrungspotentials kann nicht untersucht werden, da ein Ausschluss verzerrter Studien zu einem Studienpool mit weniger als zwei Studien führen würde.

Sensitivitätsanalyse

Die ACCORD-II-Studie wies hinsichtlich der Baseline-Patientencharakteristika ein Ungleichgewicht zwischen den Behandlungsgruppen auf. Um die Auswirkungen des Ausschlusses dieser Studie aus der Meta-Analyse zu ermitteln, wurde eine entsprechende Sensitivitätsanalyse durchgeführt (Abbildung 45).

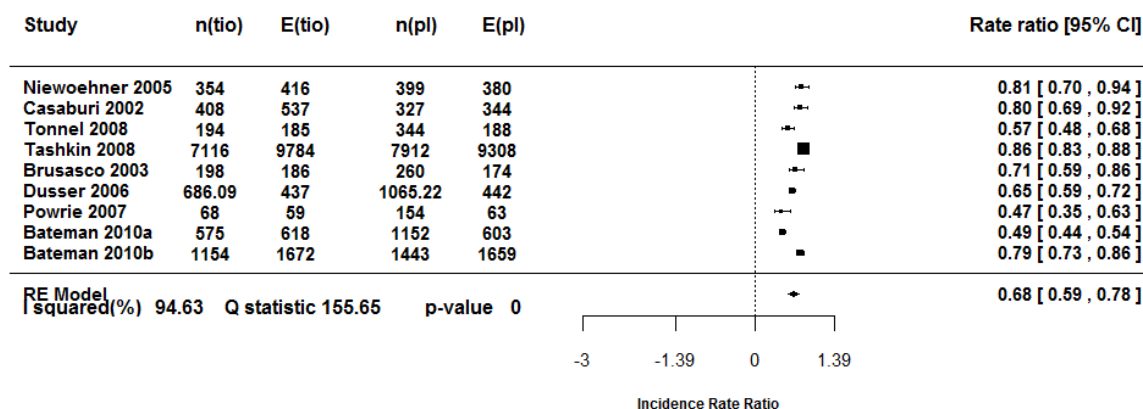


n: Anzahl der Ereignisse, E: Personenjahre, (ac): Acclidiniumbromid, (pl): Placebo

Abbildung 45: Sensitivitätsanalyse (Verzerrungspotential) für Exazerbationsraten aus RCT; Acclidiniumbromid im Vergleich zu Placebo.

Abbildung 45 zeigt, dass eine Behandlung mit Acclidiniumbromid im Vergleich zu Placebo eine stärkere Reduzierung der Exazerbationsraten bewirkt (Ratenverhältnis 0,64 [95%-KI 0,47; 0,87]). Wie aus den Konfidenzintervallen der Meta-Analyse ersichtlich ist, hatte der Ausschluss der ACCORD-II-Studie keine signifikanten Auswirkungen auf den gepoolten Effektschätzer. Der Heterogenitätstest ($Q=0,006$; $p=0,94$; $I^2=0,0\%$) ergab keine Hinweise auf signifikante Studienheterogenität.

Um zu untersuchen, ob die signifikante Heterogenität zwischen den Tiotropiumbromid-Studien auf das Verzerrungspotenzial zurückzuführen ist, wurden die Studien mit einem hohen Verzerrungspotenzial ausgeschlossen. Die Ergebnisse sind in Abbildung 46 dargestellt.



n: Anzahl der Ereignisse, E: Personenjahre, (tio): Tiotropiumbromid, (pl): Placebo

Abbildung 46: Sensitivitätsanalyse (Verzerrungspotential) für Exazerbationsraten aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.

Abbildung 46 zeigt, dass eine Behandlung mit Tiotropiumbromid im Vergleich zu Placebo eine stärkere Reduzierung der Exazerbationsrate bewirkt (Ratenverhältnis 0,68 [95%-KI 0,59; 0,78]). Der Ausschluss der Studie mit hohem Verzerrungspotenzial wirkte sich nicht signifikant auf den gepoolten Effektschätzer aus, wie die Konfidenzintervalle der Meta-Analyse zeigen. Der Heterogenitätstest ($Q=155,7$; $p<0,001$; $I^2=94,6\%$) lässt dennoch auf eine signifikante Studienheterogenität schließen.

Um die Ursache der Heterogenität genauer zu untersuchen, wurden die Studien nach Dosierung (5µg, 18 µg) bzw. Inhalationsgerät (Respimat, Handihaler) eingeteilt (Abbildung 47, Abbildung 48). Anschließend wurde pro Gruppe eine Meta-Analyse durchgeführt.

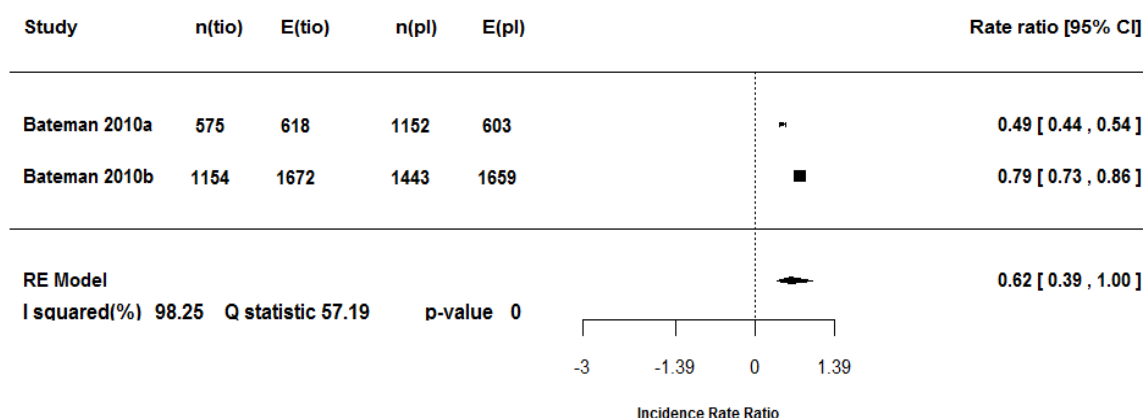


Abbildung 47: Sensitivitätsanalyse (Dosierung) für Exazerbationsraten aus RCT; Tiotropiumbromid 5µg im Vergleich zu Placebo.

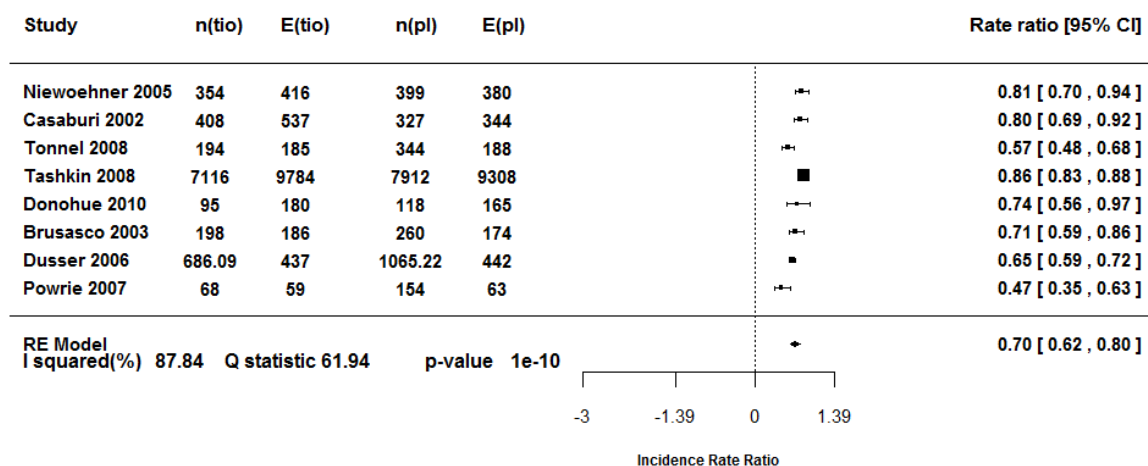


Abbildung 48: Sensitivitätsanalyse (Dosierung) für Exazerbationsraten aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.

Abbildung 47 und Abbildung 48 zeigen, dass die Behandlung mit Tiotropiumbromid 5µg sowie 18µg die Exazerbationsrate im Vergleich zu Placebo reduziert. Die separate Analyse der Tiotropiumbromid-5µg- und Tiotropiumbromid-18µg-Studien führte zu keinem signifikanten Unterschied hinsichtlich des gepoolten Effektschätzers, wie die Konfidenzintervalle der Meta-Analyse zeigen. Der Heterogenitätstest lässt trotz Separierung der Tiotropiumbromid-5µg- und Tiotropiumbromid-18µg-Studien auf eine signifikante Studienheterogenität in der Meta schließen.

Die Exazerbationsraten wurden nicht in allen Tiotropiumbromid-Studien gleich definiert. Sowohl dies, als auch die Tatsache, dass sich die Dauer der jeweiligen Studien (und somit auch die Wahrscheinlichkeit einer Exazerbation trotz Jahresbezug) unterscheiden, könnten die Heterogenität der Studien erklären. Der Gesamteffektschätzer kann dennoch als aussagefähig betrachtet werden, da alle Tiotropiumbromid-Studien die gleiche Effektrichtung aufweisen.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-121 zeigt die DIC-Schätzwerte für die geplanten NMAs. Die Definition von Exazerbationen unterscheidet sich je nach Studie und die Studiendauer. Daher empfiehlt sich

für die klinische Auswertung der Daten die Verwendung eines Modells mit zufälligen Effekten. Dieser Nachweis ist konsistent mit dem Befund der Meta-Analyse von Tiotropiumbromid 5 µg und 18µg, der eine signifikante Heterogenität ergab.

Tabelle 4.3-121: DIC für Exazerbationsraten – Indirekter Vergleich aus RCT

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung zum Modell
Ohne LABA-Begleitmedikation	nicht berechnet	174,004	Zufällige Effekte
Alle Studien	nicht berechnet	240,755	Zufällige Effekte

Die Ergebnisse der NMA werden in Tabelle 4.3-122 und Tabelle 4.3-123 dargestellt.

Tabelle 4.3-122: Exazerbationsraten - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne LABA-Begleitmedikation; Modell mit zufälligen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	0,62	
	95% KI	0,51	0,75
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	0,73	
	95% KI	0,48	1,07
	P(Überlegenheit)	95%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Aclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-123: Exazerbationsraten - Indirekter Vergleich aus RCT; Szenario 2: alle Studien; Modell mit zufälligen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	0,68	
	95% KI	0,58	0,79
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	0,73	
	95% KI	0,48	1,07
	P(Überlegenheit)	95%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Aclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Beide Szenarien der NMA (Tabelle 4.3-122 und Tabelle 4.3-123) ergeben, dass Aclidiniumbromid und Placebo gleichwertig sind (OR: 0,73 [KI 95% 0,48; 1,07]). Die obere Grenze des 95 %-Konfidenzintervalls liegt unmittelbar über der Eins und die Wahrscheinlichkeit, dass Aclidiniumbromid Placebo überlegen ist, beträgt 95% (Studien ohne LABA-Begleitmedikation, Tabelle 4.3-122). Dies deutet in die Richtung, die die Meta-Analyse vorgab. Diese ergab, dass Aclidiniumbromid effektiver ist als Placebo (OR: 0,71 [KI 95% 0,54; 0,94]). Die zu den relativen Effekten von Aclidiniumbromid im Vergleich zu Placebo ermittelten Konfidenzintervalle waren in der NMA etwas größer als in der Meta-Analyse, was zu den unterschiedlichen Schlussfolgerungen führte. Dies kann teilweise mit der Verwendung eines Modells mit zufälligen Effekten bei der NMA erklärt werden, welches aufgrund der heterogenen Tiotropiumbromid-Studien verwendet wurde. Dabei werden für alle paarweisen Vergleiche Heterogenitäten vorausgesetzt, obwohl die Heterogenität innerhalb der Aclidiniumbromid-Studien gering war (siehe Ergebnisse der Metaanalyse, Abbildung 43).

Gegenüber Tiotropiumbromid kann die Senkung der Exazerbationsraten durch Aclidiniumbromid gleichwertig angesehen werden [Odds Ratio 1,2 (95% KI: 0,7, 1,8)] (Tabelle 4.3-122).

Aclidiniumbromid zeigte bezüglich der Exazerbationsraten in der Metaanalyse einen statistisch signifikanten Vorteil, in der NMA gab es einen Trend zugunsten von Aclidiniumbromid gegenüber Placebo. Der Effekt von Aclidiniumbromid und Tiotropiumbromid auf die Exazerbationsrate war gleichwertig, da die Eins innerhalb des 95 %-Konfidenzintervalls lag. Aclidiniumbromid und Tiotropiumbromid haben einen gleichwertigen Nutzen in Bezug auf den Endpunkt Exazerbationsrate.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.9 Krankenhausaufenthalte aufgrund von Exazerbationen – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-124: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Krankenhausaufenthalte aufgrund von Exazerbationen – 12 Wochen: alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I ACCORD II	•			•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Zwei Studien berichteten über Krankenhausaufenthalte aufgrund von Exazerbationen für Aclidiniumbromid nach 12 Wochen, eine NMA war nicht durchführbar, da keine Tiotropiumbromid-Studien über diesen Endpunkt berichten.

Tabelle 4.3-125: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Krankenhausaufenthalte aufgrund von Exazerbationen – 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
2	Brusasco 2003 Vogelmeier 2008			• •	• •

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Drei Studien berichten über die Anzahl der Patienten mit Krankenhauseinweisungen aufgrund von Exazerbationen nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet); eine Acridiniumbromid-Studie und zwei Tiotropiumbromid-18µg-Studien. Keine Tiotropiumbromid-5µg-Studie berichtete über diesen Endpunkt. Abbildung 49 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 49: Netzwerk-Diagramm für die Analyse von Krankenhauseinweisungen aufgrund von Exazerbationen nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)



Tabelle 4.3-126: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Krankenhausaufenthalte aufgrund von Exazerbationen – 24 Wochen: alle Studien]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
3	Brusasco 2003			•	•
	Niewoehner 2005			•	•
	Vogelmeier 2008			•	•

Vier Studien berichteten über die Anzahl der Patienten mit Krankenhauseinweisungen aufgrund von Exazerbationen nach 24 Wochen; eine Acridiniumbromid-Studie und drei Tiotropiumbromid-18µg-Studien. Keine Tiotropiumbromid-5µg-Studie berichtete über diesen Endpunkt. Abbildung 50 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 50: Netzwerk-Diagramm für die Analyse von Krankenhauseinweisungen aufgrund von Exazerbationen nach 24 Wochen (alle Studien)



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-127: Operationalisierung von Krankenhauseinweisungen aufgrund von Exazerbationen

Studie	Operationalisierung
ACCORD I	Gesamtzahl der Tage im Krankenhaus, in der Notaufnahme und auf der Intensivstation wegen COPD-Exazerbationen. Zusätzliche Efficacy-Parameter in Bezug auf Krankenhausaufenthalte wegen COPD-Exazerbation wurden deskriptiv analysiert. Die Gesamtzahl der Tage von Krankenhausaufenthalt, Notaufnahme- und Intensivstations-Aufenthalte wegen COPD-Exazerbationen wurden für die ITT-Population und die Safety-Population bewertet.
ACCORD II	Die Anzahl der Tage von Krankenhausaufhalten (Notaufnahme oder der Intensivstation) aufgrund von COPD-Exazerbationen wurde als Summe der Krankenhausaufhaltstage (Notaufnahme oder der Intensivstation) aufgrund von COPD-Exazerbationen während der doppelblinden Behandlungsperiode berechnet.
ATTAIN	Ein Krankenhausaufenthalt wurde als eine Übernachtung im Krankenhaus oder der Notaufnahme definiert. Die Dauer der Krankenhausaufenthalte wurde mit Hilfe deskriptiver Statistik dargestellt.
Brusasco 2003	Die Informationen über Tage im Krankenhaus und die Nutzung der Intensivstation wurden zusammen mit Informationen über außerplanmäßige Arztuntersuchungen oder andere Gesundheitsdienstleister, Tage mit Behinderungs- und Beschäftigungsstatus erhoben. Es wurden Informationen über Krankenhausaufenthalte jeglicher Ursache und Exazerbationen, die eine stationäre Behandlung erforderten, erhoben.
Niewoehner 2005	Das Studienteam an jedem Standort forderte Arztbriefe für alle Krankenhausaufenthalte an, wo immer sie auftraten. Das Studienteam identifizierte Krankenhausaufenthalte aufgrund von COPD-Exazerbationen aus Ereignissen auf den Prüfbögen, die der Protokoll-Definition einer Exazerbation entsprachen und wo eine Überprüfung der Arztbriefe und anderer verfügbarer medizinischer Aufzeichnungen zeigten, dass die Ereignisse zu einem Krankenhausaufenthalt führten. Ein Ereignis wurde als Krankenhausaufenthalt gewertet, wenn ein Patient sich wegen einer akuten Atemwegserkrankung in einer Intensivpflege-Abteilung oder in einer Beobachtungseinheit für länger als 24 Stunden aufhielt und behandelt wurde. Einweisungen in Pflegeheime oder andere erweiterte Betreuungsangebote wurden nicht als Krankenhausaufenthalte berücksichtigt. Die COPD-Krankenhausaufenthalts-Endpunkte wurden mit einem Cochran-Mantel-Haenszel-Test analysiert, mit dem Studienzentrum als Stratum. In der Analyse wurde ein 2-seitiges α -Niveau von 0,05 verwendet.
Vogelmeier 2008	Ohne Angabe

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-128: Bewertung des Verzerrungspotenzials für Krankenhauseinweisungen aufgrund von Exazerbationen in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkt-erheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Unklar	Unklar	Ja	Hoch
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Brusasco 2003	Niedrig	Ja	Ja	Ja	Nein	Hoch
Niewoehner 2005	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Vogelmeier 2008	Hoch	Nein	Ja	Unklar	Nein	Hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die ACCORD I- und ACCORD II-Studie wurde das Verzerrungspotenzial für den Endpunkt Krankenhausaufenthalte aufgrund von Exazerbationen als hoch angesehen (Tabelle 4.3-128, Details sind in Anhang 4-G beschrieben). Eine von drei Tiotropiumbromid-Studien hatte auf der Studienebene ein hohes Verzerrungspotenzial. Bei der Betrachtung des Verzerrungspotenzials speziell für diesen Endpunkt hatten jedoch 2 von 3 Tiotropiumbromid-Studien ein hohes Verzerrungspotenzial (Details siehe Anhang 4-G).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-129: Ergebnisse für Patienten mit Krankenhausaufenthalten aufgrund von Exazerbation aus RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Gruppendifferenz (SE) [KI] p-Wert
ACCORD I	Aclidiniumbromid	190	2 (1,1%)	1,96 (1,23) [0,18; 21,77] p= 0,58
	Placebo	185	1 (0,5%)	
ACCORD II	Aclidiniumbromid	177	1 (0,6%)	0,34 (1,16) [0,03; 3,29] p= 0,35
	Placebo	182	3 (1,7%)	

N = Anzahl der Patienten, SE = Standardfehler, ITT = intention-to-treat population, KI = Konfidenzintervall

Tabelle 4.3-129 zeigt die Ergebnisse der Studien zur Untersuchung von Aclidiniumbromid mit Bezug auf Krankenhausaufenthalte aufgrund von Exazerbationen nach 12 Wochen. Die Tiotropiumbromid-Studien berichteten nicht über diesen Endpunkt. Die Behandlungseffekte unterschieden sich in den Studien: Aclidiniumbromid erhöhte das Risiko von Krankenhausaufenthalten in der ACCORD I-Studie und verringerten das Risiko in der ACCORD II-Studie. Die Zahl der Ereignisse war allerdings gering, was eine geringe statistische Aussagekraft und große Konfidenzintervalle ergab.

Tabelle 4.3-130: Ergebnisse für Patienten mit Krankenhausaufenthalten aufgrund von Exazerbation aus RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Gruppendifferenz (SE) [KI] p-Wert
ATTAIN	Aclidiniumbromid	269	2 (0,7%)	0,20 (0,78) [0,04; 0,91] p= 0,04
	Placebo	273	10 (3,7%)	
Vogelmeier 2008	Tiotropiumbromid 18µg	221	5 (2,3%)	1,59 (0,73) [0,38; 6,74] p=0,53
	Placebo	209	3 (1,44%)	
Niewoehner 2005	Tiotropiumbromid 18µg	914	64 (7,0%)	0,72 (0,17) [0,51; 1,00] p=0,06
	Placebo	915	87 (9,5%)	
Brusasco 2003	Tiotropiumbromid 18µg	386	12 (3,1%)	0,61 (0,38) [0,29; 1,29] p=0,19
	Placebo	362	18 (5,0%)	

N = Anzahl der Patienten, SE = Standardfehler, ITT = intention-to-treat population, KI = Konfidenzintervall

Tabelle 4.3-130 zeigt die Ergebnisse der Studien zur Untersuchung von Aclidiniumbromid und Tiotropiumbromid 18µg für Krankenhausaufenthalte aufgrund von Exazerbationen nach 24 Wochen. Die ATTAIN-Studie war die einzige Aclidiniumbromid-Studie, die über diesen Endpunkt berichtete, weshalb eine Meta-Analyse nicht möglich war. Die Behandlungseffekte variierten zwischen den Tiotropiumbromid-Studien: Eine Behandlung mit Tiotropiumbromid erhöhte in der Studie von Vogelmeier 2008 das Risiko einer Krankenseinweisung und verringerte das Risiko in den Studien von Niewoehner 2005 und Brusasco 2003. Die Zahl der Ereignisse war in der Studie von Vogelmeier 2008 allerdings gering, was eine geringe statistische Aussagekraft und große Konfidenzintervalle ergab.

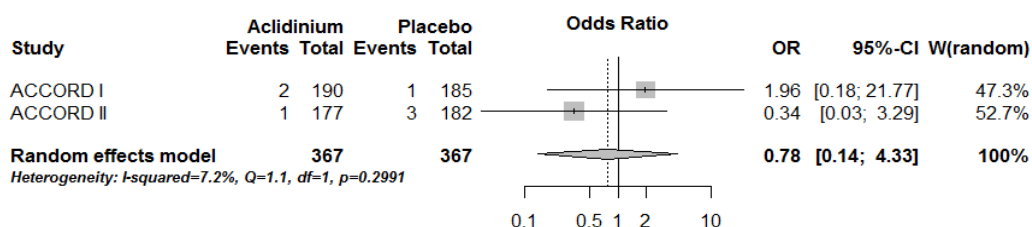


Abbildung 51: Meta-Analyse für die Anzahl von Patienten, die aufgrund von unerwünschten Ereignissen nach 12 Wochen in ein Krankenhaus eingeliefert wurden; Aclidiniumbromid im Vergleich zu Placebo

Abbildung 51 zeigt, dass Acclidiniumbromid für diesen Endpunkt nach 12 Wochen mit Placebo vergleichbar ist (gepoolte OR 0,78 [KI 95% 0,14; 4,33]). Dieses Ergebnis sollte aufgrund der geringen Zahl von Ereignissen jedoch mit Vorsicht interpretiert werden. Die geringe Anzahl von Ereignissen spiegelt sich auch in den großen Konfidenzintervallen wider. Es lag keine Studien-Heterogenität vor ($Q=1,1$; $p=0,30$; $I^2=7,2\%$).

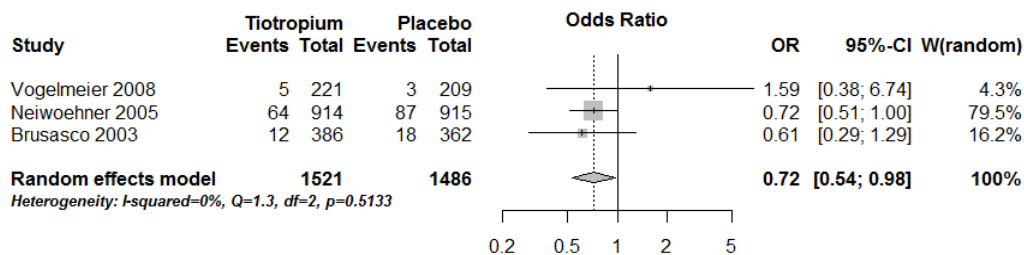


Abbildung 52: Meta-Analyse für die Anzahl von Patienten, die aufgrund von Exazerbationen in ein Krankenhaus eingeliefert wurden (24 Wochen); Tiotropiumbromid 18µg im Vergleich zu Placebo

Abbildung 52 zeigt, dass Tiotropiumbromid mit weniger Krankenhauseinweisungen nach 24 Wochen als Placebo assoziiert ist (gepoolte OR 0,72 [KI 95% 0,14; 0,98]). Diese Ergebnisse sollten mit Vorsicht interpretiert werden, da nur eine von drei Studien ein geringes Verzerrungspotenzial für dieses Outcome hatte. Das dem Behandlungseffekt in der Studie von Vogelmeier 2008 entsprechende Konfidenzintervall war sehr groß, was auf die geringe Anzahl der Ereignisse in dieser Studie zurückzuführen ist. Signifikante Heterogenität der Studien gab es nicht ($Q=1,3$; $p=0,51$; $I^2=0,0\%$).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-131 zeigt die DIC-Schätzungen für die geplanten NMAs. Nach 24 Wochen war das beste durch DIC-Ergebnisse ermittelte Modell ein Modell mit festen Effekten in Szenario 1 und Szenario 2. Dies stimmt mit den Ergebnissen der Meta-Analyse für Tiotropiumbromid 18µg überein (Abbildung 52); hier ließ der Heterogenitätstest auf keine signifikante Heterogenität schließen. Es wurde keine NMA nach 12 Wochen durchgeführt, weil nur zwei Acclidiniumbromid-Studien Daten über Krankenhausaufenthalte aufgrund von Exazerbationen nach 12 Wochen berichteten.

Tabelle 4.3-131: DIC des indirekten Vergleichs von Acclidiniumbromid und Tiotropiumbromid 18µg; Krankenhausaufenthalte aufgrund von Exazerbationen

Studienpool	Feste Effekte	Zufällige Effekte	Schlussfolgerung für die NMA-Modelle
Woche 24			
Ohne LABA-Begleitmedikation	33,6	34,0	Feste Effekte
Alle Studien	47,7	49,3	Feste Effekte

Die Ergebnisse der NMA werden in Tabelle 4.3-132 und Tabelle 4.3-133 dargestellt.

Tabelle 4.3-132: Krankenhausaufenthalte aufgrund von Exazerbationen (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 18
TIO 18	Effektschätzer	0,79	
	95% KI	0,38	1,45
	P(Überlegenheit)	81%	
AB 400	Effektschätzer	0,22	
	95% KI	0,02	0,70
	P(Überlegenheit)	99%	
		0,31	
		0,03	1,10
		97%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-133: Krankenhausaufenthalte aufgrund von Exazerbationen (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: alle Studien; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 18
TIO 18	Effektschätzer	0,73	
	95% KI	0,53	0,97
	P(Überlegenheit)	98%	
AB 400	Effektschätzer	0,22	
	95% KI	0,02	0,70
	P(Überlegenheit)	99%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen verringerte Acclidiniumbromid im Vergleich zu Placebo die Anzahl der Patienten, die aufgrund von Exazerbationen ins Krankenhaus eingewiesen wurden signifikant [OR 0,2 (95% KI: 0,03, 0,7)]. Bezüglich Tiotropiumbromid lässt das Ergebnis der NMA einen starken Trend zugunsten von Acclidiniumbromid erkennen [OR 0,31 (95% KI: 0,03, 1,1)] (Tabelle 4.3-132). Szenario 2 bestätigt diesen Trend noch etwas stärker [OR 0,31 (95% KI: 0,03, 1,0)].

Beide Therapien, Acclidiniumbromid und Tiotropiumbromid senkten die Hospitalisationsrate wegen Exazerbationen nach 24 Wochen im Vergleich zu Placebo signifikant. Der indirekte Vergleich ergab eine Gleichwertigkeit von Acclidiniumbromid und Tiotropiumbromid, bei dem eine Tendenz zur Überlegenheit von Acclidiniumbromid zu erkennen ist – die Eins lag nur knapp innerhalb des 95 %-Konfidenzintervalls.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.10 SGRQ-Gesamtscore – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-134: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Gesamtwert– 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
3	INHANCE			•	•
	TIPHON			•	•
	Verkindre 2006			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sechs Studien berichteten über die Veränderung des SGRQ-Gesamtwerts (begleitende Behandlung mit LABAs nicht gestattet) nach 12 Wochen; alle drei Aclidiniumbromid-Studien (ACCORD I, ACCORD II und ATTAIN) und drei Tiotropiumbromid-18µg-Studien (INHANCE, TIPHON, Verkindre, 2006). In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 53 zeigt das Netzwerk, das basierend auf diesen drei Studien erzeugt werden kann.

Abbildung 53: Netzwerkdiagramm für die Analyse des SGRQ-Gesamtwerts nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

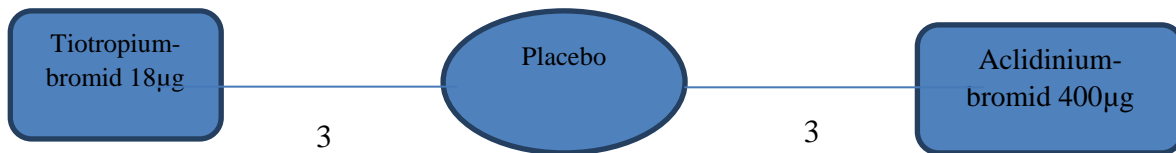


Tabelle 4.3-135: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Gesamtwert – 12 Wochen – alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
3	INHANCE			•	•
	TIPHON			•	•
	Verkindre 2006			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sechs Studien berichteten über die Veränderung des SGRQ-Gesamtwerts nach 12 Wochen; alle drei Aclidiniumbromid-Studien und drei Tiotropiumbromid-18µg-Studien (INHANCE, TIPHON, Verkindre 2006). In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 54 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Abbildung 54: Netzwerkdiagramm für die Analyse des SGRQ-Gesamtwerts nach 12 Wochen (alle Studien)

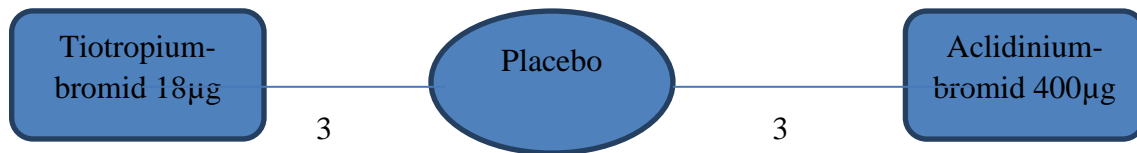


Tabelle 4.3-136: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Gesamtwert – 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
6	Brusasco 2003			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	Donohue 2002			•	•
	TIPHON			•	•
	Vogelmeier 2008			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sieben Studien berichteten über die Veränderung des SGRQ-Gesamtwerts (begleitende Behandlung mit LABAs nicht gestattet) nach 24 Wochen; eine Acridiniumbromid-Studie (ATTAIN) und sechs Tiotropiumbromid-18µg-Studien. In keiner der Tiotropiumbromid-5µg-Studien wurde dieser Endpunkt berichtet. Abbildung 55 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 55: Netzwerkdiagramm für die Analyse des SGRQ-Gesamtwerts nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

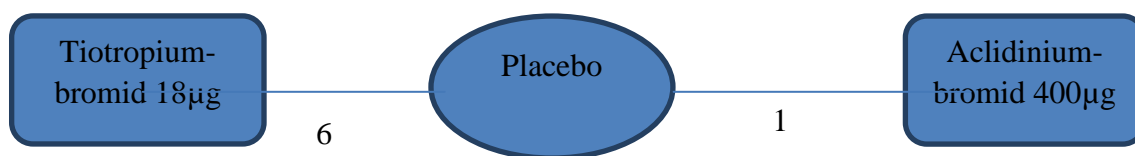


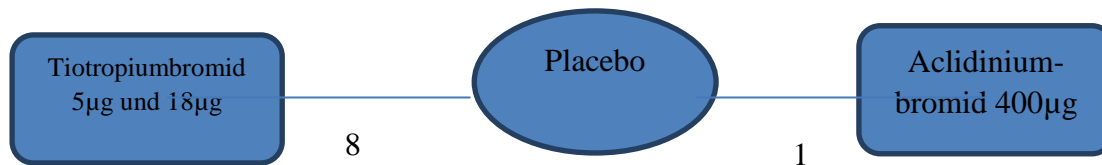
Tabelle 4.3-137: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Gesamtwert – 24 Wochen – alle Studien]

Anzahl Studien	Referenzstudien	Acclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
7	Brusasco 2003			•	•
	Casaburi 2002			•	•
	INHANCE			•	•
	Donohue 2002			•	•
	UPLIFT			•	•
	TIPHON			•	•
	Vogelmeier 2008			•	•
1	Bateman 2010b		•		•

Neun Studien berichteten über die Veränderung des SGRQ-Gesamtwerts nach 24 Wochen; eine Acclidiniumbromid-Studie (ATTAIN), sieben Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie (Bateman 2010b).

Abbildung 56 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 56: Netzwerkdiagramm für die Analyse des SGRQ-Gesamtwerts nach 24 Wochen (alle Studien)



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-138: Operationalisierung des SGRQ-Gesamtwerts

Studie	Operationalisierung
ACCORD I	Der SGRQ bietet ein objektives Lebensqualitätsmaß für Patienten mit COPD. Der Patient füllte den SGRQ nach 12 Wochen aus. Alle Bereiche des SGRQ-Fragebogens wurden ausgefüllt.
ACCORD II	Der SGRQ bietet ein objektives Lebensqualitätsmaß für Patienten mit COPD. Es ist ein sensitiver, zuverlässiger und validierter, aus 2 Teilen bestehender Fragebogen: einer Symptomauswertung und einer Auswertung der Aktivität und Auswirkungen. Die Patienten mussten den Fragebogen bei den Visiten 2, 4, 5, 6 (12 Wochen, Ende von Teil 1 der Studie), 7, 8, 9 und 10 (40 Wochen, Ende von Teil 2 der Studie) oder bei vorzeitiger Beendigung vervollständigen. Alle Bereiche des SGRQ-Fragebogens wurden ausgefüllt.
ATTAIN	Der krankheitsspezifische Gesundheitszustand wurde mittels eines selbst zu bedienenden Instrumentes, des SGRQ, bewertet. Dieser Fragebogen ist ein standardisiertes selbst vom Patienten auszufüllendes Instrument zur Messung von Gesundheitsstörungen und wahrgenommenem Wohlbefinden ("QoL") bei Erkrankungen der Atemwege. Die relevante gültige Sprachversion des Fragebogens wurde dem Patienten während der Visite 2 (Bewertung der Ausgangswerte), Visite 4 (Woche 4), Visite 6 (Woche 12) und Visite 8 (Woche 24) zur Verfügung gestellt. Bei den letzten Visiten wurden den Patienten keine Informationen zu den Ausgangswert-Fragebogen (Visite 2) zur Verfügung gestellt.
Brusasco 2003	Die HRQoL wurde mit dem St. Georgs-Respiratory Fragebogen (SGRQ) bestimmt. Der SGRQ ist ein krankheitsspezifisches Instrument, das 50 Items in drei Subskalen (Symptome, Aktivität und Auswirkungen) enthält. Jede Antwort hat ein empirisch abgeleitetes Gewicht. Die Gesamtnote wird aus den Antworten aller 50 Items berechnet. Fehlende individuelle Fragen wurden nach den Richtlinien des Entwicklers des Fragebogens hinzugefügt. Die Daten aus fehlenden SGRQ Fragebögen wurden unter Verwendung der zuletzt durchgeführten Beobachtung übertragen.

INHANCE	Der Gesundheitszustand wurde mit dem St. George's Respiratory Fragebogen (SGRQ) gemessen. Der SGRQ wurde von den Patienten bei den Visiten an Tag 1 und in den Wochen 4, 8, 12 und 26 ausgefüllt.
Donohue 2002	Die HRQoL wurde mit dem St. George's Respiratory Fragebogen (SGRQ) gemessen. Der SGRQ ist ein krankheitsspezifisches Instrument, das 50 Items in drei Subskalen (Symptome, Aktivität und Auswirkungen) enthält. Die Gesamtnote kann aus den Antworten aus allen 50 Items berechnet werden. Ein niedrigerer Wert bedeutet eine Verbesserung.
UPLIFT	Gesundheitsbezogene Lebensqualität wurde mit dem SGRQ vor der Präbronchodilator-Spirometrie-Untersuchung bei Studienbeginn und alle 6 Monate gemessen. Die Werte auf der SGRQ reichen von 0 bis 100, wobei niedrigere Werte eine Verbesserung angeben und eine Änderung von 4 oder mehr Einheiten als klinisch aussagekräftig angesehen werden; SGRQ Daten aus der Türkei wurden wegen fehlerhafter Validierung des Fragebogens ausgeschlossen.
TIPHON	Der SGRQ besteht aus einem 76-Item-Fragebogen, der in 3 Komponenten unterteilt ist: Symptome (zur Beurteilung der Behinderung wegen Atemwegsbeschwerden), Aktivität (Beurteilung der Auswirkungen der Atemlosigkeit auf Mobilität und körperliche Aktivität) und Auswirkungen (Beurteilung der psychosozialen Auswirkungen der Erkrankung). Die SGRQ-Gesamtpunktzahl setzt sich aus 3 Teilwerten zusammen. Die Werte werden gewichtet und reichen von 0 bis 100, wobei höhere Werte eine schlechtere gesundheitsbezogene Lebensqualität angeben.
Verkindre 2006	Die HRQoL wurde mit dem St. George's Respiratory Fragebogen (SGRQ) vor der Behandlung an den Tagen 1, 42 und 84 des Behandlungszeitraumes gemessen. Der SGRQ besteht aus einem 76-Item-Fragebogen, der in 3 Komponenten unterteilt ist: Symptome zur Beurteilung der Behinderung wegen Atemwegsbeschwerden; Aktivität zur Beurteilung der Auswirkungen der Atemlosigkeit auf Mobilität und körperliche Aktivität und Auswirkungen zur Beurteilung der psychosozialen Auswirkungen der Erkrankung. Die SGRQ-Gesamtpunktzahl besteht aus allen 3 Teilwerten. Die Werte werden gewichtet und reichen von 0 bis 100, wobei höhere Werte eine schlechtere gesundheitsbezogene Lebensqualität angeben.
Bateman 2010b	Die HRQoL wurde mit dem St. George's Respiratory Fragebogen (SGRQ) gemessen.
Vogelmeier 2008	Die krankheitsspezifische Lebensqualität (QoL) wurde mit dem St.-Georgs-Respiratory Fragebogen (SGRQ) beurteilt. Der SGRQ wurde zu Studienbeginn und nach 12 und 24 Behandlungswochen eingesetzt. Negative Werte weisen auf eine verbesserte Lebensqualität im Vergleich zu Placebo hin.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-139: Bewertung des Verzerrungspotenzials für den SGRQ-Gesamtwert in RCT mit dem zu bewertenden und zu vergleichenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Brusasco 2003	Niedrig	Ja	Unklar	Unklar	Nein	Niedrig
INHANCE	Niedrig	Nein	Nein	Nein	Nein	Hoch
Donohue 2002	Niedrig	Ja	Nein	Nein	Nein	Hoch
UPLIFT	Niedrig	Ja	Nein	Nein	Nein	Hoch
TIPHON	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Verkindre 2006	Niedrig	Ja	Nein	Nein	Nein	Hoch
Vogelmeier 2008	Hoch	Nein	Ja	Unklar	Nein	Hoch
Bateman 2010b	Niedrig	Ja	Nein	Unklar	Ja	Hoch
Bateman 2010a	Niedrig	Ja	Unklar	Nein	Ja	Hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei der ACCORD II-Studie wurde das Verzerrungspotenzial für den SGRQ-Score als hoch eingestuft. Obwohl beim Randomisierungsprozess und bei der Behandlungsallokation dieser Studie bewährte Methoden angewandt wurden, war die Vergleichbarkeit der Patientencharakteristika in beiden Studienarmen zu Studienbeginn eingeschränkt.

Auf Studienebene wiesen alle Tiotropiumbromid-Studien mit Ausnahme der Studie von Vogelmeier 2008 ein geringes Verzerrungspotenzial auf. Auf Ebene des SGRQ-Scores wiesen 6 von 8 Tiotropiumbromid-Studien ein hohes Verzerrungspotenzial auf. Dies ist darauf zurückzuführen, dass entweder das ITT-Prinzip nicht für die Outcome-Analyse angewendet oder die SGRQ-Messung für die Behandlungsallokation nicht verblindet wurde (Tabelle 4.3-139). Eine Studie (Bateman 2010b) gab außerdem nicht die Anzahl der in die Analyse eingeschlossenen Patienten an. (Detaillierte Erklärung im Anhang 4-G).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-140: Ergebnisse für SGRQ-Gesamtwert (12 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mittelwert (SD)	Mittelwert (SD)	CFB (SD)	Gruppendifferenz (SE) [KI] p-Wert
ACCORD I	Aclidiniumbromid	190	48 (17,8)	43,17(-)	-4,83 (10,3)	-2,97 (1,13)
	Placebo	185	45 (16,2)	43,14(-)	-1,86 (10,2)	[-5,05; -0,89] p = 0,019
ACCORD II	Aclidiniumbromid	177	50,4 (16,9)	45(-)	-5,40 (13,3)	-1,10 (1,02)
	Placebo	182	49,2 (17,4)	44,9(-)	-4,30 (13,5)	[-3,87; 1,67] p = 0,429
ATTAIN	Aclidiniumbromid	269	47,4 (18)	40,95(17,2)	-6,45 (11,8)	-4,09 (1,41)
	Placebo	273	44,9 (16,5)	42,54(15,7)	-2,36 (11,9)	[-6,09; -2,09] p < 0,001
INHANCE (Donohue 2010)	Tiotropiumbromid 18µg	415	(-)	39,40 (-)	(-)	-1,10 (0,85)
	Placebo	418	(-)	40,40 (16,2)	(-)	[-2,79; 0,59] p = 0,436
TIPHON (Tonnel 2008)	Tiotropiumbromid 18µg	247	45,8 (17,7)	(-)	-6,97 (6,3)	-3,51 (1,3)
	Placebo	245	48,9 (18,4)	(-)	-3,46 (5,4)	[-6,01; -1,01] p = 0,006
Verkindre 2006	Tiotropiumbromid 18µg	46	48,7 (-)	41,2 (2,4)	(-)	-6,5 (2,8)
	Placebo	54	48,7 (-)	47,7 (2,3)	(-)	[-12,18; -0,82] p = 0,05

N=Anzahl der Patienten, CFB=Veränderung gegenüber dem Ausgangswert, SE=Standardfehler SD=Standardabweichung, ITT=Intention-to-treat-Population, KI=Konfidenzintervall

Tabelle 4.3-140 zeigt den SGRQ-Gesamtwert nach 12 Wochen für die Studien Aclidiniumbromid (400µg; BID) und Tiotropiumbromid (18 µg; OD), welche diesen Endpunkt berichten. In allen Studien zeigten die mit Aclidiniumbromid behandelten Patienten eine Verbesserung der Lebensqualität (d. h. einen niedrigeren SGRQ-Score). In ACCORD I und ATTAIN war die Verbesserung gegenüber Placebo signifikant.

Tabelle 4.3-141: Ergebnisse für SGRQ (24 Wochen) - Gesamtwert aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten	Baseline	Follow-up	Veränderung	
		(ITT)	Mittelwert (SD)	Mittelwert (SD)	CFB (SD)	Gruppendifferenz (SE) [KI] p-Wert
		N				
ATTAIN	Acridiniumbromid	269	47,4 (18,4)	40,4 (18,1)	-7,41 (13,4)	-4,29 (1,08) [-6,42; -2,16] p < 0,001
	Placebo	273	44,9 (16,7)	42,8 (17,3)	-2,79 (13,5)	
Bateman 2010b	Tiotropiumbromid 5µg	1989	(-)	(-)	-4,8 (-)	-2,20 (0,46) [-1,30; -3,10] p < 0,001
	Placebo	2002	(-)	(-)	-2,6 (-)	
Brusasco 2003	Tiotropiumbromid 18µg	386	(-)	(-)	4,2 (13,8)	-2,70 (1,0) [-4,64; -0,76] p = 0,006
	Placebo	362	(-)	(-)	1,5 (13,3)	
Casaburi 2002	Tiotropiumbromid 18µg	507	(-)	(-)	-3,53 (-)	-3,08 (0,93) [-4,94; -1,22] p = 0,001
	Placebo	332	(-)	(-)	-0,45 (-)	
Donohue 2002	Tiotropiumbromid 18µg	209	45,4 (-)	(-)	-5,14 (-)	-2,71 (1,6) [-5,39; -0,03] p = 0,05
	Placebo	201	45,4 (-)	(-)	-2,43 (-)	
INHANCE (Donohue 2010)	Tiotropiumbromid 18µg	415	(-)	39,40 (-)	(-)	-1,0 (0,9) [-2,8; 0,8] p = 0,276
	Placebo	418	(-)	40,40 (16,2)	(-)	
TIPHON (Tonnel 2008)	Tiotropiumbromid 18µg	247	45,8 (17,7)	(-)	-7,77 (6,8)	-3,51 (1,3) [-6,02; -1,0] p = 0,006
	Placebo	245	48,9 (18,4)	(-)	-4,26 (7,1)	
UPLIFT (Tashkin 2008)	Tiotropiumbromid 18µg	2987	45,7 (17,0)	42,22 (10,9)	(-)	-2,86 (0,4) [-3,68; -2,03] p < 0,0001
	Placebo	3006	46,0 (17,2)	39,37 (8,2)	(-)	
Vogelmeier 2008	Tiotropiumbromid 18µg	221	(-)	(-)	(-)	-2,09 (1,2) [-4,58; 0,40] p = 0,100
	Placebo	209	(-)	(-)	(-)	

N=Anzahl der Patienten, CFB=Veränderung gegenüber dem Ausgangswert, SE=Standardfehler SD=Standardabweichung, ITT=Intention-to-treat-Population, KI=Konfidenzintervall

Die Tabelle 4.3-141 zeigt den Endpunkt SGRQ-Gesamtscore nach 24 Wochen für alle Acridinium- und Tiotropiumstudien, die diesen Endpunkt berichteten. Mit einer Ausnahme (Vogelmeier 2008) wurde in allen Studien der SGRQ-Gesamtscore im Vergleich zu Placebo signifikant gesenkt. Da nur eine Acridiniumbromidstudie den Endpunkt SGRQ-Gesamtscore nach 24 Wochen untersuchte, wurde für diesen Endpunkt nach 24 Wochen mit Acridiniumbromid keine Metaanalyse durchgeführt.

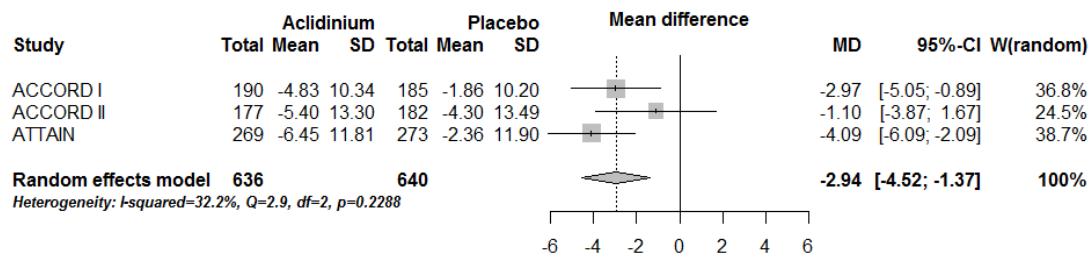


Abbildung 57: Meta-Analyse für den SGRQ-Score (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Die Metaanalyse in Abbildung 57 zeigt, dass Acclidiniumbromid den Endpunkt SGRQ-Gesamtscore nach 12 Wochen stärker reduziert als Placebo (Differenz -2,94 [95 % CI: -4,52; -1,37]). Eine signifikante Studienheterogenität existierte nicht ($Q = 2,9$, $p = 0,23$, $I^2 = 32,2\%$).

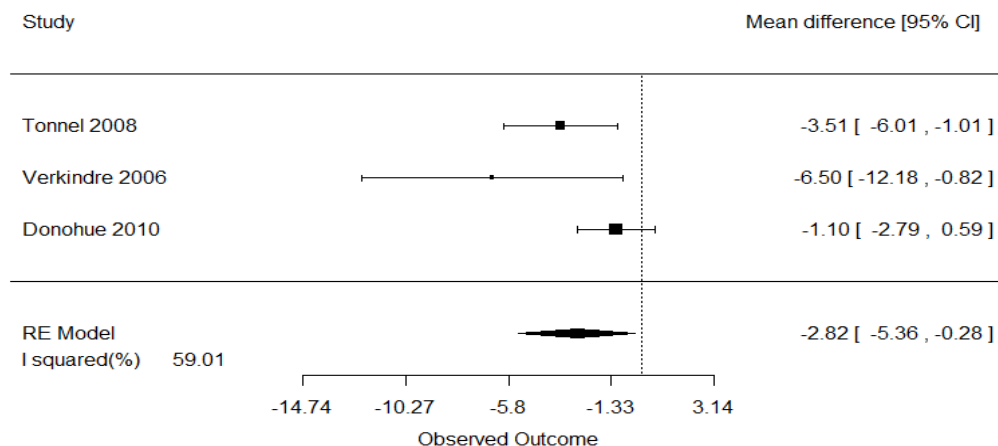


Abbildung 58: Meta-Analyse für SGRQ-Score (12 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo

Die Metaanalyse (Abbildung 58) zeigt, dass Tiotropiumbromid 18 µg den Endpunkt SGRQ-Gesamtscore nach 12 Wochen stärker senkt als Placebo (mittlere Differenz: -2,82 [CI 95%: -5,36; -0,28]). Eine signifikante Heterogenität existierte nicht ($Q = 4,88$, $p = 0,09$, $I^2 = 59,0\%$).

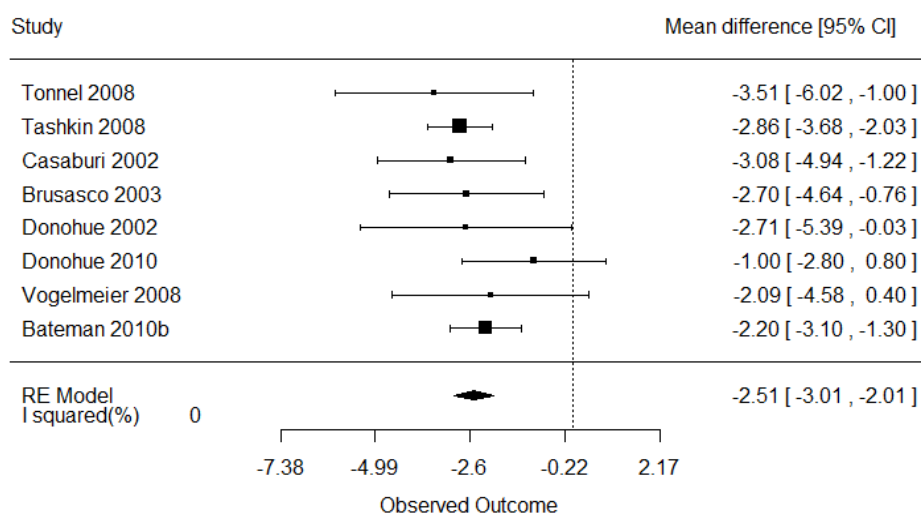


Abbildung 59: Meta-Analyse für den SGRQ-Score (24 Wochen) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo.

Tiotropiumbromid reduziert den Endpunkt SGRQ-Gesamtscore nach 24 Wochen stärker als Placebo (mittlere Differenz: -2,51 [CI 95 % -3,01; -2,01]. Eine signifikante Heterogenität der Studien lag nicht vor ($Q = 5,0$, $p = 0,66$, $I^2 = 0\%$).

Sensitivitätsanalyse

Die ACCORD-II-Studie wies hinsichtlich der Baseline-Patientencharakteristika ein Ungleichgewicht zwischen den Behandlungsgruppen auf. Um die Auswirkungen des Ausschlusses dieser Studie aus der Meta-Analyse zu analysieren, haben wir eine Sensitivitätsanalyse durchgeführt (Abbildung 60).

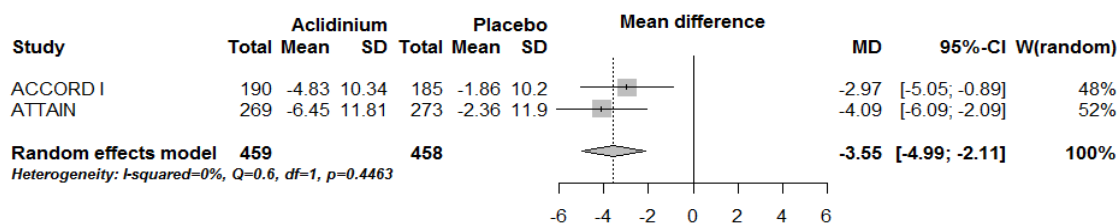


Abbildung 60: Sensitivitätsanalyse (Verzerrungspotential) für SGRQ-Gesamtwert nach 12 Wochen (ml) aus RCT; Acclidiniumbromid im Vergleich zu Placebo.

Acidiniumbromid (Abbildung 60) senkte den SGRQ-Gesamtscore nach 12 Wochen effektiver als Placebo. (mittlere Differenz -3,55 [95 % CI: -4;99; -2,11]). Der Ausschluss der ACCORD II-Studie beeinflusste die kombinierten Effektschätzwerte nicht signifikant, wie die Konfidenzintervalle der Metaanalyse zeigen. Eine signifikante Heterogenität lag nicht vor ($Q = 0,8$, $p < 0,38$, $I^2 = 0,0$ %).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-142 zeigt die DIC-Schätzwerte für die geplanten NMAs. Für den Endpunkt SGRQ-Gesamtscore ergaben sich nach 12 Wochen für die Random-Effekt-Analyse die besten DIC-Schätzwerte. Die Ergebnisse des Heterogenitätstests der Meta-Analyse ($p = 0,09$) sprechen ebenfalls dafür. Szenario 2 (alle Studien) wurde nach 12 Wochen nicht durchgeführt, da die Patienten in allen Studien, die Daten zu dem SGRQ-Gesamtwert nach 12 Wochen berichteten, keine Begleitmedikation mit LABA einsetzten. Nach 24 Wochen ergaben sich für die Fixed-Effekt-Analyse sowohl in Szenario 1 (ohne begleitende LABA-Therapie) als auch in Szenario 2 (alle Studien) die besten DIC-Schätzwerte (Tabelle 4.3-142). Das negative Ergebnis des Heterogenitätstests der Meta-Analyse bestätigt diese Entscheidung.

Tabelle 4.3-142: DIC des SGRQ-Gesamtwerts – indirekte Vergleich aus RCT

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung für die NMA-Modelle
Woche 12			
Ohne LABA-Begleitmedikation	29,779	28,644	Zufällige Effekte
Alle Studien	-	-	-
Woche 24			
Ohne LABA-Begleitmedikation	25,8	27,2	Feste Effekte
Alle Studien	27,395	28,803	Feste Effekte

Die Ergebnisse der Netzwerk-Meta-Analyse sind in Tabelle 4.3-143, Tabelle 4.3-144 und Tabelle 4.3-145 aufgeführt.

Tabelle 4.3-143: SGRQ-Gesamtwert (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit zufälligen Effekten

Intervention		Kontrolle			
		Pbo		TIO 18	
TIO 18	Effektschätzer	-2,81			
	95% KI	-6,36	-0,01		
	P(Überlegenheit)	98%			
AB 400	Effektschätzer	-2,60		0,22	
	95% KI	-5,62	0,30	-3,85	4,85
	P(Überlegenheit)	97%		48%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg;
Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Acclidiniumbromid reduzierte in Szenario 1 (ohne begleitende LABA-Therapie) den SGRQ-Gesamtscore in vergleichbarer Weise wie Placebo [Differenz: -2.6 Punkte (95 %-KI: -5.6, 0.3)] bei hoher Wahrscheinlichkeit einer Überlegenheit gegenüber Placebo (97 %). Dies ist zu den Ergebnissen der Meta-Analyse konsistent, in der Acclidiniumbromid gegenüber Placebo überlegen war. Der paarweise Vergleich von Acclidiniumbromid mit Tiotropiumbromid ergab eine Gleichwertigkeit der beiden Therapien [Differenz 0,2 Punkte (95 %-KI: -3,9, 4,9)] (Tabelle 4.3-143).

Tabelle 4.3-144: SGRQ-Gesamtwert (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle			
		Pbo		TIO 18	
TIO 18	Effektschätzer	-2,39			
	95% KI	-3,26	-1,52		
	P(Überlegenheit)	>99%			
AB 400	Effektschätzer	-4,29		-1,90	
	95% KI	-6,41	-2,18	-4,17	0,38
	P(Überlegenheit)	>99%		95%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg;
Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-145: SGRQ-Gesamtwert (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: Alle Studien; Modell mit festen Effekten

Intervention		Kontrolle			
		Pbo		TIO 5 und 18	
TIO 5 und 18	Effektschätzer	-2,52			
	95% KI	-3,01	-2,03		
	P(Überlegenheit)	>99%			
AB 400	Effektschätzer	-4,29		-1,77	
	95% KI	-6,41	-2,18	-3,94	0,40
	P(Überlegenheit)	>99%		95%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Acridiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen war Acridiniumbromid beim Endpunkt SGRQ-Gesamtscore effektiver als Placebo (mittlere Differenz: -4.3 Punkte (95 %-KI: -6,4, -2,2) und gleichwertig zu Tiotropiumbromid [mittlere Differenz: -1,9 Punkte (95 %-KI: -4,2, 0,4)] (Tabelle 4.3-144). Die Ergebnisse von Szenario 2 (alle Studien; Tabelle 4.3-145) bestätigen die Ergebnisse von Szenario 1 und zeigen ebenfalls eine gleichwertige Effektivität von Acridiniumbromid und Tiotropiumbromid bezüglich der Lebensqualität nach 24 Wochen.

Sowohl Acridiniumbromid als auch Tiotropiumbromid führten zu einer statistisch signifikant besseren Lebensqualität (gemessen als SGRQ) als Placebo. Der indirekte Vergleich zeigte, dass beide Substanzen die Lebensqualität in gleicher Weise gegenüber Placebo verbessern. Sowohl nach 12 als auch nach 24 Wochen schloss das 95 %-Konfidenzintervall die Null ein, d. h. bei der Wirkung auf den SGRQ ergaben sich keine statistisch signifikanten Unterschiede. Der Nutzen von Acridiniumbromid und Tiotropiumbromid auf die Lebensqualität ist demnach gleichwertig.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.11 SGRQ-Responder – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-146: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Responder (%) – 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
2	TIPHON			•	•
	Verkindre 2006			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Fünf Studien berichteten über die Anzahl der SGRQ-Responder nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet; SGRQ-Response: Verbesserung des SGRQ-Gesamtwerts ≥ 4 Punkte); alle drei Acridiniumbromid-Studien und zwei Tiotropiumbromid-18µg-Studien (TIPHON, Verkindre 2006). In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 61 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 61: Netzwerkdiagramm für die Analyse der SGRQ-Responder (%) nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

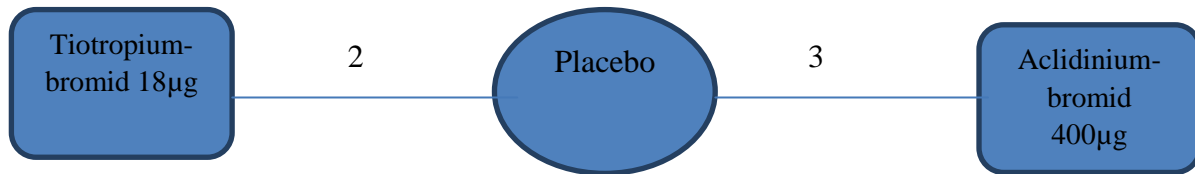


Tabelle 4.3-147: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Responder (%) – 12 Wochen – alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
3	ACCORD I	•			•
	ACCORD II	•			•
	ATTAIN	•			•
2	TIPHON			•	•
	Verkindre 2006			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Fünf Studien berichteten über die Anzahl der SGRQ-Responder nach 12 Wochen (Patienten mit einer Verbesserung des SGRQ-Gesamtwerts ≥ 4 Punkten); alle drei Aclidiniumbromid-Studien und zwei Tiotropiumbromid-18µg-Studien (TIPHON, Verkindre 2006). In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 62 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 62: Netzwerkdiagramm für die Analyse der SGRQ-Responder (%) nach 12 Wochen (alle Studien)



Tabelle 4.3-148: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Responder (%) – 24 Wochen – begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
2	Brusasco 2003			•	•
	Donohue 2002			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Drei Studien berichteten über die Anzahl der SGRQ-Responder nach 24 Wochen (Begleitende Behandlung mit LABAs nicht gestattet; Patienten mit einer Verbesserung des SGRQ-Gesamtwerts von ≥ 4 Punkten); eine Acridiniumbromid-Studie (ATTAIN) und zwei Tiotropiumbromid-18µg-Studien (Brusasco 2003, Donohue 2002). In keiner der Tiotropiumbromid-5µg-Studien wurde dieses Outcome berichtet. Abbildung 62 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Da in keiner der Studien eine Begleitmedikation mit LABA zugelassen war, sind Szenario 1 und Szenario 2 für diesen Endpunkt identisch.

Abbildung 63: Netzwerkdiagramm für die Analyse der SGRQ-Responder (%) nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)



Tabelle 4.3-149: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [SGRQ-Responder (%) – 24 Wochen - alle Studien]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
2	Brusasco 2003 Donohue 2002			• •	• •
1	Bateman 2010b		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Vier Studien berichteten über die Anzahl der SGRQ-Responder nach 24 Wochen (Verbesserung des SGRQ-Gesamtwerts ≥ 4 Punkte); eine Acridiniumbromid-Studie (ATTAIN), zwei Tiotropiumbromid-18µg-Studien (Brusasco 2003, Donohue 2002) und eine Tiotropiumbromid-5µg-Studie (Bateman 2010b). Abbildung 64 zeigt das Netzwerk, das basierend auf diesen Studien erzeugt werden kann.

Abbildung 64: Netzwerkdiagramm für die Analyse der SGRQ-Responder (%) nach 24 Wochen (alle Studien)



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-150: Operationalisierung von SGRQ Respondern

Studie	Operationalisierung
ACCORD I	Der SGRQ bietet ein objektives Lebensqualitätsmaß für Patienten mit COPD. Der Patient füllte den SGRQ nach 12 Wochen aus. Alle Bereiche des SGRQ-Fragebogens wurden ausgefüllt. SGRQ-Responder: Prozentsatz der Patienten, die eine Verringerung von mindestens 4 Einheiten gegenüber dem Ausgangswert der SGRQ-Gesamtpunktzahl in Woche 12 erreicht haben
ACCORD II	Der SGRQ bietet ein objektives Lebensqualitätsmaß für Patienten mit COPD. Es ist ein sensitiver, zuverlässiger und validierter, aus 2 Teilen bestehender Fragebogen: einer Symptomauswertung und einer Auswertung der Aktivität und Auswirkungen. Die Patienten mussten den Fragebogen bei den Visiten 2, 4, 5, 6 (12 Wochen, Ende von Teil 1 der Studie), 7, 8, 9 und 10 (40 Wochen, Ende von Teil 2 der Studie) oder bei vorzeitiger Beendigung vervollständigen. Alle Bereiche des SGRQ-Fragebogens wurden ausgefüllt. SGRQ-Responder: Prozentsatz der Patienten, die eine Verringerung von mindestens 4 Einheiten gegenüber dem Ausgangswert der SGRQ-Gesamtpunktzahl in Woche 12 erreicht haben.
ATTAIN	Der krankheitsspezifische Gesundheitszustand wurde mittels eines vom Patienten selbst auszufüllenden Instruments, des SGRQ, bewertet. Dieser Fragebogen ist ein standardisiertes selbst auszufüllendes Instrument zur Messung von Gesundheitsstörungen und wahrgenommenem Wohlbefinden ("QoL") bei Erkrankungen der Atemwege. Die relevante gültige Sprachversion des Fragebogens wurde dem Patienten während der Visite 2 (Bewertung der Ausgangswerte), Visite 4 (Woche 4), Visite 6 (Woche 12) und Visite 8 (Woche 24) zur Verfügung gestellt. Bei den letzten Visiten wurden den Patienten keine Informationen zu den Ausgangswert-Fragebogen (Visite 2) zur Verfügung gestellt. SGRQ-Responder: eine Änderung von 4 Einheiten im SGRQ wurde als Schwelle für eine klinisch relevante Veränderung des Gesundheitszustandes festgelegt. Daher wurde eine Verringerung um 4 Einheiten verwendet, um Patienten mit einer klinisch relevanten Verbesserung im SGRQ zu erfassen. Die Patienten mit einem geringeren Rückgang wurden als Non-Responder eingestuft.
Brusasco 2003	Die HRQoL wurde mit dem St. Georgs-Respiratory Fragebogen (SGRQ) bestimmt. Der SGRQ ist ein krankheitsspezifisches Instrument, das 50 Items in drei Subskalen (Symptome, Aktivität und Auswirkungen) enthält. Jede Antwort hat ein empirisch abgeleitetes Gewicht. Die Gesamtnote wird aus den Antworten aller 50 Items berechnet. Fehlende individuelle Fragen wurden nach den Richtlinien des Entwicklers des Fragebogens hinzugefügt. Die Daten aus fehlenden SGRQ Fragebögen wurden unter Verwendung der zuletzt durchgeführten Beobachtung übertragen. Ein minimaler klinisch relevanter Unterschied in SGRQ wurde als eine Verbesserung von mindestens 4 Einheiten in der SGRQ-Gesamtpunktzahl festgelegt.
Donohue 2002	Die HRQoL wurde mit dem St. George's Respiratory Fragebogen (SGRQ) gemessen. Der SGRQ ist ein krankheitsspezifisches Instrument, das 50 Items in drei Subskalen (Symptome, Aktivität und Auswirkungen) enthält. Die Gesamtnote wird aus den Antworten aller 50 Items berechnet. Ein niedrigerer Wert bedeutet eine Verbesserung. Eine Änderung von 4 Einheiten wird als klinisch signifikant angesehen.
TIPHON	Die SGRQ besteht aus einem 76-Item-Fragebogen, der in 3 Komponenten unterteilt ist: Symptome (zur Beurteilung der Behinderung wegen Atemwegsbeschwerden), Aktivität (Beurteilung der Auswirkungen der Atemlosigkeit auf Mobilität und körperliche Aktivität) und Auswirkungen (Beurteilung der psychosozialen Auswirkungen der Erkrankung). Die SGRQ-Gesamtpunktzahl setzt sich aus 3 Teilwerten zusammen. Die Werte werden gewichtet und reichen von 0 bis 100, wobei höhere Werte eine schlechtere gesundheitsbezogene Lebensqualität angeben. Derzeit gilt nach dem Autor des Fragebogens (Jones, 2002) eine Reduktion von ≥ 4 Einheiten als minimale klinisch bedeutsame Abweichung. Fehlende individuelle SGRQ-Items wurden gemäß den Empfehlungen der Autoren des Fragebogens wie folgt behandelt: fehlende Items aus Teil 1 oder Teil 2 (Abschnitte 1 und 7 des Fragebogens wurden negativem Response zugerechnet. Wenn einzelne SGRQ-Items aus Teil 2

	<p>(Abschnitte 2-6 und 7 des Fragebogens) fehlten, wurde das Gesamtgewicht unter Berücksichtigung der fehlenden Werte angepasst. Wenn mehr als 10 Items in diesen Abschnitten fehlten, dann wurde der Wert der Komponente nicht in die anschließende Analyse einbezogen.</p> <p>Die maximale Anzahl fehlender Items betrug 2, 4, und 6 für die Symptom-, Aktivitäts- bzw. Auswirkungs-Komponenten. Darüber hinaus wurden fehlende SGRQ-Komponenten oder Gesamtpunktzahlen unter Heranziehung der Fortschreibungsmethode anhand der zuletzt durchgeführten Beobachtung berechnet (außer, wenn Tag 1 die letzte Visite war, in diesem Fall wurde der Patient ausgeschlossen). Die VSRQ-Gesamtpunktzahl wurde aus den 8 Kategorie-punktzahlen berechnet. Da es keine spezifische Empfehlung für den Umgang mit fehlenden Daten gab, wurde die Methode der Fortschreibung anhand der letzten durchgeführten Beobachtung verwendet.</p>
Verkindre 2006	<p>Die HRQoL wurde mit dem St. George's Respiratory Fragebogen (SGRQ) vor der Behandlung an den Tagen 1, 42 und 84 des Behandlungszeitraumes gemessen. Der SGRQ besteht aus einem 76-Item-Fragebogen, der in 3 Komponenten unterteilt ist: Symptome zur Beurteilung der Behinderung wegen Atemwegsbeschwerden; Aktivität zur Beurteilung der Auswirkungen der Atemlosigkeit auf Mobilität und körperliche Aktivität und Auswirkungen zur Beurteilung der psychosozialen Auswirkungen der Erkrankung. Die SGRQ-Gesamtpunktzahl besteht aus allen 3 Teilwerten. Die Werte werden gewichtet und reichen von 0 bis 100, wobei höhere Werte eine schlechtere gesundheitsbezogene Lebensqualität angeben. Ein Unterschied von 4 Einheiten wird als klinisch relevant angesehen.</p>
Bateman 2010b	<p>Die HRQoL wurde mit dem St. George's Respiratory Fragebogen (SGRQ) gemessen. SGRQ-Responder wurden als solche definiert, die eine Änderung von 4 Einheiten in der SGRQ-Gesamtpunktzahl (minimale klinisch bedeutende Differenz [MCID]) erreichten.</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-151: Bewertung des Verzerrungspotenzials für den SGRQ-Wert in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Brusasco 2003	Niedrig	Ja	Unklar	Unklar	Nein	Niedrig
Donohue 2002	Niedrig	Ja	Nein	Nein	Nein	Hoch
TIPHON	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Verkindre 2006	Niedrig	Ja	Nein	Nein	Nein	Hoch
Bateman 2010b	Niedrig	Ja	Nein	Unklar	Ja	Hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Bei der ACCORD II-Studie wurde das Verzerrungspotenzial in Bezug auf den Prozentanteil der SGRQ-Responder als hoch eingestuft. Alle Tiotropiumbromid-Studien wiesen auf Studienebene ein geringes Verzerrungspotenzial auf. Drei von fünf Studien wiesen jedoch besonders bezüglich der ergebnisbezogenen SGRQ-Responder ein hohes Verzerrungspotenzial auf; dies ist hauptsächlich darauf zurückzuführen, dass das ITT-Prinzip nicht für die Endpunkt-Analyse angewendet wurde (detaillierte Erklärung im Anhang 4-G).

Stellen Sie die Ergebnisse für SGRQ-Responder für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-152: Ergebnisse für SGRQ-Responder aus RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N Responder (%)	Gruppendifferenz (SE) [KI] p-Wert
ACCORD I	Aclidiniumbromid	190	84 (44,4%)	1,40 (0,21) [0,92; 2,11] p=0,130
	Placebo	185	67 (35,9%)	
ACCORD II	Aclidiniumbromid	177	82 (44,8%)	1,41 (0,21) [0,93; 2,15] p=0,120
	Placebo	182	69 (38,8%)	
ATTAIN	Aclidiniumbromid	269	153 (56,9%)	2,02 (0,18) [1,43; 2,84] p<0,001
	Placebo	273	108 (39,5%)	
TIPHON (Tonnel 2008)	Tiotropiumbromid 18µg	247	158 (64,0%)	1,80 (0,17) [1,29; 2,53] p<0,001
	Placebo	245	129 (52,7%)	
Verkindre 2006	Tiotropiumbromid 18µg	46	27 (58,7%)	2,62 (0,41) [1,16; 5,89] p=0,02
	Placebo	54	19 (35,2%)	

N=Anzahl der Patienten, Responder = Anzahl der Patienten mit einer Reduzierung des SGRQ-Gesamtwerts von mindestens 4 Einheiten im Vergleich zum Ausgangswert, SE=Standardfehler ITT=Intention-to-treat-Population, KI=Konfidenzintervall

Tabelle 4.3-152 zeigt die Anzahl der SGRQ-Responder (d. h. Patienten mit einer Reduzierung des SGRQ-Gesamtwerts von mindestens 4 Einheiten) nach 12 Wochen in jeder Studie, die dieses Outcome berichtete. Aclidiniumbromid erhöhte die Responsewahrscheinlichkeit, dies war in einer Studie statistisch signifikant. In den Tiotropiumbromid-Studien signifikante Verbesserung der Lebensqualität in beiden Studien beobachtet.

Tabelle 4.3-153: Ergebnisse für SGRQ-Responder aus RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N Responder (%)	Gruppendifferenz (SE) [KI] p-Wert
ATTAIN	Acclidiniumbromid	269	154 (57,3%)	1,93 (0,17) [1,37; 2,71] p< 0,001
	Placebo	273	112 (41,0%)	
Bateman 2010b	Tiotropiumbromid 5µg	1989	1012 (50,9%)	1,43 (0,06) [1,26; 1,62] p< 0,001
	Placebo	2002	841 (42%)	
Brusasco 2003	Tiotropiumbromid 18µg	386	197 (51,0%)	1,49 (0,14) [1,12; 1,97] p= 0,006
	Placebo	362	157 (43,4%)	
Donohue 2002	Tiotropiumbromid 18µg	209	107 (51%)	1,46 (0,20) [0,99; 2,16] p=0,057
	Placebo	201	84 (42%)	

N=Anzahl der Patienten, Responder = Anzahl der Patienten mit einer Reduzierung des SGRQ-Gesamtwerts von mindestens 4 Einheiten im Vergleich zum Ausgangswert, SE=Standardfehler ITT=Intention-to-treat-Population, KI=Konfidenzintervall

Tabelle 4.3-153 zeigt die Anzahl der SGRQ-Responder nach 24 Wochen für alle Studien, die dieses Outcome berichten. Mit Acclidiniumbromid wurde die Responsewahrscheinlichkeit signifikant erhöht. In zwei von drei Tiotropiumbromid-Studien zeigte sich bei den Patienten eine signifikante Verbesserung der Lebensqualität. Da nur eine Acclidiniumbromid-Studie diesen Endpunkt nach 24 Wochen berichtete, war eine Meta-Analyse für Acclidiniumbromid-Studien nach 24 Wochen nicht möglich.

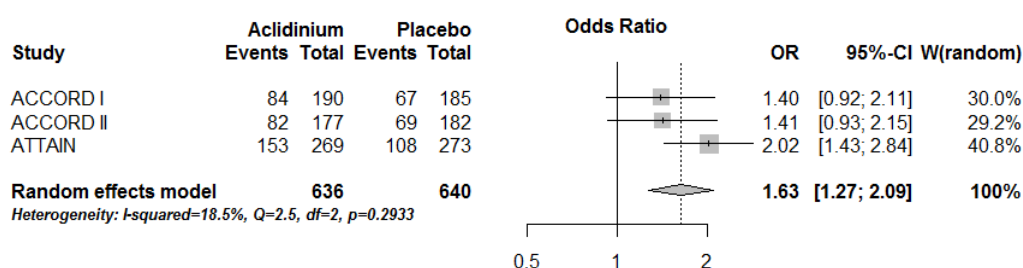


Abbildung 65: Meta-Analyse für den Prozentsatz der SGRQ-Responder (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Acclidiniumbromid zeigt bei der Erhöhung des Anteils der SGRQ-Responder eine höhere Effektivität als Placebo (OR = 1,63 [95 % CI: 1,27; 2,09]). Eine signifikante Heterogenität der Studien gab es nicht (Q = 2,5, p = 0,29, I² = 18,5 %) (Abbildung 65).

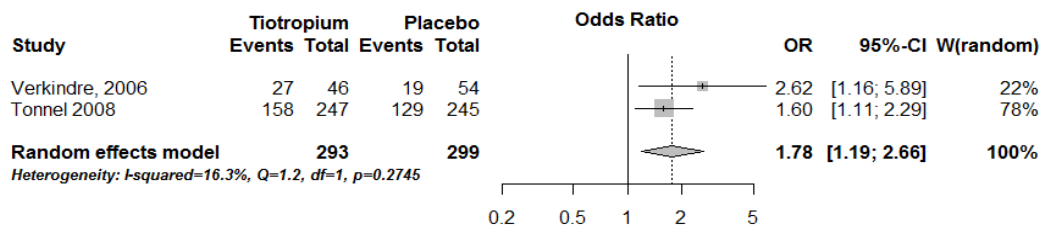


Abbildung 66: Meta-Analyse für den Prozentsatz der SGRQ-Responder (12 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.

Tiotropiumbromid ist in Bezug auf den Endpunkt Anteil der SGRQ-Responder nach 12 Wochen effektiver als Placebo (OR = 1,78 [95 % CI: 1,19; 2,66]). Eine signifikante Heterogenität der Studien gab es nicht ($Q = 1,2$ $p = 0,27$, $I^2 = 16,3$ %) (Abbildung 66).

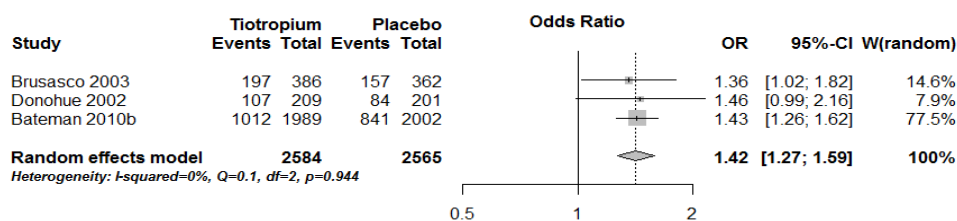


Abbildung 67: Meta-Analyse für den Prozentsatz der SGRQ-Responder (24 Wochen) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo.

Tiotropiumbromid ist in Bezug auf den Endpunkt Anteil der SGRQ-Responder nach 24 Wochen effektiver als Placebo (OR = 1,42 [95 % CI: 1,27; 1,59]). Eine signifikante Heterogenität der Studien gab es nicht ($Q = 0,1$, $p = 0,94$, $I^2 = 0$ %) (Abbildung 67).

Eine Sensitivitätsanalyse der Tiotropiumbromid-Studien bezüglich der verzerrten Studien war aufgrund der niedrigen Anzahl von Studien nicht möglich.

Sensitivitätsanalyse

Die ACCORD-II-Studie wies hinsichtlich der Baseline-Patientencharakteristika ein Ungleichgewicht zwischen den Behandlungsgruppen auf. Um die Auswirkungen des Ausschlusses dieser Studie aus der Meta-Analyse zu analysieren, haben wir eine Sensitivitätsanalyse durchgeführt (Abbildung 68).

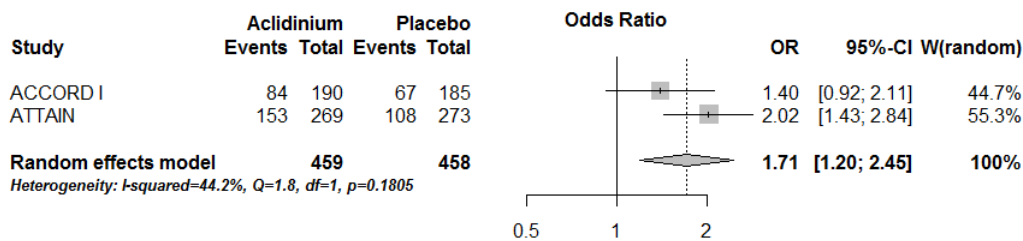


Abbildung 68: Sensitivitätsanalyse (Verzerrungspotential) für SGRQ-Responder nach 12 Wochen (ml) aus RCT; Acridiniumbromid im Vergleich zu Placebo.

Wie Abbildung 68 zeigt, erhöht Acridiniumbromid die Wahrscheinlichkeit eines SGRQ-Ansprechens nach 12 Wochen im Vergleich zu Placebo (OR 1,71 [95 % CI: 1,20; 2,45]). Der Ausschluss der ACCORD-II-Studie beeinflusst den gepoolten Effektschätzwerte nicht signifikant, wie die Konfidenzintervalle der Metaanalyse zeigen. Der Heterogenitätstest zeigt nach Ausschluss der ACCORD-II-Studie jedoch eine Heterogenität der Studien ($Q = 1,8$, $p = 0,18$, $I^2 = 44,2\%$), die beim gesamten Studienpool nicht zu beobachten war. Das Ergebnis der Meta-Analyse wird aufgrund des signifikanten Effektschätzers dennoch als relevant eingestuft.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-154 zeigt die DIC-Schätzwerte für die geplanten NMAs. Nach 12 Wochen ergab die Fixed-Effekt-Analyse die besten DIC-Ergebnisse (Tabelle 4.3-154). Das war mit den Metaanalyse-Ergebnissen konsistent. Szenario 2 wurde nicht nach 12 Wochen durchgeführt, da alle Patienten in den Studien, die Daten zu dem SGRQ-Gesamtwert nach 12 Wochen berichteten, keine Begleitmedikation mit LABA einsetzten. Nach 24 Wochen ergab die Fixed-Effekt-Analyse sowohl für Szenario 1 als auch für Szenario 2 die besten DIC-Ergebnisse (Tabelle 4.3-84). Dies ist mit den Meta-Analysen konsistent, da darin keine signifikanten Heterogenitäten festgestellt wurden.

Tabelle 4.3-154: DIC für SGRQ-Responder - indirekter Vergleich aus RCT

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung für die NMA-Modelle
Woche 12			
Ohne LABA-Begleitmedikation	72,7	73,7	Feste Effekte
Alle Studien	-	-	-
Woche 24			
Ohne LABA-Begleitmedikation	46,4	47,9	Feste Effekte
Alle Studien	64,568	66,527	Feste Effekte

Die Ergebnisse der Netzwerk-Meta-Analyse sind in Tabelle 4.3-155, Tabelle 4.3-156 und Tabelle 4.3-157 aufgeführt.

Tabelle 4.3-155: SGRQ-Responder (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 18
TIO 18	Effektschätzer	1,77	
	95% KI	1,25	2,44
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	1,65	
	95% KI	1,31	2,06
	P(Überlegenheit)	>99%	
		0,96	
		0,63	1,41
		39%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 12 Wochen lag die Odds Ratio des Endpunkts SGRQ-Responder von Acclidiniumbromid versus Placebo oberhalb von 1 und die Grenzwerte des 95 %-Konfidenzintervalls umfassten die 1 nicht. Deshalb kann geschlussfolgert werden, dass Acclidiniumbromid die Zahl der SGRQ-Responder im Vergleich zu Placebo erhöhte [OR 1,65 (95 %-KI: 1,3, 2,1)]. Die NMA-Ergebnisse des Vergleichs Acclidiniumbromid und Tiotropiumbromid versus Placebo waren mit den Ergebnissen der Metaanalysen konsistent. Sowohl Acclidiniumbromid als auch Tiotropiumbromid erhöhten die Zahl der SGRQ-Responder nach 12 Wochen im Vergleich zu Placebo. (Tabelle 4.3-155, Abbildung 65, Abbildung 66). Die Grenzwerte des 95 %-Konfidenzintervalls der NMA Acclidiniumbromid versus Tiotropiumbromid schlossen die 1 mit ein. Die Zahl der SGRQ-Responder unter Acclidiniumbromid ist mit der unter Tiotropiumbromid als gleichwertig zu betrachten [OR 0,96 (95 %-KI: 0,6, 1,4)] (Tabelle 4.3-155).

Tabelle 4.3-156: SGRQ-Responder (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 18
TIO 18	Effektschätzer	1,41	
	95% KI	1,11	1,76
	P(Überlegenheit)	>99%	
AB 400	Effektschätzer	1,96	
	95% KI	1,37	2,72
	P(Überlegenheit)	>99%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Aclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-157: SGRQ-Responder (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: Alle Studien; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	1,42	
	95% KI	1,27	1,59
	P(Überlegenheit)	100%	
AB 400	Effektschätzer	1,96	
	95% KI	1,37	2,72
	P(Überlegenheit)	>99%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg;

AB 400= Aclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen war Aclidiniumbromid ebenfalls effektiver als Placebo [OR 1,96 (95 %-KI: 1,4, 2,7)]. Die NMA ist mit den Ergebnissen der Aclidiniumbromid-Meta-Analysen (Abbildung 67) konsistent. Gegenüber Tiotropiumbromid ist hinsichtlich der SGRQ-Responder unter Aclidiniumbromid eine Gleichwertigkeit mit Tendenz zur Überlegenheit festzustellen [OR 1,4 (95 %-KI: 0,9, 2,1)] (Tabelle 4.3-156). Die Ergebnisse von Szenario 2 bestätigten die Ergebnisse von Szenario 1 (Tabelle 4.3-157).

Sowohl Aclidiniumbromid als auch Tiotropiumbromid steigerten die Rate der Patienten mit SGRQ-Response statistisch signifikant im Vergleich zu Placebo. Der indirekte Vergleich zeigte, dass beide Substanzen die Rate der SGRQ-Responder in gleichwertiger Weise gegenüber Placebo erhöhen. Sowohl nach 12 als auch nach 24 Wochen schloss das 95 %-Konfidenzintervall die Eins ein, d. h. bei der Wirkung auf die SDRQ-Responderrate ergaben sich keine statistisch signifikanten Unterschiede. Der Nutzen von Aclidiniumbromid und Tiotropiumbromid auf den Endpunkt SGRQ ist also gleichwertig.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.12 Unerwünschte Ereignisse – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Tabelle 4.3-158: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Unerwünschte Ereignisse - 12 Wochenbegleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I	•			•
	ACCORD II	•			•
2	Beeh 2006			•	•
	Casaburi 2000			•	•
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Fünf Studien berichteten über die Anzahl der Patienten, die nach 12 Wochen (LABAs waren nicht gestattet) über unerwünschte Ereignisse berichtet haben; zwei Aclidiniumbromid-Studien, zwei Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 69 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 69: Netzwerk-Diagramm für die Analyse unerwünschter Ereignisse nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

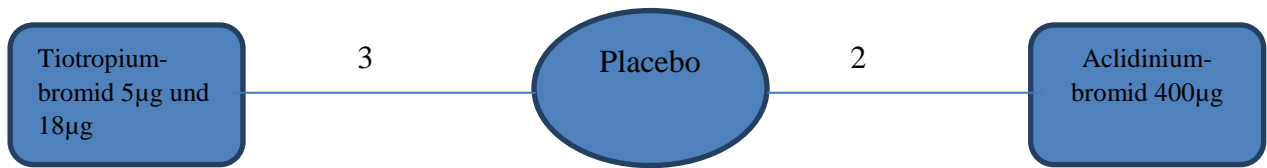


Tabelle 4.3-159: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Unerwünschte Ereignisse - 12 Wochen - Alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I	•			•
	ACCORD II	•			•
5	SPRUCE			•	•
	Johansson 2008			•	•
	SAFE Portugal			•	•
	Beeh 2006			•	•
	Casaburi 2000			•	•
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Acht Studien berichteten über die Anzahl der Patienten mit unerwünschten Ereignissen nach 12 Wochen; zwei Aclidiniumbromid-Studien, fünf Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 70 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 70: Netzwerk-Diagramm für die Analyse unerwünschter Ereignisse nach 12 Wochen

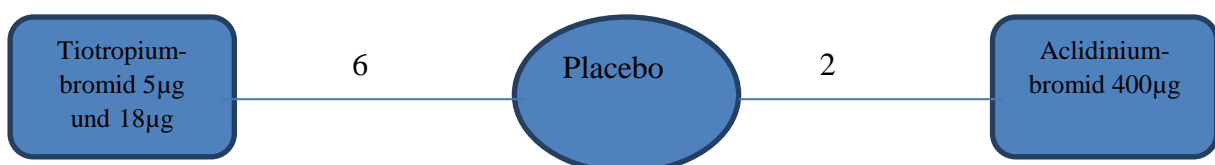


Tabelle 4.3-160: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Unerwünschte Ereignisse - 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
2	INHANCE			•	•
	Vogelmeier 2008			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Drei Studien berichteten über die Anzahl der Patienten mit unerwünschten Ereignissen nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet); eine Aclidiniumbromid-Studie (ATTAIN) und zwei Tiotropiumbromid-18µg-Studien (INHANCE, Vogelmeier 2008). Keine Tiotropiumbromid-5µg-Studie berichtete über diesen Endpunkt. Abbildung 71 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 71: Netzwerk-Diagramm für die Analyse unerwünschter Ereignisse nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)



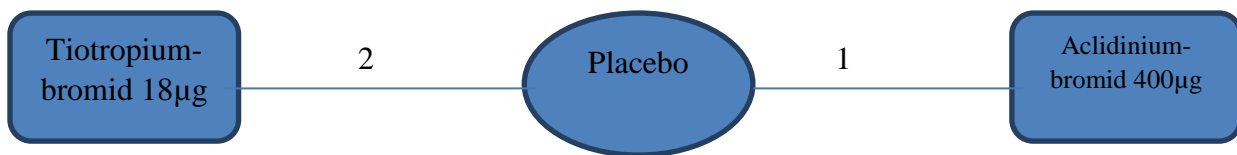
Tabelle 4.3-161: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Unerwünschte Ereignisse - 24 Wochen – alle Studien, inkl. alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
2	INHANCE			•	•
	Vogelmeier 2008			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Drei Studien berichteten über die Anzahl der Patienten mit unerwünschten Ereignissen nach 24 Wochen; eine Aclidiniumbromid-Studie (ATTAIN) und zwei Tiotropiumbromid-18µg-Studien (INHANCE, Vogelmeier 2008). Keine Tiotropiumbromid-5µg-Studie berichtete über dieses Outcome. Abbildung 72 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 72: Netzwerk-Diagramm für die Analyse unerwünschter Ereignisse nach 24 Wochen (alle Studien)



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-162: Operationalisierung von unerwünschten Ereignissen

Studie	Operationalisierung
ACCORD I	<p>Ein AE [unerwünschtes Ereignis] ist jedes unerwünschte medizinische Geschehen bei einem Patienten bei Verabreichung eines pharmazeutischen Produkts, das nicht unbedingt in kausalem Zusammenhang mit dieser Behandlung stehen muss. Ein AE kann daher jedes ungünstige und ungewollte Anzeichen (beispielsweise ein abnormaler Laborbefund), Symptom oder jede Erkrankung sein, die zeitlich entweder mit der Verwendung eines medizinischen Erzeugnisses verbunden ist oder als nicht mit dem medizinischen Erzeugnis in Zusammenhang stehend angesehen wird. Jedes unerwünschte Geschehen oder ungünstiges und unbeabsichtigtes Anzeichen, Symptom oder jede Krankheit, das bis 30 Tage nach der letzten Dosis der Behandlung auftritt, wird als Nebenwirkung angesehen. Bei jeder Untersuchung wurden die Patienten zu allen AEs befragt, die seit der letzten Untersuchung aufgetreten sind.</p>
ACCORD II	<p>Ein AE [unerwünschtes Ereignis] ist jedes unerwünschte medizinische Geschehen bei einem Patienten bei Verabreichung eines pharmazeutischen Produkts, das nicht unbedingt in kausalem Zusammenhang mit dieser Behandlung stehen muss. Ein AE kann daher jedes ungünstige und ungewollte Anzeichen (beispielsweise ein abnormaler Laborbefund), Symptom oder jede Erkrankung sein, die zeitlich entweder mit der Verwendung eines medizinischen Erzeugnisses verbunden ist oder als nicht mit dem medizinischen Erzeugnis in Zusammenhang stehend angesehen wird. Jedes unerwünschte medizinische Geschehen oder ungünstiges und unbeabsichtigtes Anzeichen, Symptom oder jede Krankheiten, das vom Zeitpunkt der Unterzeichnung der Einwilligungserklärung bis 30 Tage nach der letzten Dosis des Untersuchungsproduktes auftritt, wird als Nebenwirkung angesehen.</p> <p>Bei jeder Untersuchung wurden die Patienten zu allen AEs/SAEs [schwerwiegenden AEs] befragt, die seit der letzten Untersuchung aufgetreten sind. Die Patienten wurden gebeten, Antworten auch auf eine nicht-suggestive Fragen zu geben. Das Personal des Studienzentrums zeichnet alle sachdienlichen Informationen in den eCRFs [elektronischen Patienten-Prüfbogen] auf. Alle durch den Patienten (oder einen Patientenvertreter) berichteten oder beobachteten oder anderweitig durch die Untersuchungsperson (oder anderes Prüfpersonal) bei einer Prüfuntersuchung bemerkten oder während irgendeiner anderen Kommunikation mit dem Patienten (oder einem Patientenvertreter) außerhalb der festgelegten Prüfuntersuchungen vorgekommenen AEs/SAEs (vom Zeitpunkt der Unterzeichnung der Einwilligungserklärung bis 30 Tage nach der letzten Gabe des Prüfmedikaments) müssen dokumentiert werden. Alle AEs mussten in einem entsprechenden AE-Berichts-Formular in dem eCRF des Patienten erfasst werden, unabhängig davon, ob sie in ursächlichem Zusammenhang mit dem Prüfpräparat stehen oder nicht. Die Safety-Analyse (unerwünschte Ereignisse einschließlich COPD-Exazerbationen) wurde unter Verwendung der Safety-Population durchgeführt und durch eine Häufigkeitsverteilung (Anzahl und Prozentsatz der Patienten) zusammengefasst.</p>
ATTAIN	<p>Unerwünschte Ereignisse wurden vom Zeitpunkt der Unterzeichnung der Einwilligungserklärung bis zu zwei Wochen nach der Gabe des letzten Studienmedikaments oder einem vorzeitigen Abbruch aufgezeichnet. Die unerwünschten Ereignisse wurden erfasst, indem den Patienten nicht-suggestive Fragen gestellt wurden sowie durch die spontanen Aufzeichnungen der Patienten, die unterwiesen waren, AE's in den Patient Diary Cards täglich zwischen den Untersuchungen aufzuzeichnen.. Jegliche auf der Patienten-Tagebuch-Karte aufgezeichneten AEs wurden von der Untersuchungsperson mit medizinischen Standardbegriffen in die eCRF übertragen. Diese Abweichungen in Laborversuchen, ECGs und körperlichen Untersuchungen, die von der Untersuchungsperson als klinisch signifikant betrachtet wurden oder die ein Eingreifen oder einen Diagnose-Test verlangten oder die aus einem Abbruch des medizinischen Untersuchungsproduktes resultierten, wurden als AEs erfasst. Ein AE ist definiert als jedes unerwünschte medizinische Vorkommnis in einer klinischen Patientenstudie (unabhängig von der Anwendung des IMP und seines kausalen Zusammenhangs mit dem berichteten Ereignis). Ein AE könnte daher jedes ungünstige medizinische Ereignis in der Patientenstudie sein, einschließlich der Verschlechterung einer bereits bestehenden Erkrankung, eines abnormen Werts einer</p>

	Laboruntersuchung, einer EKG-Anomalie oder eines abnormen Befundes bei der körperlichen Untersuchung. Die Untersuchungsperson überwacht engmaschig alle AEs und ergreift die notwendigen klinischen Maßnahmen, um die Sicherheit der Patienten zu gewährleisten. Unerwünschte Ereignisse müssen in zeitlichem Zusammenhang mit der Teilnahme des Patienten an der Studie stehen, d. h. sie müssen zwischen der Unterzeichnung des ICF und dem letzten planmäßigen Kontakt mit dem Patienten auftreten. Zum Zeitpunkt des Auftretens eines AE darf die Verabreichung des IMP daher noch nicht eingeleitet worden sein. Wenn sie bereits eingeleitet wurde, muss es nicht unbedingt eine positive kausale Beziehung zu dem Ereignis haben. Die Analysen von Safety Outcomes wurden über die Safety-Population durchgeführt.
INHANCE	Die Patienten verwendeten Tagebuchkarten zur Aufzeichnung unerwünschter Ereignisse. Sicherheitsrelevante Daten wurden mit einem adjustiertem Modell aus Behandlung, Raucherstatus und Land als feststehende Einflüsse und Ausgangs-FEV ₁ und Reversibilität als Kovariaten analysiert. Fehlende Werte in Woche 12 wurden mit dem vorherigen Wert (ab Woche 2 oder höher) ersetzt.
SPRUCE	Bei allen Untersuchungen wurden alle unerwünschten Ereignisse, schwere und nicht-schwerwiegende, unabhängig von der Kausalität gesammelt.
Johansson 2008	Die unerwünschten Ereignisse wurden in jedem Zentrum aufgezeichnet und am Studienende zusammengetragen.
SAFE Portugal	Während der Studie wurden unerwünschte Ereignisse gesammelt. Eine beschreibende Statistik wurde für die Safety- Variablen verwendet.
Vogelmeier 2008	Alle unerwünschten Ereignisse (AEs) wurden aufgezeichnet Die Intent-to-treat (ITT)-Population umfasste alle randomisierten Patienten, die mindestens eine Gabe des Studienmedikaments erhielten. Diese Population wurde für Wirksamkeit und Sicherheitsanalysen eingesetzt. Alle Analysen wurden mit einem 5% Signifikanzniveau (2-seitig) durchgeführt.
Voshaar 2008	Eine beschreibende Statistik wurde für die Safety-Variablen verwendet. Unerwünschte Ereignisse wurden mit dem Medical Dictionary for Regulatory Activities (MedDRA) Version 8.0 codiert und nach Organsystem und bevorzugten Bedingungen klassifiziert. Alle randomisierten Patienten mit jeglichen verfügbaren Daten wurden in die Sicherheitsbewertung einbezogen.
Beeh 2006	Die unerwünschten Ereignisse wurden den Behandlungsgruppen wie behandelt zugeordnet und ebenfalls für alle behandelten Patienten ausgewertet.
Casaburi 2000	Untersuchungen zur Bewertung und Aufzeichnung unerwünschter Ereignisse.in der Studienklinik erfolgten am ersten Tag der Studienmedikaments-Verabreichung, am Ende der ersten Therapiewoche und alle 3 Wochen danach während der 13-wöchigen Behandlungsperiode.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-163: Bewertung des Verzerrungspotenzials für unerwünschte Ereignisse in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Ja	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
INHANCE	Niedrig	Nein	Ja	Nein	Nein	Hoch
SPRUCE	Hoch	Ja	Nein	Unklar	Nein	Hoch
Johansson 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
SAFE Portugal	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Vogelmeier 2008	Hoch	Nein	Ja	Unklar	Nein	Hoch
Voshaar 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Beeh 2006	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
Casaburi 2000	Niedrig	Ja	Unklar	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die ACCORD II-Studie wurde das Verzerrungspotenzial für unerwünschte Ereignisse als hoch angesehen (Tabelle 4.3-163), weil die Behandlungsgruppen bei Studienbeginn nicht vollständig vergleichbar waren. Zwei der acht Tiotropiumbromid-Studien hatten ein hohes Verzerrungspotenzial auf der Studienebene. Das Verzerrungspotenzial war speziell für den Endpunkt unerwünschte Ereignisse bei drei von acht Tiotropiumbromid-Studien hoch. Diese Studien waren nicht verblindet oder verstießen gegen das ITT-Prinzip bei der Analyse dieses Endpunktes (Details sind in Anhang 4-G beschrieben).

Stellen Sie die Ergebnisse für den Endpunkt unerwünschte Ereignisse für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-164: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ACCORD I	Acidiniumbromid	190	85 (44,7%)	0,74 (0,20) [0,50;1,11] p=0,139
	Placebo	186	97 (52,2%)	
ACCORD II	Acidiniumbromid	177	90 (50,8%)	1,06 (0,21) [0,70; 1,60] p=0,782
	Placebo	182	90 (49,5%)	
SAFE Portugal	Acidiniumbromid	147	17 (11,6%)	0,69 (0,34) [0,36; 1,34] p=0,268
	Placebo	164	26 (15,9%)	
SPRUCE	Tiotropiumbromid 18µg	200	102 (51,0%)	0,65 (0,20) [0,44; 0,97] p=0,033
	Placebo	195	120 (61,5%)	
Johansson 2008	Tiotropiumbromid 18µg	107	50 (46,7%)	1,22 (0,27) [0,72; 2,06] p=0,458
	Placebo	117	49 (41,9%)	
Beeh 2006	Tiotropiumbromid 18µg	1236	321 (26,0%)	0,86 (0,13) [0,67; 1,10]
	Placebo	403	117 (29,0%)	
Casabury 2000	Tiotropiumbromid 18µg	279	172 (61,6%)	0,81 (0,20) [0,55; 1,19]
	Placebo	191	127 (66,5%)	
Voshaar 2008	Tiotropiumbromid 5µg	180	95 (52,8%)	0,77 (1,24) [0,51; 1,17] p= 0,22
	Placebo	181	107 (59,1%)	

N=Anzahl der Patienten SE= Standardfehler, ITT= Intention-to-treat-Population [KI]=Konfidenzintervall

Tabelle 4.3-164 zeigt den Prozentsatz der Patienten, die in den Acidiniumbromid- und Tiotropiumbromid-5µg und 18µg-Studien nach 12 Wochen unerwünschte Ereignisse berichtet haben. Die Behandlungen weisen bis auf eine Studie (SPRUCE) keinen signifikanten Unterschied zu Placebo auf.

Tabelle 4.3-165: Ergebnisse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ATTAIN	Acidiniumbromid	269	144 (53,5%)	0,86 (0,17) [0,62; 1,21] p= 0,40
	Placebo	273	156 (57,1%)	
Vogelmeier 2008	Tiotropiumbromid 18µg	221	79 (35,7%)	0,86 (0,20) [0,58; 1,27] p=0,451
	Placebo	209	82 (39,2%)	
INHANCE	Tiotropiumbromid 18µg	415	279 (67,2%)	1,17 (0,15) [0,88; 1,56] p=0,282
	Placebo	418	266 (63,6%)	

N=Anzahl der Patienten SE= Standardfehler, ITT= Intention-to-treat-Population [KI]=Konfidenzintervall

Tabelle 4.3-165 zeigt den Prozentsatz der Patienten, die in den Acidiniumbromid- und Tiotropiumbromid-5µg und 18µg-Studien nach 24 Wochen unerwünschte Ereignisse berichtet haben. Die Behandlungseffekte unterschieden sich in den Tiotropiumbromid-18µg-Studien. Eine Meta-Analyse war für Acidiniumbromid nicht möglich, weil nur die ATTAIN-Studie über dieses Ergebnis berichtete.

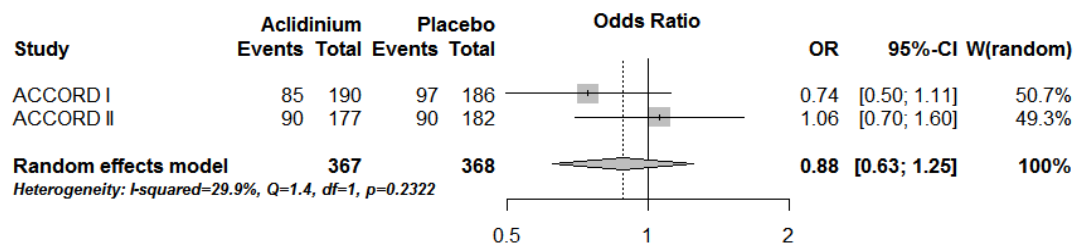


Abbildung 73: Meta-Analyse für unerwünschte Ereignisse (12 Wochen) aus RCT; Acidiniumbromid im Vergleich zu Placebo

Abbildung 73 zeigt, dass Acidiniumbromid beim Endpunkt unerwünschte Ereignisse mit Placebo vergleichbar ist (gepooltes OR 0,88 [95%-KI 0,63; 1,25]). Zwischen den beiden Studienergebnissen gab es deutliche Unterschiede. Während Acidiniumbromid in ACCORD I die Rate unerwünschter Ereignisse im Vergleich zu Placebo senkte, war diese Rate in der ACCORD II-Studie tendenziell höher. Für diesen Endpunkt bestand jedoch keine signifikante Studienheterogenität (Q=1,4; p=0,23; I²=29,9%). Allerdings könnte die

unterschiedliche Krankheitsschwere in der ACCORD II-Studie zu Studienbeginn eine Rolle gespielt haben.

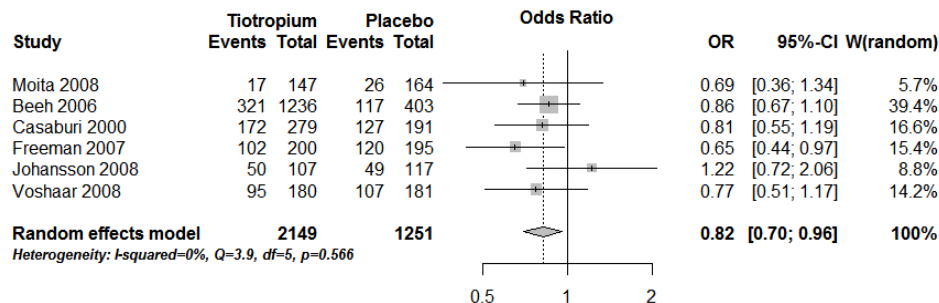


Abbildung 74: Meta-Analyse für unerwünschte Ereignisse (12 Wochen) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo

Abbildung 74 zeigt, dass die Tiotropiumbromid-Behandlung mit weniger unerwünschten Ereignissen assoziiert war als eine Placebo-Behandlung (gepooltes OR 0,82 [95%-KI 0,70; 0,96]). Zwischen den Punktschätzern der einzelnen Studienergebnisse gab es deutliche Unterschiede, von einem Schutzeffekt durch Tiotropiumbromid (SAFE Portugal, Beeh 2006, Casaburi 2000, SPRUCE und Voshaar 2008) bis hin zu einem zunehmenden Risiko unerwünschter Ereignisse (Johansson 2008). Für dieses Outcome bestand keine signifikante Studienheterogenität ($Q=3,9$; $p=0,57$; $I^2=0\%$).

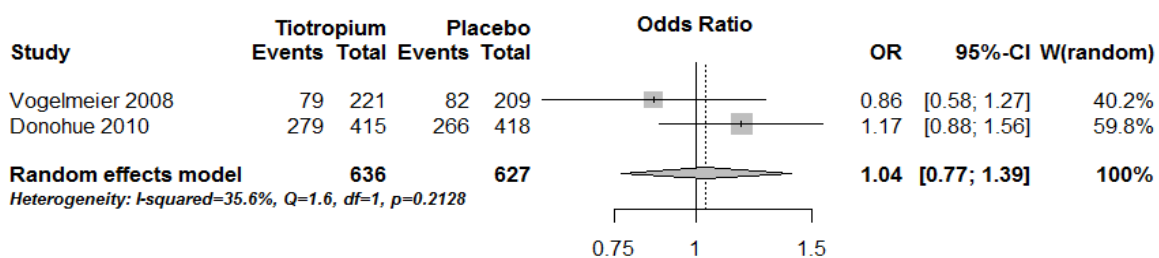


Abbildung 75: Meta-Analyse für unerwünschte Ereignisse (24 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo

Abbildung 75 zeigt, dass die zwei Studien, die in diese Meta-Analyse eingeschlossen wurden, gegensätzliche Behandlungseffekte hatten. Die gepoolte OR betrug 1,04 [95%-KI 0,77; 1,39]. Beide Studien wiesen ein hohes Verzerrungspotenzial für diesen Endpunkt auf. Es bestand keine signifikante Studienheterogenität ($Q=1,6$; $p=0,21$; $I^2=35,6\%$). Eine Sensitivitätsanalyse war nicht durchführbar, da nur zwei Studien diesen Endpunkt berichteten.

Sensitivitätsanalyse

Um die Auswirkungen des Schweregrads der Erkrankung auf den gepoolten Effekt zu untersuchen, wurden zwei Studien, in der überwiegend Patienten mit mittelschwerer COPD behandelt wurden (Freeman 2007 (SPRUCE) und Johansson 2008), ausgeschlossen.

Die Ergebnisse dieser Analyse sind in Abbildung 76 dargestellt.

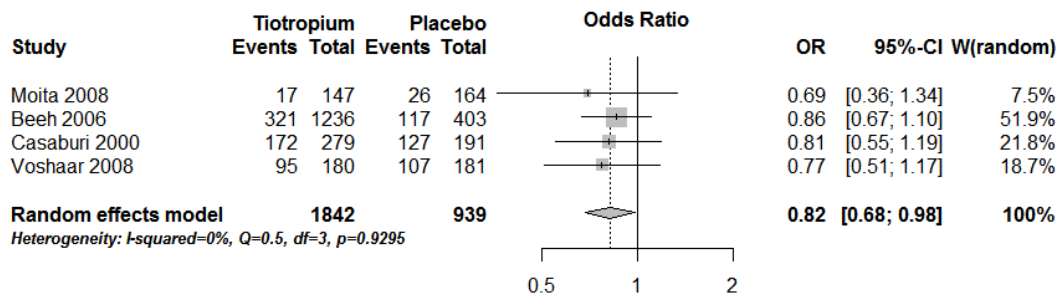


Abbildung 76: Sensitivitätsanalyse (COPD-Schweregrad) für unerwünschte Ereignisse (12 Wochen) aus RCT; Tiotropiumbromid 5µg oder 18µg im Vergleich zu Placebo

Abbildung 76 zeigt, dass die Tiotropiumbromid-Behandlung mit weniger unerwünschten Ereignissen assoziiert war als eine Placebo-Behandlung (OR 0,82 [95%-KI 0,68; 0,98]). Wie in der Meta-Analyse, welche die Studien von Freeman (SPRUCE) und Johansson 2008 einschloss, zeigten die mit Tiotropiumbromid behandelten Patienten auch in der Sensitivitätsanalyse ein geringeres Risiko für. Die Ergebnisse ähneln denen der Meta-Analyse, in denen sie oben genannten Studien eingeschlossen wurden, daher wird davon ausgegangen, dass die Studienpopulation von Freeman (SPRUCE) und Johansson 2008 nur geringe Auswirkungen auf die Ergebnisse hat. Für diesen Endpunkt bestand keine signifikante Studienheterogenität ($Q=0.5$; $p=0.93$; $I^2=0\%$).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-194 zeigt die DIC-Schätzungen für die geplanten NMAs. Nach 12 Wochen war das beste durch DIC-Ergebnisse ermittelte Modell ein Modell mit festen Effekten mit Szenario 1 (Studien ohne LABA-Begleitmedikation) und Szenario 2 (alle Studien, inkl. LABA-Begleitmedikation; Tabelle 4.3-166). Die Verwendung eines Modells mit festen Effekten stimmte mit den Ergebnissen der Meta-Analyse der Tiotropiumbromid-5µg- und Tiotropiumbromid-18µg-Studien überein, in welcher der Heterogenitätstest auf keine signifikante Heterogenität schließen ließ (Abbildung 74). Nach 24 Wochen ergab sich aus den DIC-Ergebnissen als bestes Modell für Szenario 1 ein Modell mit festen Effekten (Tabelle 4.3-194). Dies stimmte auch mit den Ergebnissen der Meta-Analyse für Tiotropiumbromid 18µg überein; hier ließ der Heterogenitätstest auf keine signifikante Heterogenität schließen (Abbildung 75). Szenario 2 (alle Studien) wurde nach 24 Wochen nicht durchgeführt, da kein Patient in den Studien, in denen über unerwünschte Ereignisse nach 24 Wochen berichtet wurde, eine Begleitmedikation mit LABA erhalten hat.

Tabelle 4.3-166: DIC des indirekten Vergleichs von Acclidiniumbromid und Tiotropiumbromid 5µg und 18µg (gepoolt); unerwünschte Ereignisse

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung für die NMA-Modelle
Woche 12			
Ohne LABA-Begleitmedikation	74,9	76,6	Feste Effekte
Alle Studien	86,321	87,246	Feste Effekte
Woche 24			
Ohne LABA-Begleitmedikation	48,0	48,3	Feste Effekte
Alle Studien	-	-	-

Die Ergebnisse der NMA werden in Tabelle 4.3-167, Tabelle 4.3-168 und Tabelle 4.3-169 dargestellt.

Tabelle 4.3-167: Unerwünschte Ereignisse nach 12 Wochen - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Komparator		
		PLA		TIO 5 und 18
TIO 5 und 18	Effektschätzer	0,83		
	95% KI	0,69	1,00	
	P(Überlegenheit)	97%		
Acidiniumbromid 400	Effektschätzer	0,89		1,08
	95% KI	0,66	1,18	0,75 1,50
	P(Überlegenheit)	81%		37%

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg; AB 400= Acidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-168: Unerwünschte Ereignisse nach 12 Wochen - Indirekter Vergleich aus RCT; Szenario 2: alle Studien: Modell mit festen Effekten

Intervention		Kontrolle		
		Pbo		TIO 5 und 18
TIO 5 und 18	Effektschätzer	0,82		
	95% KI	0,70	0,96	
	P(Überlegenheit)	99%		
AB 400	Effektschätzer	0,89		1,09
	95% KI	0,66	1,18	0,77 1,50
	P(Überlegenheit)	81%		33%

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg; AB 400= Acidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 12 Wochen war Acidiniumbromid mit Placebo [OR 0,9 (95% KI: 0,7, 1,2)] und mit Tiotropiumbromid [OR 1,1 (95% KI: 0,8, 1,5)] vergleichbar (Tabelle 4.3-167). Szenario 2 ergab ähnliche Ergebnisse für diese paarweisen Vergleiche (Tabelle 4.3-168). Die bei der NMA für den Vergleich von Acidiniumbromid und Placebo erhaltenen Ergebnisse sind mit den Ergebnissen der Meta-Analysen konsistent.

Tabelle 4.3-169: Unerwünschte Ereignisse nach 24 Wochen - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA: Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 18
TIO 18	Effektschätzer	1,06	
	95% KI	0,84	1,33
	P(Überlegenheit)	67%	
AB 400	Effektschätzer	0,88	
	95% KI	0,61	1,22
	P(Überlegenheit)	80%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen waren die unerwünschten Ereignisse unter Acclidiniumbromid und Placebo ebenfalls vergleichbar (Tabelle 4.3-169). Dies ist konsistent mit den Meta-Analyse Ergebnissen. Acclidiniumbromid und Tiotropiumbromid können im Hinblick auf unerwünschte Ereignisse als gleichwertig angesehen werden.

Die Raten unerwünschter Ereignisse mit Acclidiniumbromid und Placebo waren nach 12 und 24 Wochen vergleichbar (das 95 % Konfidenzintervall schloss die Eins ein). Im indirekten Vergleich waren die Raten unerwünschter Ereignisse mit Acclidiniumbromid und Tiotropiumbromid vergleichbar. Sowohl nach 12 Wochen als auch nach 24 Wochen lag die Eins innerhalb des 95 %-Konfidenzintervalls. Der Nutzen beider Therapien im Hinblick auf diesen Endpunkt ist deshalb gleichwertig.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.13 Schwere unerwünschte Ereignisse – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-170: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Schwere unerwünschte Ereignisse - 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I	•			•
	ACCORD II	•			•
1	Casaburi 2000			•	•
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Vier Studien berichten über die Anzahl der Patienten mit schweren unerwünschten Ereignissen nach 12 Wochen (LABAs waren nicht gestattet); zwei Acridiniumbromid-Studien, eine Tiotropiumbromid-18µg-Studie und eine Tiotropiumbromid-5µg-Studie. Abbildung 77 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 77: Netzwerk-Diagramm für die Analyse schwerwiegender unerwünschter Ereignisse nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

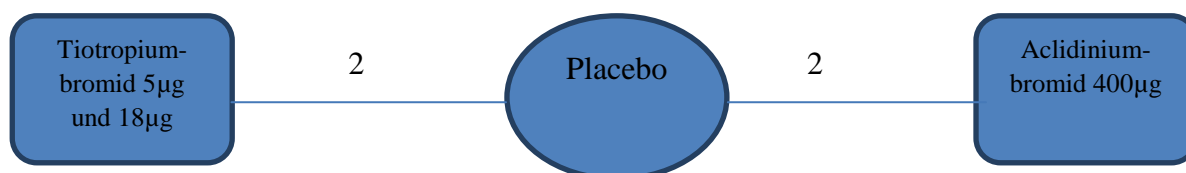


Tabelle 4.3-171: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Schwere unerwünschte Ereignisse - 12 Wochen - alle Studien]

Anzahl Studien	Referenzstudien	Acidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I ACCORD II	• •			• •
4	Casaburi 2000 Covelli 2005 Johansson 2008 SAFE Portugal			• • • •	• • • •
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Sieben Studien berichten über die Anzahl der Patienten mit schweren unerwünschten Ereignissen nach 12 Wochen; zwei Acidiniumbromid-Studien, vier Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 78 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 78: Netzwerk-Diagramm für die Analyse schwerer unerwünschter Ereignisse nach 12 Wochen

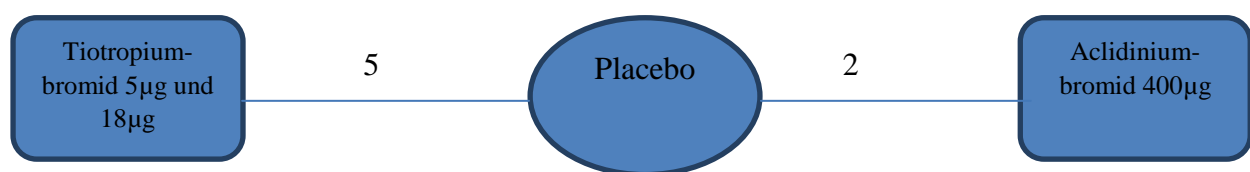


Tabelle 4.3-172: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Schwere unerwünschte Ereignisse - 24 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
1	INHANCE			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Zwei Studien berichten über die Anzahl der Patienten mit schweren unerwünschten Ereignissen nach 24 Wochen (LABAs waren nicht gestattet); eine Acridiniumbromid-Studie (ATTAIN) und eine Tiotropiumbromid-18µg-Studie (INHANCE). Keine Tiotropiumbromid-5µg-Studie berichtete über diesen Endpunkt. Abbildung 79 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 79: Netzwerk-Diagramm für die Analyse schwerer unerwünschter Ereignisse nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

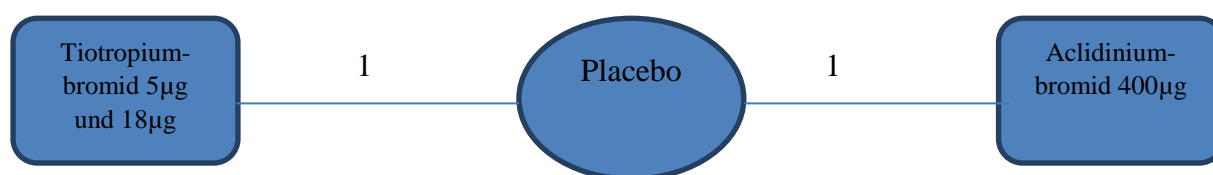


Tabelle 4.3-173: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Schwere unerwünschte Ereignisse - 24 Wochen – alle Studien]

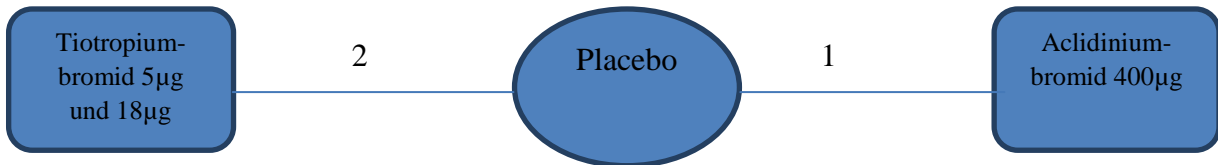
Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
2	INHANCE			•	•
	Niewoehner 2005			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Drei Studien berichten über die Anzahl der Patienten mit schweren unerwünschten Ereignissen nach 24 Wochen; eine Acridiniumbromid-Studie (ATTAIN) und zwei

Tiotropiumbromid-18 μ g-Studien (INHANCE, Niewoehner 2005). Keine Tiotropiumbromid-5 μ g-Studie berichtete über diesen Endpunkt. Abbildung 80 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 80: Netzwerk-Diagramm für die Analyse schwerer unerwünschter Ereignisse nach 24 Wochen



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-174: Operationalisierung von schweren unerwünschten Ereignissen

Studie	Operationalisierung
ACCORD I	<p>Ein SAE wurde definiert als eines der folgenden Ereignisse, bei irgendeiner Gabe des Medikaments:</p> <ul style="list-style-type: none"> • Tod • Lebensbedrohliche unerwünschte Medikamenten-Nebenwirkung • Stationärer Krankenhausaufenthalt oder Verlängerung eines bestehenden Krankenhausaufenthalts • Verursachung einer/eines angeborenen Anomalie/Geburtsdefekts • Wichtige medizinische Ereignisse, die nicht zum Tod führten, nicht lebensbedrohlich waren oder keinen Krankenhausaufenthalt erforderten, konnten als ernsthaft betrachtet werden, wenn sie, basierend auf einer entsprechenden medizinischen Beurteilung, den Patienten oder den Studiengegenstand gefährdeten und ein medizinisches oder chirurgisches Eingreifen zur Verhinderung eines der in dieser Definition verzeichneten Ergebnisse erforderten. Beispiele für solche medizinischen Ereignisse schließen allergische Bronchospasmen ein, die eine intensive Behandlung in einer Notaufnahme oder zu Hause erfordern, Blutkrankheiten oder Krämpfe, die nicht aus stationärer Behandlung resultieren, oder die Ausbildung einer Drogenabhängigkeit oder ein Drogenmissbrauch. Besuche in der Notaufnahme, die nicht zu einer Aufnahme ins Krankenhaus führten, wurden als anderes schwerwiegendes - in dieser Definition aufgeführtes - Ergebnis (z. B. lebensbedrohend, anderes schwerwiegendes [medizinisch bedeutendes] Ereignis) ausgewertet. <p>Folgende Ereignisse wurde nicht als SAE gewertet:</p> <ul style="list-style-type: none"> • Krankenhausaufenthalt für eine Behandlung/Operation, die ausgewählt oder für eine bereits bestehende Erkrankung vorgeplant war, die sich während der Teilnahme an der Studie nicht verschlechtert hat • Krankenhausaufenthalt oder Verlängerung eines bestehenden Krankenhausaufenthaltes für eine Kurzzeitpflege (z.B. wenn die Patienten zu weit entfernt vom Krankenhaus lebten oder keine Bezugsperson zu Hause hatten) <p>Bei jeder Visite wurden die Patienten zu allen Nebenwirkungen befragt, die seit der letzten Untersuchung aufgetreten sind.</p>
ACCORD II	<p>Ein SAE ist jedes ungünstige medizinische Ereignis, das bei irgendeiner Medikamentengabe:</p> <ul style="list-style-type: none"> • den Tod zur Folge hat • lebensbedrohlich ist • einen stationären Krankenhausaufenthalt oder die Verlängerung eines bestehenden Krankenhausaufenthalts verlangt • eine dauerhafte oder bedeutende Behinderung/Invalidität ergibt, oder • eine angeborene Anomalie/einen Geburtsdefekt verursachte • Wichtige medizinische Ereignisse, die nicht zum Tod führen, nicht lebensbedrohlich sind oder keinen Krankenhausaufenthalt erfordern, können als ernsthaft betrachtet werden, wenn sie, basierend auf einer entsprechenden medizinische Beurteilung, den Patienten oder den Studiengegenstand gefährden können und ein medizinisches oder chirurgisches Eingreifen zur Verhinderung eines der in dieser Definition verzeichneten Ergebnisse erfordern würden. Zu den Beispielen solcher medizinischen Ereignisse gehören allergische Bronchospasmen, die eine intensive Behandlung in einer Notaufnahme oder zu Hause erfordern, Blutkrankheiten oder Krämpfe, die keinen stationären Krankenhausaufenthalt bewirken, oder auf die Entwicklung einer Drogenabhängigkeit oder einen Drogenmissbrauch hinauslaufen. Besuche in der Notaufnahme, die keinen Krankenhausaufenthalt (d. h. Übernachtung) zur Folge haben, sollten als ernste Ergebnisse bewertet werden um herauszufinden, ob sie SAEs darstellen. Vorgeplante Krankenhausaufenthalte (d. h. elektive Eingriffe für vorher existierende Erkrankungen, die sich nicht verschlechterten) werden vom SAE-Bericht ausgeschlossen. <p>Bei jeder Untersuchung wurden die Patienten zu allen AEs/SAEs [schwerwiegenden AEs] befragt, die seit der letzten Untersuchung aufgetreten sind. Die Patienten wurden gebeten,</p>

	<p>nicht-suggestive Fragen zu beantworten. Das Personal des Studienzentrums zeichnet alle entsprechenden Informationen in den eCRFs [elektronischen Patienten-Prüfbogen] auf. Alle durch den Patienten (oder einen Patientenvertreter) berichteten oder beobachteten oder anderweitig durch die Untersuchungsperson (oder anderes Prüfpersonal) bei einer Prüfuntersuchung bemerkten oder während irgendeiner anderen Kommunikation mit dem Patienten (oder einem Patientenvertreter) außerhalb der festgelegten Prüfvisiten vorgekommenen AEs/SAEs (vom Zeitpunkt der Unterzeichnung der Einwilligungserklärung bis 30 Tage nach der letzten Gabe des Prüfmedikaments) müssen dokumentiert werden. Alle AEs mussten in einem entsprechenden AE-Berichts-Formular in der eCRF des Patienten erfasst werden, unabhängig davon ob sie in ursächlichem Zusammenhang mit dem Prüfpräparat stehen oder nicht. Die Zahl und der Prozentsatz von Patienten mit SAEs bei der Therapie wurden unter einem Fachbegriff und einer Behandlungsgruppe zusammengefasst.</p>
ATTAIN	<p>Ein SAE war jedes AE, das:</p> <ul style="list-style-type: none"> • tödlich war • lebensbedrohlich war. • eine stationäre Behandlung oder einen verlängerten bestehenden Krankenhausaufenthalt erforderte • zu einer bleibenden oder schwerwiegenden Behinderung oder Invalidität geführt hat • eine angeborene Anomalie/einen Geburtsdefekt verursachte • eine anderes medizinisch wichtiges Ereignis war, das den Patienten gefährdet haben könnte, oder einen Eingriffe zur Verhinderung eines der anderen oben genannten Ergebnisse erfordert hat. <p>Lebensbedrohlich in diesem Zusammenhang bedeutet, dass der an der Studie teilnehmende Patient zum Zeitpunkt des Ereignisses in unmittelbarer Todesgefahr schwebte. Das bedeutete nicht, dass das Ereignis hypothetisch den Tod verursacht haben könnte, wenn es schwerer war oder länger gedauert hat. Ein Krankenhausaufenthalt wurde als eine Übernachtung im Krankenhaus oder der Notaufnahme definiert. Eine Verlängerung des Krankenhausaufenthaltes wurde definiert als jede Verlängerung eines stationären Krankenhausaufenthaltes über den erwarteten/erforderlichen Aufenthalt bezogen auf den ursprünglichen Grund für die erstmalige Einweisung, die durch die Untersuchungsperson oder den behandelnden Arzt festgelegt wurde. Die Analysen von Safety-Outcomes wurden über die Safety-Population durchgeführt.</p>
Casaburi 2000	Ohne Angabe
Covelli 2005	Ohne Angabe
INHANCE	Ohne Angabe
Johansson 2008	Ohne Angabe
SAFE Portugal	Ohne Angabe
Niewoehner 2005	Schwerwiegende unerwünschte Ereignisse, einschließlich Todesfällen, wurden einer Behandlungsgruppe zugeordnet, wenn sie innerhalb von 30 Tagen nach der letzten Gabe der Studienmedikation aufgetreten sind.
Voshaar 2008	Ohne Angabe

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-175: Bewertung des Verzerrungspotenzials für schwere unerwünschte Ereignisse in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Casaburi 2000	Niedrig	Ja	Nein	Nein	Nein	Hoch
Covelli 2005	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
INHANCE	Niedrig	Nein	Ja	Nein	Nein	Hoch
Johansson 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
SAFE Portugal	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Niewoehner 2005	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Voshaar 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für ACCORD II wurde das Verzerrungspotenzial für das Outcome schwerwiegender unerwünschte Ereignisse als hoch angesehen (Tabelle 4.3-175), weil die Vergleichbarkeit der Behandlungsgruppen zu Studienbeginn nicht vollständig gegeben war. Alle Tiotropiumbromid-Studien hatten auf Studienebene ein geringes Verzerrungspotenzial. Allerdings hatten 2 von 7 Tiotropiumbromid-Studien ein hohes Verzerrungspotenzial speziell für den Endpunkt schwere unerwünschte Ereignisse, da für die Endpunkt-Analyse nicht das ITT-Prinzip verwendet wurde oder die Endpunkterhebung nicht verblindet war.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-176: Ergebnisse für schwere unerwünschte Ereignisse aus RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ACCORD I	Aclidiniumbromid	190	6 (3,7%)	1,48 (0,66) [0,41; 5,35] p=0,550
	Placebo	186	4 (2,2%)	
ACCORD II	Aclidiniumbromid	177	8 (4,5%)	0,67 (0,47) [0,27; 1,68] p=0,39
	Placebo	182	12 (6,6%)	
SAFE Portugal	Tiotropiumbromid 18µg	147	6 (4,1%)	2,28 (0,72) [0,56; 9,30] p=0,25
	Placebo	164	3 (1,8%)	
Covelli 2005	Tiotropiumbromid 18µg	94	5 (5,3%)	0,89 (0,65) [0,25; 3,18] p=0,857
	Placebo	84	5 (5,9%)	
Casaburi 2000	Tiotropiumbromid 18µg	276	19 (6,8%)	1,00 (0,37) [0,48; 2,07] p=1,0
	Placebo	188	13 (6,8%)	
Johansson 2008	Tiotropiumbromid 18µg	107	3 (2,8%)	3,35 (1,16) [0,34; 32,67] p=0,299
	Placebo	117	1 (0,9%)	
Voshaar 2008	Tiotropiumbromid 5µg	180	1 (0,7%)	1,00 (1,42) [0,06; 16,11] p=1,0
	Placebo	181	1 (0,6%)	

N=Anzahl der Patienten SE= Standardfehler, ITT= Intention-to-treat-Population [KI]=Konfidenzintervall

Tabelle 4.3-176 zeigt die Ergebnisse der Studien mit Aclidiniumbromid und Tiotropiumbromid 5µg und 18µg für den Endpunkt schwere unerwünschte Ereignisse nach 12 Wochen. Die Anzahl von Ereignissen war in fast allen Studien gering, was die statistische Aussagekraft begrenzte. Der Prozentsatz der Ereignisse unterschied sich beträchtlich, sowohl zwischen den Aclidiniumbromid-Studien wie auch zwischen den Tiotropiumbromid-Studien.

Tabelle 4.3-177: Ergebnisse für schwere unerwünschte Ereignisse aus RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ATTAIN	Aclidiniumbromid	269	14 (5,2%)	0,94 (0,38) [0,45; 2,00] p= 0,88
	Placebo	273	15 (5,5%)	
Niewoehner 2005	Tiotropiumbromid 18µg	914	162 (17,7%)	1,05 (0,12) [0,82; 1,33]
	Placebo	915	156 (17,0%)	
INHANCE	Tiotropiumbromid 18µg	415	34 (8,2%)	0,98 (0,25) [0,60; 1,60] p=0,936
	Placebo	418	35 (8,4%)	

N=Anzahl der Patienten SE= Standardfehler, ITT= Intention-to-treat-Population [KI]=Konfidenzintervall

Tabelle 4.3-177 zeigt die Ergebnisse der Studien mit Aclidiniumbromid und Tiotropiumbromid 5µg und 18µg für den Endpunkt schwere unerwünschte Ereignisse nach 24 Wochen. Die Anzahl der Ereignisse war in fast allen Studien gering und mit Placebo vergleichbar. Da nur eine Aclidiniumbromid-Studie über diesen Endpunkt berichtete, war für Aclidiniumbromid keine Meta-Analyse möglich.

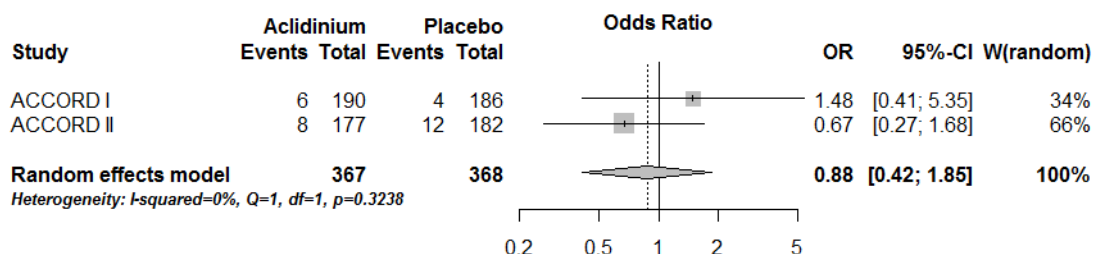


Abbildung 81: Meta-Analyse für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Aclidiniumbromid im Vergleich zu Placebo

Die Meta-Analyse (Abbildung 81) zeigt, dass die Inzidenz schwerer unerwünschter Ereignisse zwischen dem Aclidiniumbromid- und Placeboarm vergleichbar war (gepoolte OR 0,88 [95%-KI 0,42; 1,85]). Es bestand keine Studienheterogenität (Q=1; p=0,32; I²=0%).

Unter Ausschluss der ACCORD II-Studie kann die Sensitivitätsanalyse nicht durchgeführt werden, da nur zwei Aclidiniumbromid-Studien diesen Endpunkt berichten.

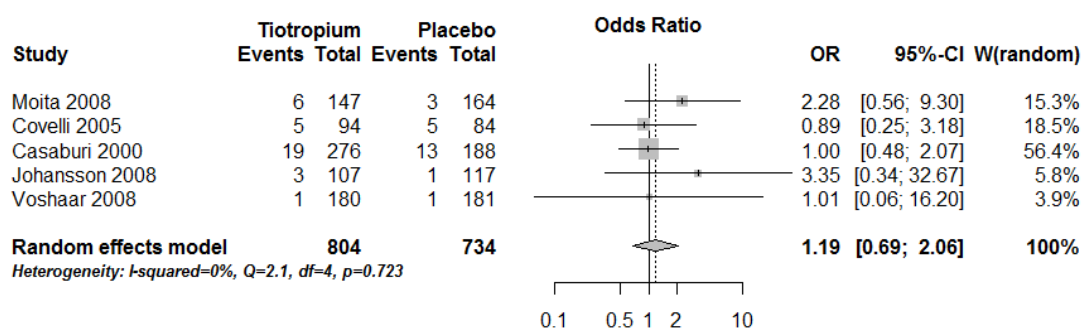


Abbildung 82: Meta-Analyse für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Tiotropiumbromid 5µg und 18µg im Vergleich zu Placebo

Numerisch wurden schwere unerwünschte Ereignisse bei mehr Tiotropium- als Placebo-Patienten dokumentiert. Der Unterschied war jedoch nicht signifikant (Abbildung 82; gepoolte OR 1,19 [95%-KI 0,69; 2,06]). Der Heterogenitätstest ergab keine Studienheterogenität für diesen Endpunkt ($Q=2,1$; $p=0,72$; $I^2=0\%$).

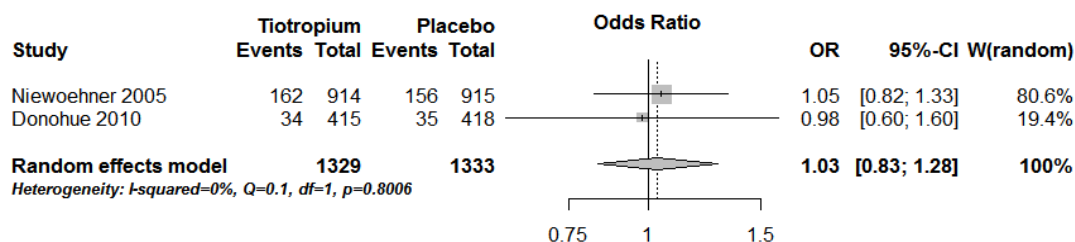


Abbildung 83: Meta-Analyse für schwere unerwünschte Ereignisse (24 Wochen) aus RCT; Tiotropiumbromid 18µg im Vergleich zu Placebo.

Abbildung 83 zeigt, dass Tiotropiumbromid für diesen Endpunkt mit Placebo vergleichbar ist (gepoolte OR 1,03 [KI 95% 0,83; 1,28]). Es lag keine Studien-Heterogenität vor ($Q=0,1$; $p=0,80$; $I^2=0\%$).

Sensitivitätsanalyse

Um den Einfluss des Schweregrads der Erkrankung auf den gepoolten Behandlungseffekt zu ermitteln, wurde eine Studie mit nahezu ausschließlicher Behandlung von Patienten mit mittelschwerer COPD (Johansson 2008) ausgeschlossen. Die Ergebnisse dieser Analyse sind in Abbildung 84 aufgeführt.

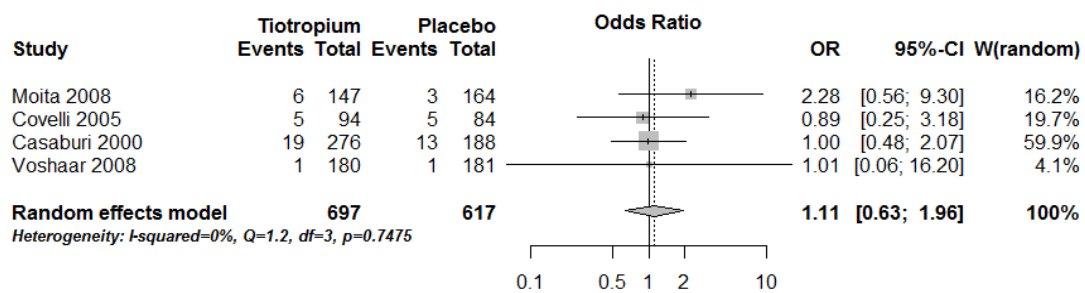


Abbildung 84 : Sensitivitätsanalyse (COPD-Schweregrad) für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Tiotropiumbromid 5µg und 18µg OD im Vergleich zu Placebo.

Abbildung 84 zeigt, dass die Inzidenz schwerer unerwünschter Ereignisse mit Tiotropiumbromid auch bei Ausschluss der Studie, die überwiegend Patienten mit mittelschwerer COPD einschloss, mit Placebo vergleichbar ist (gepoolte OR 1,11 [KI 95% 0,63; 1,96]). Das Ergebnis ist dem Ergebnis der Meta-Analyse ähnlich, die Johansson 2008 einschloss. Somit hatte der Schweregrad der Erkrankung keine Auswirkungen auf den gepoolten Behandlungseffekt für diesen Endpunkt. Studienheterogenität bestand nicht ($Q=1,2$; $p=0,75$; $I^2=0\%$).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-178 zeigt die DIC-Schätzwerte für die geplanten NMAs. Nach 12 Wochen war das beste durch DIC-Ergebnisse ermittelte Modell ein Modell mit festen Effekten sowohl für Szenario 1 als auch für Szenario 2 (Tabelle 4.3-178). Die Verwendung eines Modells mit festen Effekten war konsistent mit den Befunden der Meta-Analyse von Tiotropiumbromid 5µg bzw. 18µg und Acclidiniumbromid, wobei keine dieser beiden Meta-Analysen auf eine signifikante Heterogenität schließen ließ.

Nach 24 Wochen ergab sich aus den DIC-Ergebnissen als bestes Modell entweder ein Modell mit festen oder zufälligen Effekten für Szenario 1 und ein Modell mit festen Effekten für

Szenario 2 (Tabelle 4.3-178). Da sich durch den Heterogenitätstest keine signifikante Heterogenität in der Tiotropium-18µg-Meta-Analyse ergab, wurde für die NMA auch ein Modell mit festen Effekten verwendet.

Tabelle 4.3-178: DIC des indirekten Therapievergleichs von Acridiniumbromid und Tiotropiumbromid 5µg und 18µg (gepoolt); Unterschied des mittleren CFB (ml) bei schwerwiegenden unerwünschten Ereignissen.

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung für die NMA-Modelle
Woche 12			
Ohne LABA-Begleitmedikation	41,124	42,452	Feste Effekte
Alle Studien	67,758	69,356	Feste Effekte
Woche 24			
Ohne LABA-Begleitmedikation	27,6	27,6	Feste Effekte
Alle Studien	43,0	44,6	Feste Effekte

Die Ergebnisse der NMA werden in Tabelle 4.3-179, Tabelle 4.3-180, Tabelle 4.3-181 und Tabelle 4.3-182 dargestellt.

Tabelle 4.3-179: Schwere unerwünschte Ereignisse nach 12 Wochen - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 AND 18
TIO 5 und 18	Effektschätzer	1,08	
	95% KI	0,50	2,09
	P(Überlegenheit)	50%	
AB 400	Effektschätzer	0,94	
	95% KI	0,41	1,85
	P(Überlegenheit)	36%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg; AB 400= Acridiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-180: Schwere unerwünschte Ereignisse nach 12 Wochen - Indirekter Vergleich aus RCT; Szenario 2: alle Studien: Modell mit festen Effekten

Intervention		Kontrolle		
		Pbo		TIO 5 und 18
TIO 5 und 18	Effektschätzer	1,28		
	95% KI	0,71	2,14	
	P(Überlegenheit)	77%		
AB 400	Effektschätzer	0,94		0,80
	95% KI	0,41	1,86	0,28 1,82
	P(Überlegenheit)	37%		24%

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 12 Wochen war Acclidiniumbromid mit Placebo [OR 0,9 (95% KI: 0,4, 1,9)] und mit Tiotropiumbromid vergleichbar [OR 1,0 (95% KI: 0,3, 2,4)] (Tabelle 4.3-179). Szenario 2 (alle Studien) ergab ähnliche Ergebnisse (Tabelle 4.3-180). Die bei der NMA für den Vergleich von Tiotropiumbromid bzw. Acclidiniumbromid und Placebo erhaltenen Ergebnisse stimmen mit den Ergebnissen der Meta-Analysen überein. Beide Behandlungen waren mit einer mit Placebo vergleichbaren Anzahl schwerwiegender unerwünschter Ereignisse assoziiert (Tabelle 4.3-179, Abbildung 81, Abbildung 82).

Tabelle 4.3-181: Schwere unerwünschte Ereignisse nach 24 Wochen - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA: Modell mit festen Effekten

Intervention		Komparator		
		PLA		TIO 18
TIO 18	Effektschätzer	1,01		
	95% KI	0,59	1,60	
	P(Überlegenheit)	46%		
Acclidiniumbromid 400µg	Effektschätzer	1,01		1,07
	95% KI	0,44	2,02	0,39 2,38
	P(Überlegenheit)	44%		47%

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-182: Schwere unerwünschte Ereignisse nach 24 Wochen - Indirekter Vergleich aus RCT; Szenario 2: alle Studien: Modell mit festen Effekten

Intervention		Komparator	
		Pbo	TIO 18
TIO 18	Effektschätzer	1,04	
	95% KI	0,83	1,29
	P(Überlegenheit)	62%	
Acidiniumbromid 400µg	Effektschätzer	1,02	
	95% KI	0,44	2,01
	P(Überlegenheit)	44%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen war Acidiniumbromid in beiden Szenarien ebenfalls mit Placebo und Tiotropiumbromid 18µg vergleichbar (Tabelle 4.3-181 und Tabelle 4.3-182). In beiden Analysen (NMA Szenario 1, NMA Szenario 2 und Meta-Analyse [Abbildung 83]) war die Inzidenz schwerer unerwünschter Ereignisse mit Tiotropiumbromid mit der Inzidenz unter Placebo vergleichbar.

Die Raten schwerer unerwünschter Ereignisse mit Acidiniumbromid bzw. Tiotropiumbromid und Placebo waren nach 12 und 24 Wochen vergleichbar (das 95 % Konfidenzintervall schloss die Eins ein). Im indirekten Vergleich waren die Raten schwerer unerwünschter Ereignisse unter Acidiniumbromid und Tiotropiumbromid vergleichbar. Sowohl nach 12 Wochen als auch nach 24 Wochen lag die Eins innerhalb des 95 %-Konfidenzintervalls. Der Nutzen beider Therapien im Hinblick auf diesen Endpunkt ist deshalb gleichwertig.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.14 Studienabbrüche aufgrund unerwünschter Ereignisse – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-183: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Ausschlüsse aufgrund von AEs – 12 Wochen - begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I	•			•
	ACCORD II	•			•
2	Casaburi 2000			•	•
	Verkindre 2006			•	•
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Fünf Studien berichteten über die Anzahl der Patienten, die wegen unerwünschter Ereignisse nach 12 Wochen (LABAs waren nicht gestattet) aus der Studie ausgeschlossen wurden; zwei Acidiniumbromid-Studien, zwei Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 85 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 85: Netzwerk-Diagramm für die Analyse von Ausschlüssen wegen unerwünschter Ereignisse nach 12 Wochen (begleitende Behandlung mit LABAs nicht gestattet)

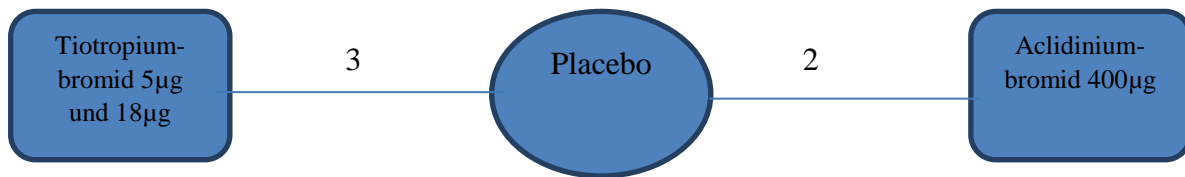


Tabelle 4.3-184: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Ausschlüsse aufgrund von AEs – 12 Wochen – alle Studien]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I ACCORD II	• •			• •
5	Casaburi 2000 SPRUCE Johansson 2008 SAFE Portugal Verkindre 2006			• • • • •	• • • • •
1	Voshaar 2008		•		•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Acht Studien berichteten über die Anzahl der Patienten, die wegen unerwünschter Ereignisse nach 12 Wochen aus der Studie ausgeschlossen wurden; zwei Aclidiniumbromid-Studien, fünf Tiotropiumbromid-18µg-Studien und eine Tiotropiumbromid-5µg-Studie. Abbildung 86 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 86: Netzwerk-Diagramm für die Analyse von Ausschlüssen wegen unerwünschter Ereignisse nach 12 Wochen

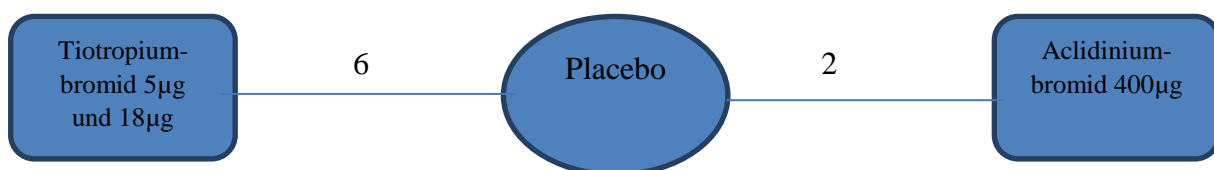


Tabelle 4.3-185: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Ausschlüsse aufgrund von AEs – 24 Wochen – begleitende Behandlung mit LABAs nicht gestattet]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
3	INHANCE			•	•
	Donohue 2002			•	•
	Vogelmeier 2008			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Vier Studien – die die begleitende Behandlung mit LABAs nicht gestatten - berichteten über die Anzahl der Patienten, die nach 24 Wochen wegen unerwünschter Ereignisse von der Studie ausgeschlossen wurden; eine Acridiniumbromid-Studie (ATTAIN) und drei Tiotropiumbromid-18µg-Studien (Donohue 2002, INHANCE, Vogelmeier 2008). Keine Tiotropiumbromid-5µg-Studie berichtete über diesen Endpunkt. Abbildung 87 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 87: Netzwerk-Diagramm für die Analyse von Ausschlüssen wegen unerwünschter Ereignisse nach 24 Wochen (begleitende Behandlung mit LABAs nicht gestattet)



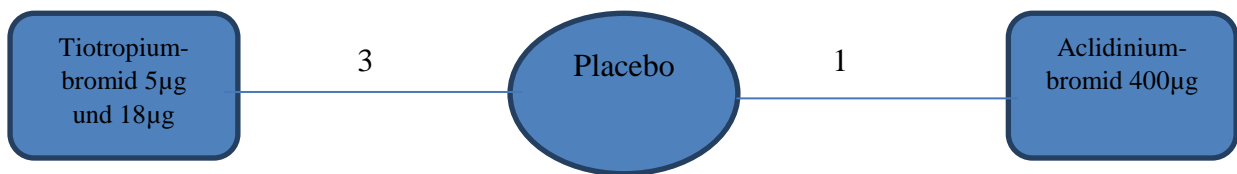
Tabelle 4.3-186: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Ausschlüsse aufgrund von AEs – 24 Wochen – alle Studien]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•
3	INHANCE			•	•
	Donohue 2002			•	•
	Vogelmeier 2008			•	•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Vier Studien (alle Studien, begleitende LABA-Therapie erlaubt) berichteten über die Anzahl der Patienten, die nach 24 Wochen wegen unerwünschter Ereignisse von der Studie ausgeschlossen wurden; eine Acclidiniumbromid-Studie (ATTAIN) und drei Tiotropiumbromid-18 μ g-Studien (Donohue 2002, INHANCE, Vogelmeier 2008). Keine Tiotropiumbromid-5 μ g-Studie berichtete über dieses Outcome. Abbildung 88 zeigt das Netzwerk, das auf der Grundlage dieser Studien erzeugt werden kann.

Abbildung 88: Netzwerk-Diagramm für die Analyse von Ausschlüssen wegen unerwünschter Ereignisse nach 24 Wochen



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-187: Operationalisierung von Studienabbrüchen aufgrund unerwünschter Ereignisse

Studie	Operationalisierung
ACCORD I	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
ACCORD II	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
ATTAIN	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
Casaburi 2000	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
Donohue 2002	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
INHANCE	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
SPRUCE	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
Johansson 2008	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
SAFE Portugal	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
Verkindre 2006	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
Vogelmeier 2008	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.
Voshaar 2008	Alle Patienten, die die Studienbehandlung wegen eines unerwünschten Ereignisses abbrachen.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-188: Bewertung des Verzerrungspotenzials für Ausschlüsse aufgrund von unerwünschten Ereignissen in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studien-ebene	Verblindung Endpunkt-erheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnis-gesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Casaburi 2000	Niedrig	Ja	Nein	Nein	Nein	Hoch
INHANCE	Niedrig	Nein	Ja	Nein	Nein	Hoch
Donohue 2002	Niedrig	Ja	Nein	Unklar	Nein	Hoch
SPRUCE	Hoch	Ja	Nein	Unklar	Nein	Hoch
Johansson 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
SAFE Portugal	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Verkindre 2006	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Vogelmeier 2008	Hoch	Nein	Ja	Nein	Nein	Hoch
Voshaar 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für ACCORD II wurde das Verzerrungspotenzial für diesen Endpunkt als hoch angesehen (Tabelle 4.3-188), weil die Vergleichbarkeit der Behandlungsgruppen zu Studienbeginn nicht vollständig gegeben war.

Alle Tiotropiumbromid-Studien außer SPRUCE und Vogelmeier 2008 hatten auf Studienebene ein geringes Verzerrungspotenzial. Wenn man allerdings das Verzerrungspotenzial speziell für den Endpunkt Studienabbruch aufgrund unerwünschter Ereignisse betrachtete, hatten 5 von 9 Tiotropiumbromid-Studien ein hohes Verzerrungspotenzial, weil das ITT-Prinzip nicht für die Endpunkt-Analyse verwendet wurde oder die Endpunkterhebung für die Behandlungsallokation nicht verblindet war.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-189: Ergebnisse für Studienabbrüche aufgrund von unerwünschten Ereignissen aus der RCT für indirekte Vergleiche (12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ACCORD I	Acidiniumbromid	190	8 (4.2%)	0.54 (0.46) [0.22; 1.32] p=0.178
	Placebo	186	14 (7.6%)	
ACCORD II	Acidiniumbromid	177	13 (7.4%)	1.72 (0.46) [0.70; 4.27] p=0.240
	Placebo	182	8 (4.4%)	
SAFE Portugal	Tiotropiumbromid 18µg	147	3 (2.0%)	1.69 (0.92) [0.28; 10.24]
	Placebo	164	2 (1.2%)	
SPRUCE	Tiotropiumbromid 18µg	200	8 (4.0%)	0.54 (0.46) [0.22; 1.31]
	Placebo	195	14 (7.2%)	
Verkindre 2006	Tiotropiumbromid 18µg	46	1 (2.2%)	0.18 (1.11) [0.02; 1.53] p=0.121
	Placebo	54	6 (11.1%)	
Casaburi 2000	Tiotropiumbromid 18µg	276	7 (2.5%)	0.42 (0.49) [0.16; 1.10]
	Placebo	188	11 (5.8%)	
Johansson 2008	Tiotropiumbromid 18µg	107	0 (0%)	0.21 (1.56) [0.01; 4.53] p=0.317
	Placebo	117	2 (1.7%)	
Voshaar 2008	Tiotropiumbromid 5µg	180	13 (7.2%)	0.80 (0.39) [0.37; 1.72] p= 0.57
	Placebo	181	16 (8.8%)	

N=Anzahl der Patienten, SE= Standardfehler ITT= Intention-to-treat-Population KI=Konfidenzintervall

Tabelle 4.3-189 zeigt die Anzahl von Patienten, die aufgrund unerwünschter Ereignisse nach 12 Wochen ausgeschlossen wurden. Die Zahl der ausgeschlossenen Patienten war gering, was die statistische Aussagekraft begrenzte.

Tabelle 4.3-190: Ergebnisse für Studienabbrüche aufgrund von unerwünschten Ereignissen aus der RCT für indirekte Vergleiche (24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ATTAIN	Acidiniumbreomid	269	8 (3,0%)	0,73 (0,47) [0,29; 1,84] p= 0,51
	Placebo	273	11 (4,0%)	
Vogelmeier 2008	Tiotropiumbromid 18µg	221	13 (5,9%)	1,57 (0,46) [0,64; 3,87] p=0,326
	Placebo	209	8 (3,8%)	
INHANCE	Tiotropiumbromid 18µg	415	17 (4,1%)	0,35 (0,29) [0,20; 0,63] p<0,001
	Placebo	418	45 (10,8%)	
Donohue 2002	Tiotropiumbromid 18µg	209	12 (5,7%)	0,25 (0,35) [0,13; 0,50]
	Placebo	201	39 (19,4%)	

N=Anzahl der Patienten, SE= Standardfehler ITT= Intention-to-treat-Population KI=Konfidenzintervall

Tabelle 4.3-190 zeigt die Ergebnisse der einzelnen Studien nach 24 Wochen für den Endpunkt Studienabbruchs wegen eines unerwünschten Ereignisses. Bis auf Vogelmeier 2008 wurden unter Studienmedikation weniger Studienabbrüche als unter Placebo beobachtet.

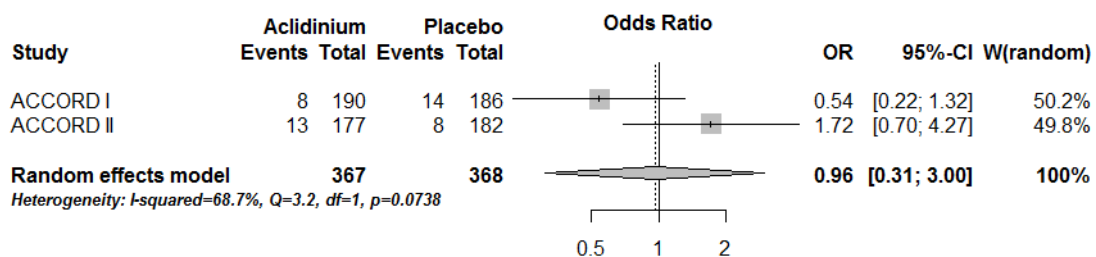


Abbildung 89: Meta-Analyse für Studienabbrüche aufgrund von unerwünschten Ereignissen (12 Wochen); Acidiniumbromid im Vergleich zu Placebo.

Abbildung 89 zeigt, dass Acidiniumbromid für diesen Endpunkt mit Placebo vergleichbar ist (gepoolte OR 0,96 [KI 95% 0,31; 3,00]). Da die Anzahl der Ereignisse in den einzelnen Studien klein war, sollte dieses Ergebnis mit Vorsicht interpretiert werden. Für diesen Endpunkt gab es eine signifikante Studienheterogenität (Q=3,2; p=0,07; I²=68,7%). Es konnte

allerdings keine Sensitivitätsanalyse durchgeführt werden, da nur zwei Studien über diesen Endpunkt berichteten.

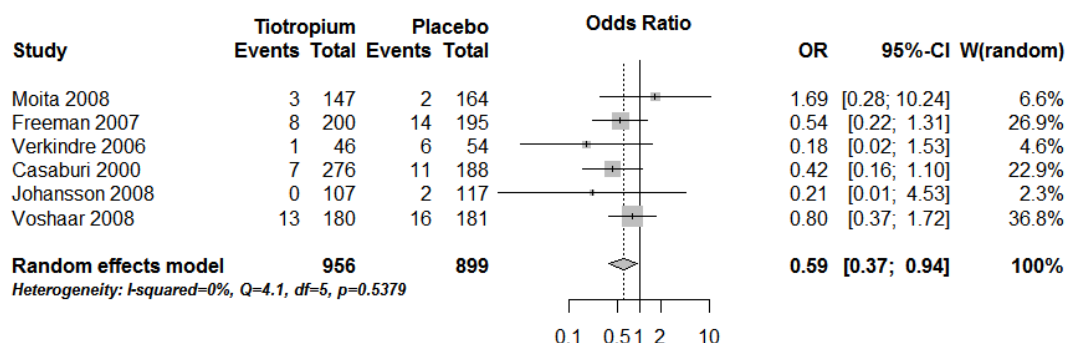


Abbildung 90: Meta-Analyse für Studienabbrüche aufgrund von unerwünschten Ereignissen (12 Wochen); Tiotropiumbromid 5µg im Vergleich zu Placebo

Abbildung 90 zeigt, dass Tiotropiumbromid die Zahl der Studienabbrüche wegen unerwünschter Ereignisse nach 12 Wochen im Vergleich zu Placebo reduziert (gepoolte OR 0,59 [KI 95% 0,37; 0,94]). Lediglich in einer Studie (SAFE Portugal) gab es in der Tiotropiumbromid-Gruppe mehr Studienabbrüchen aufgrund unerwünschter Ereignisse als in der Placebogruppe. Die Anzahl der Ereignisse war jedoch niedrig. Für den Engpunkt wurde keine Heterogenität festgestellt ($Q=4,1$; $p=0,54$; $I^2=0\%$).

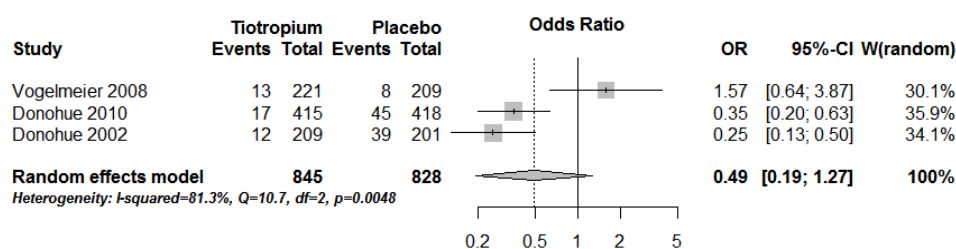


Abbildung 91: Meta-Analyse für Studienabbrüche aufgrund von unerwünschten Ereignissen (24 Wochen); Tiotropiumbromid 18µg im Vergleich zu Placebo

Abbildung 91 zeigt, dass Tiotropiumbromid mit einer mit Placebo vergleichbaren Zahl von Studienabbrüchen wegen unerwünschter Ereignisse nach 24 Wochen assoziiert ist (gepoolte OR 0,49 [KI 95% 0,19; 1,27]). Alle Tiotropium-Studien hatten jedoch ein hohes Verzerrungspotenzial für diesen Endpunkt und es gab eine signifikante Heterogenität zwischen den Studien ($Q=10,7$; $p=0,005$; $I^2=81,3\%$). Eine Sensitivitätsanalyse bezüglich des

Verzerrungspotenzials war nicht möglich, da alle für diesen Endpunkt herangezogenen Studien ein hohes Verzerrungspotenzial hatten.

Sensitivitätsanalyse

Um den Einfluss des Schweregrads der Erkrankung auf den gepoolten Behandlungseffekt für den Endpunkt „Studienabbrüche aufgrund unerwünschter Ereignisse“ nach 12 Wochen zu ermitteln, wurden zwei Studien mit nahezu ausschließlicher Behandlung von Patienten mit mittelschwerer COPD (Freeman 2007 (SPRUCE) und Johansson 2008) ausgeschlossen. Die Ergebnisse dieser Analyse sind in Abbildung 4 dargestellt.

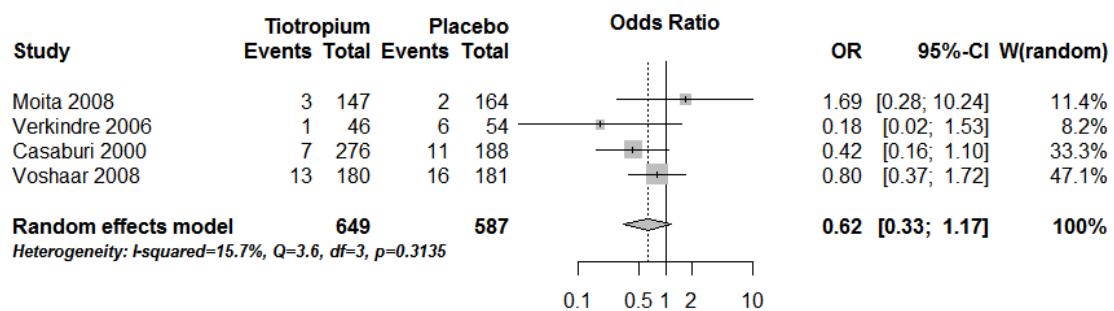


Abbildung 92: Sensitivitätsanalyse (COPD-Schweregrad) für Studienabbrüche aufgrund von unerwünschten Ereignissen (12 Wochen); Tiotropiumbromid 5µg im Vergleich zu Placebo

Wie Abbildung 92 zeigt, war die Tiotropiumbromid-Behandlung auch in der Sensitivitätsanalyse mit einer geringeren Inzidenz von Studienabbrüchen wegen unerwünschter Ereignisse nach 12 Wochen assoziiert als Placebo (gepoolte OR 0,62 [KI 95% 0,33; 1,17]). Der Gesamteffektschätzer ist jedoch mit einem breiteren Konfidenzintervall behaftet, was auf die Größe der ausgeschlossenen Studien und die geringe Anzahl von Ereignissen zurück zu führen ist. Von einem Einfluss des Schweregrades der Erkrankung auf den Behandlungseffekt von Tiotropiumbromid kann bei diesen Endpunkt nicht ausgegangen werden. Wie der Heterogenitätstest zeigte, gab es keine signifikante Heterogenität zwischen den Studien ($Q=3.6$; $p=0.31$; $I^2=15.7\%$).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*

- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Tabelle 4.3-191 zeigt die DIC-Schätzungen für die geplanten NMAs. Nach 12 Wochen war das beste durch DIC-Ergebnisse ermittelte Modell ein Modell mit zufälligen Effekten für Szenario 1 und ein Modell mit festen Effekten für Szenario 2. Das Modell mit zufälligen Effekten (Szenario 1) ergaben breite Konfidenzintervalle und dadurch wenig aussagekräftige Ergebnissen. Als Konsequenz daraus basierten die in Tabelle 4.3-192 und Tabelle 4.3-193 berichteten Ergebnisse auf einem Modell mit festen Effekten. Nach 24 Wochen war das beste durch DIC-Ergebnisse ermittelte Modell ein Modell mit zufälligen Effekten. Die so durchgeführte NMA führte aufgrund zu breiter Konfidenzintervallen zu wenig aussagekräftigen Ergebnissen. Als Konsequenz daraus basierten die in Tabelle 4.3-194 berichteten Ergebnisse auf einem Modell mit festen Effekten.

Tabelle 4.3-191: DIC des indirekten Vergleichs von Aclidiniumbromid und Tiotropiumbromid 5µg bzw. 18µg (gepoolt); Studienabbruch aufgrund unerwünschter Ereignisse

Szenario	Feste Effekte	Zufällige Effekte	Schlussfolgerung für die NMA-Modelle
Woche 12			
Ohne LABA-Begleitmedikation	58,810	58,396	Zufällige Effekte
Alle Studien	84,448	84,994	Feste Effekte
Woche 24			
Ohne LABA-Begleitmedikation	59,271	52,552	Zufällige Effekte
Alle Studien	-	-	-

Die Ergebnisse der NMA werden in Tabelle 4.3-192, Tabelle 4.3-193 und Tabelle 4.3-194 dargestellt. Szenario 2 wurde nach 24 Wochen nicht durchgeführt, da keine der Studien mit Daten über Ausschlüsse aufgrund unerwünschter Ereignisse nach 24 Wochen eine begleitende LABA-Therapie gestatteten.

Tabelle 4.3-192: Studienabbrüche aufgrund unerwünschter Ereignisse (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA; Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	0,57	
	95% KI	0,30	0,96
	P(Überlegenheit)	98%	
AB 400	Effektschätzer	1,01	
	95% KI	0,51	1,78
	P(Überlegenheit)	56%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Tabelle 4.3-193: Studienabbrüche aufgrund unerwünschter Ereignisse (12 Wochen) - Indirekter Vergleich aus RCT; Szenario 2: alle Studien: Modell mit festen Effekten

Intervention		Kontrolle	
		Pbo	TIO 5 und 18
TIO 5 und 18	Effektschätzer	0,58	
	95% KI	0,35	0,89
	P(Überlegenheit)	99%	
AB 400	Effektschätzer	1,00	
	95% KI	0,51	1,78
	P(Überlegenheit)	56%	

TIO 5= Tiotropiumbromid 5 µg; TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 12 Wochen war die Studienabbruchrate unter Acclidiniumbromid mit der unter Placebo [OR 1,0 (95% KI: 0,5, 1,8)] und der von Tiotropiumbromid vergleichbar [OR 1,93 (95% KI: 0,75 4,11)] (Tabelle 4.3-192). Die Anzahl von Studienabbrüchen aufgrund unerwünschter Ereignisse war allerdings in allen Studien sehr gering, was die Interpretation der Ergebnisse erschwert. In Szenario 2 war Acclidiniumbromid mit Placebo sowie mit Tiotropiumbromid vergleichbar (Tabelle 4.3-193).

Tabelle 4.3-194: Studienabbrüche aufgrund unerwünschter Ereignisse (24 Wochen) - Indirekter Vergleich aus RCT; Szenario 1: Studien ohne Begleitmedikation mit LABA: Modell mit festen Effekten

Intervention		Kontrolle			
		Pbo		TIO 18	
TIO 18	Effektschätzer	0,42			
	95% KI	0,28	0,60		
	P(Überlegenheit)	>99%			
AB 400	Effektschätzer	0,81		1,99	
	95% KI	0,27	1,83	0,62	4,77
	P(Überlegenheit)	75%		14%	

TIO 18= Tiotropiumbromid 18 µg; AB 400= Acclidiniumbromid 400 µg; Pbo= Placebo; KI= Konfidenzintervall, P= Wahrscheinlichkeit

Nach 24 Wochen war Acclidiniumbromid mit Placebo und Tiotropiumbromid [OR 1,99 (95% KI: 0.6, 4.8)] vergleichbar (Tabelle 4.3-194).

Die Ergebnisse der NMA für den paarweisen Vergleich von Acclidiniumbromid 400µg mit Placebo stimmen mit den Ergebnissen der Meta-Analysen überein, in den Acclidiniumbromid mit Placebo vergleichbar war. Die Meta-Analyse der Tiotropiumbromid-Studien ergab, dass Tiotropiumbromid 18µg hinsichtlich der Studienabbrüche aufgrund unerwünschter Ereignisse vergleichbar mit Placebo war, während Tiotropiumbromid 18µg in der NMA wirksamer als Placebo war (Tabelle 4.3-194). Der Effekt wird wahrscheinlich in der NMA überschätzt. Die enthaltenen Tiotropiumbromid-Studien berichten einen Schätzwert zugunsten von Tiotropiumbromid in Donohue 2002 und INHANCE und einen Schätzwert zugunsten von Placebo in Vogelmeier 2008. Da das Konfidenzintervall bei Donohue 2002 und INHANCE nicht mit dem von Vogelmeier 2008 überlappt, ist die Heterogenität hoch und Schlussfolgerungen sollten nur mit äußerster Vorsicht gezogen werden.

Die Raten der Studienabbrüche aufgrund unerwünschter Ereignisse mit Acclidiniumbromid und Placebo waren nach 12 und 24 Wochen vergleichbar (95 % Konfidenzintervall schloss die Eins ein). Im indirekten Vergleich waren die Raten der Studienabbrüche aufgrund unerwünschter Ereignisse mit Acclidiniumbromid und Tiotropiumbromid vergleichbar. Sowohl nach 12 Wochen als auch nach 24 Wochen lag die Eins innerhalb des 95 %-Konfidenzintervalls. Der Nutzen beider Therapien im Hinblick auf diesen Endpunkt ist deshalb vergleichbar.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.15 Kardiovaskuläre Morbidität und Mortalität – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-195: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Kardiovaskuläre Morbidität und Mortalität – 12 Wochen]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I ACCORD II	• •			• •
1	SAFE Portugal			•	•

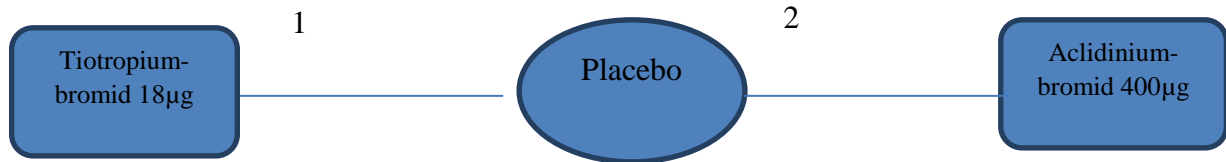
Tabelle 4.3-196: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Kardiovaskuläre Morbidität und Mortalität – 24 Wochen]

Anzahl Studien	Referenzstudien	Aclidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•			•

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Eine Tiotropiumbromid-18µg-Studie (SAFE Portugal) und zwei Aclidiniumbromid-Studien (ACCORD I, ACCORD II) berichteten Daten für kardiovaskuläre Morbidität und Mortalität nach 12 Wochen. Daraus kann folgendes (Abbildung 93) Netzwerk erstellt werden.

Abbildung 93: Netzwerk-Diagramm für die Analyse kardiovaskulärer Morbidität und Mortalität nach 12 Wochen



Nach 24 Wochen berichtete nur eine Studie mit Acridiniumbromid (ATTAIN) Daten über kardiovaskuläre Morbidität und Mortalität. Daher konnte nach 24 Wochen kein indirekter Vergleich mit Tiotropiumbromid durchgeführt werden und auch kein Netzwerk-Diagramm aufgezeigt werden.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-197: Operationalisierung von kardiovaskulärer Morbidität und Mortalität

Studie	Operationalisierung
ACCORD I	Herztod und schwerwiegende unerwünschte Ereignisse, die als kardiale (einschließlich kongestiver Herzinsuffizienz, Hypotonie) oder Gefäßerkrankungen (einschließlich Schlaganfall) eingestuft wurden.
ACCORD II	Herztod und schwerwiegende unerwünschte Ereignisse, die als kardiale (einschließlich Herz-Atemstillstand, akuter Myokardinfarkt, Angina pectoris, Vorhofflimmern) oder Gefäßerkrankungen (einschließlich Aorten-Aneurysma, transitorische ischämische Attacke) eingestuft wurden.
ATTAIN	Herztod und schwerwiegende unerwünschte Ereignisse, die als kardiale (einschließlich Myokardischämie, Angina pectoris, kardiale Beschwerden, akute Herzinsuffizienz, kongestive Herzinsuffizienz, Sick-Sinus-Syndrom, linksventrikuläre Herzinsuffizienz, Myokardinfarkt) oder Gefäßerkrankungen (einschließlich Arteriosklerose, Aortenstenose, Stenose der Arteria subclavia) eingestuft wurden.
SAFE Portugal	Anzahl der Patienten, die über als Herzerkrankungen eingestufte unerwünschte Ereignisse berichtet haben. Die Mortalität wird nicht nach Organ-Klassen angegeben und ist daher nicht Bestandteil dieses Ergebnisses.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-198: Bewertung des Verzerrungspotenzials für für kardiovaskuläre Morbidität und Mortalität in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
SAFE Portugal	Niedrig	Ja	Ja	Unklar	Ja	Hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

In einer der drei Acridiniumbromid-Studien (ACCORD II) wurde das Verzerrungspotenzial für das Outcome "kardiovaskuläre Morbidität und Mortalität" als hoch angesehen. Obwohl das Randomisierungsverfahren und die Behandlungsallokation dieser Studie etablierten Methoden entsprachen, war die Vergleichbarkeit der Baseline-Eigenschaften der Patienten in den zwei Behandlungsarmen mäßig (Tabelle 4-16; Details sind in Anhang 4-G beschrieben).

Die einzige Tiotropiumbromid-Studie, die über dieses Outcome berichtete, war SAFE Portugal. Diese Studie erzielte auf Studienebene ein geringes Verzerrungspotenzial. Speziell hinsichtlich des Verzerrungspotenzials für kardiovaskuläre Morbidität und Mortalität hatte diese Studie ein hohes Verzerrungspotenzial (Details sind in Anhang 4-G beschrieben).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-199: Ergebnisse für kardiovaskuläre Morbidität und Mortalität aus RCT für indirekte Vergleiche (CFB nach 12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ACCORD I	Aclidiniumbromid	190	2 (1,05%)	1,96 (1,23) [0,18; 21,77] p= 0,58
	Placebo	185	1 (0,54%)	
ACCORD II	Aclidiniumbromid	177	3 (1,69%)	1,03 (0,82) [0,21; 5,17] p= 0,97
	Placebo	182	3 (1,65%)	
SAFE Portugal	Tiotropiumbromid 18µg	144	2 (1,39%)	2,24 (1,23) [0,20; 24,96] p= 0,51
	Placebo	160	1 (0,63%)	

N=Anzahl der Patienten, SE= Standardfehler ITT= Intention-to-treat-Population KI=Konfidenzintervall

Tabelle 4.3-199 stellt die Ergebnisse der Aclidiniumbromid- und Tiotropiumbromid-Studien dar, die über das Outcome kardiovaskulärer Morbidität und Mortalität nach einer Behandlung von 12 Wochen berichteten. Die Anzahl von Ereignissen war gering, was die statistische Aussagekraft begrenzte. Eine auf den Tiotropiumbromid-Studien basierende Meta-Analyse war nicht möglich, weil über dieses Outcome nur in einer Studie berichtet wurde.

Tabelle 4.3-200: Ergebnisse für kardiovaskuläre Morbidität und Mortalität aus RCT für indirekte Vergleiche (CFB nach 24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ATTAIN	Aclidiniumbromid	269	7 (2,6%)	3,60 (0,81) [0,75; 18,6] p=0,11
	Placebo	273	2 (0,73%)	

N=Anzahl der Patienten, SE= Standardfehler ITT= Intention-to-treat-Population KI=Konfidenzintervall

Tabelle 4.3-200 stellt die Ergebnisse von ATTAIN dar, der einzigen Studie, die nach 24 Wochen Behandlung über kardiovaskuläre Morbidität und Mortalität berichtet. Die Anzahl der Patienten mit einem kardiovaskulären Ereignis war im Aclidiniumbromid-Arm höher. Die statistische Aussagekraft wurde jedoch von der geringen Anzahl von Patienten begrenzt. Daher sollten dieses Ergebniss mit Vorsicht interpretiert werden.

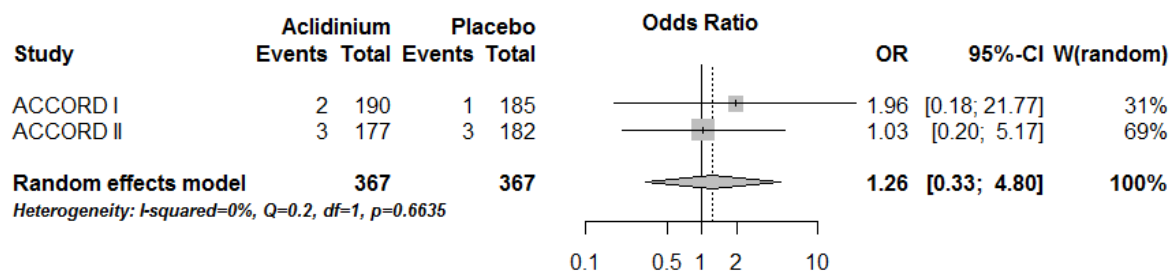


Abbildung 94: Meta-Analyse für die Anzahl von Patienten mit kardiovaskulärer Morbidität und Mortalität (12 Wochen); Acridiniumbromid im Vergleich zu Placebo

Abbildung 94 zeigt die Ergebnisse der Meta-Analyse unter Berücksichtigung von zwei Acridiniumbromid-Studien. Die Anzahl der Ereignisse war sehr niedrig, was die statistische Aussagekraft begrenzte (und große Konfidenzintervalle bewirkte). Für dieses Outcome gab es keine signifikante Heterogenität in der Studie ($Q=0,2$; $p=0,66$; $I^2=0,0\%$).

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Wegen der sehr geringen absoluten Anzahl der Ereignisse und des hohen Verzerrungsrisikos bei kleiner absoluter Zahl von Ereignissen wurde nach 12 Wochen kein indirekter Vergleich durchgeführt. Nach 24 Wochen war kein indirekter Vergleich möglich, da keine Tiotropiumstudie den Endpunkt Kardiovaskuläre Morbidität und Mortalität nach 24 Wochen berichtete.

In allen Studien waren Ereignisse, die dem Endpunkt kardiovaskuläre Morbidität zugeordnet werden konnten selten. Die Zahl der Ereignisse nach 12 Wochen war in beiden Acridiniumbromid-Studien (ACCORD I, ACCORD II) und der Tiotropium-Studie (SAFE Portugal) mit Placebo vergleichbar (die Eins lag innerhalb des 95 % Konfidenzintervalls).

Wegen der geringen Zahl von Ereignissen und des damit verbundenen hohen Verzerrungsrisikos wurde kein indirekter Vergleich durchgeführt.

Nach 24 Studien berichtete nur eine Acridiniumbromid-Studie diesen Endpunkt. Die Zahl der Ereignisse war klein und in beiden Behandlungsgruppen vergleichbar (die Eins lag innerhalb des 95 % Konfidenzintervalls).

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.16 Letalität und Gesamtmortalität – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche.

Tabelle 4.3-201: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Letalität und Gesamtmortalität (12 Wochen)]

Anzahl Studien	Referenzstudien	Acidiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
2	ACCORD I ACCORD II	• •			• •
2	Covelli 2005 SAFE Portugal			• •	• •
1	Voshaar 2008		•		•

Insgesamt 5 Studien berichteten Daten für diesen Endpunkt nach 12 Wochen; zwei Studien zu Acidiniumbromid (ACCORD I, ACCORD II), zwei Studien über Tiotropiumbromid 18µg (Covelli 2005, SAFE Portugal) und eine Studie über Tiotropiumbromid 5µg (Voshaar 2008).

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Abbildung 95: Netzwerk-Diagramm für die Analyse der Letalität und der Gesamt-Mortalität nach 12 Wochen



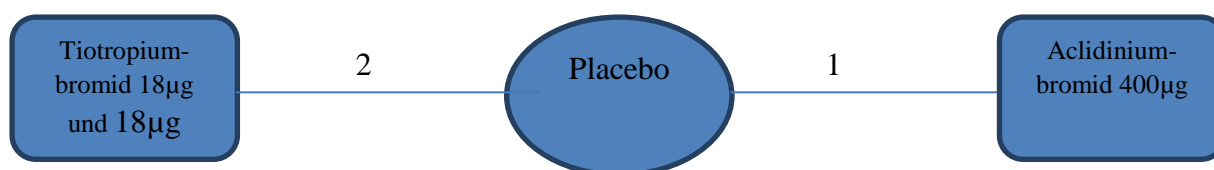
Tabelle 4.3-202: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden [Letalität und Gesamtmortalität (24 Wochen)]

Anzahl Studien	Referenzstudien	Acridiniumbromid 400 µg BID	Tiotropiumbromid 5µg; OD	Tiotropiumbromid 18µg; OD	Placebo
1	ATTAIN	•	Vogelmeier 2008		•
2	INHANCE Vogelmeier 2008			• •	• •

Die ATTAIN-Studie war die einzige Acridiniumbromid-Studie, die Daten für den Zeitpunkt 24 Wochen berichtete. Für Tiotropiumbromid 18µg berichteten zwei Studien (INHANCE, Vogelmeier 2008) Daten.

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Abbildung 96: Netzwerk-Diagramm für die Analyse der Letalität und der Gesamt-Mortalität nach 24 Wochen



Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-203: Operationalisierung von Letalität und Gesamtmortalität

Studie	Operationalisierung
ACCORD I	Sterblichkeit während der Doppelblind-Behandlungsperiode
ACCORD II	Sterblichkeit während der Doppelblind-Behandlungsperiode.
ATTAIN	Sterblichkeit während der Doppelblind-Behandlungsperiode.
Covelli 2005	Nicht beschrieben.
INHANCE	Nicht beschrieben.
SAFE Portugal	Nicht beschrieben.
Vogelmeier 2008	Nicht beschrieben.
Voshaar 2008	Nicht beschrieben.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-204: Bewertung des Verzerrungspotenzials für Letalität und Gesamtmortalität in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotenzial beeinflussende Punkte	Verzerrungspotenzial Endpunkt
ACCORD I	Niedrig	Ja	Ja	Nein	Nein	Niedrig
ACCORD II	Niedrig	Ja	Ja	Nein	Ja	Hoch
ATTAIN	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Covelli 2005	Niedrig	Ja	Unklar	Nein	Nein	Niedrig
INHANCE	Niedrig	Unklar	Unklar	Nein	Nein	Niedrig
SAFE Portugal	Niedrig	Ja	Ja	Nein	Nein	Niedrig
Vogelmeier 2008	Hoch	Unklar	Ja	Unklar	Nein	Niedrig
Voshaar 2008	Niedrig	Ja	Unklar	Nein	Nein	Niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In einer der drei Acclidiniumbromid-Studien (ACCORD II) wurde das Verzerrungspotenzial für das Outcome "Letalität und Gesamt-Mortalität" als hoch angesehen. Obwohl das Randomisierungsverfahren und die Behandlungsallokation dieser Studie etablierten Methoden entsprachen, war die Vergleichbarkeit der Baseline-Eigenschaften der Patienten in den zwei Behandlungsarmen mäßig (Tabelle 4.3-204; Details sind in Anhang 4-G beschrieben).

Die Tiotropiumbromid-Studien, die über dieses Outcome berichteten, wurden mit einem geringen Verzerrungspotenzial für das Outcome "Letalität und Gesamt-Mortalität" bewertet. Wenn man aber das studienbezogene Verzerrungspotenzial betrachtet, hatte die Studie von Vogelmeier 2008 ein hohes Verzerrungspotenzial (Details sind in Anhang 4-G beschrieben).

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4.3-205: Ergebnisse für Letalität und Gesamt-Mortalität aus RCT für indirekte Vergleiche (CFB nach 12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ACCORD I	Aclidiniumbromid	190	1 (0,53%)	-
	Placebo	185	0	
ACCORD II	Aclidiniumbromid	177	1 (0,56%)	1,03
	Placebo	182	1 (0,55%)	
Covelli 2005	Tiotropiumbromid	94	0 (0%)	-
	18µg	84	0 (0%)	
	Placebo			
SAFE Portugal	Tiotropiumbromid	144	2 (1,39%)	-
	18µg	160	0 (0%)	
	Placebo			
Voshaar 2008	Tiotropiumbromid	175	0 (0%)	-
	18µg	171	0 (0%)	
	Placebo			

N=Anzahl der Patienten, SE= Standardfehler ITT= Intention-to-treat-Population KI=Konfidenzintervall

Tabelle 4.3-206: Ergebnisse für Letalität und Gesamt-Mortalität aus RCT für indirekte Vergleiche (CFB nach 24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Veränderung
		N	N (%)	Odds Ratio (SE) [KI] p-Wert
ATTAIN	Aclidiniumbromid	269	1 (0,37%)	-
	Placebo	273	0	
INHANCE	Tiotropiumbromid 18µg	415	2 (0,48%)	-
	Placebo	418	0 (0%)	
Vogelmeier 2008	Tiotropiumbromid 18µg	221	0 (0%)	-
	Placebo	209	1 (0%)	

N=Anzahl der Patienten, SE= Standardfehler ITT= Intention-to-treat-Population KI=Konfidenzintervall

Die Tabelle 4.3-205 und Tabelle 4.3-206 stellen die Ergebnisse der Aclidiniumbromid- und Tiotropiumbromid-Studien dar, die über das Outcome "Letalität und Gesamt-Mortalität nach einer Behandlung von 12 bzw. 24 Wochen berichteten. Die Anzahl von Ereignissen war sehr gering oder Null, was die statistische Aussagekraft einschränkte. Eine NMA ist auf der Grundlage dieser geringen Anzahl von Ereignissen nicht informativ, da das Konfidenzintervall größer ist als das Ergebnis der geringen Zahl von Patienten, die ein Ereignis erlebt haben.

Wegen der sehr niedrigen Zahl der Ereignisse wurde kein indirekter Vergleich durchgeführt. *Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an die übliche Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:*

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Aufgrund der geringen Anzahl von Ereignissen wurde kein indirekter Vergleich durchgeführt.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.17 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.4.

Für die Darstellung der Subgruppenergebnisse für die Studien LAS-MD-33, 34 und 38A musste aufgrund zu kleiner Patientenzahlen in den vorab definierten Subgruppen von dem im Abschnitt 4.2.5.5 beschriebenen Vorgehen abgewichen werden.

Tabelle 4.3-207: Darstellung der Patientenpopulation der Studien ACCORD I, ATTAIN und ACCORD II nach definierten Subgruppen

Subgruppen		Placebo	Acridinium- bromid 400	Gesamt
ACCORD I				
1	Stufe II	n=111	n=118	n=229
2	Stufe III mit ≤ 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=70	n=66	n=136
3	Stufe III mit > 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=2	n=2	n=4
4	Stufe IV mit ≤ 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=1	n=1	n=2
5	Stufe IV mit > 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=0	n=0	n=0
Gesamt		n=184	n=187	n=371
ATTAIN				
1	Stufe II	n=178	n=184	n=362
2	Stufe III mit ≤ 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=90	n=81	n=171
3	Stufe III mit > 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=1	n=2	n=3
4+5	Stufe IV	n=0	n=0	n=0
Gesamt		n=269	n=267	n=536
ACCORD II				
1	Stufe II	n=113	n=79	n=192
2	Stufe III mit ≤ 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=67	n=90	n=157
3	Stufe III mit > 2 Exazerbationen in den letzten 12 Monaten vor Studienbeginn	n=0	n=6	n=6
4+5	Stufe IV	n=0	n=0	n=0

Gesamt	n=180	n=175	n=355
--------	-------	-------	-------

Wie in Tabelle 4.3-207 dargestellt, befinden sich lediglich in der LAS-MD-33 Patienten der COPD-Stufe IV (Subgruppe 4 und 5). Patienten der COPD-Stufe III mit mehr als zwei Exazerbationen in den letzten 12 Monaten wurden in allen Studien nur in kleinen Fallzahlen (n=3, 4, 6) beobachtet.

Aufgrund dessen werden im Folgenden die Subgruppenergebnisse der Studien ATTAIN, ACCORD I und II wie in Abschnitt 4.2.5.5 beschrieben, nach den Schweregraden II und III dargestellt.

4.3.2.1.3.17.1 CFB Trough-FEV₁Tabelle 4.3-208: Ergebnisse der Subgruppenanalyse für CFB Trough-FEV₁ aus RCT für indirekte Vergleiche

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	Gruppendifferenz ¹ (SE) KI [] p-Wert
ACCORD I					
12 Wochen					
Stufe II	AB (n=118)	1536 (480)	1643 (485)	107 (217,3)	134 (28,2)
	Pbo (n=111)	1630 (557)	1603 (551)	-27 (210,7)	KI [78,60; 189,4] p=<0,001
Stufe III	AB (n=68)	976 (271)	1076 (321)	100 (164,9)	129 (28,30)
	Pbo (n=72)	983 (314)	955 (365)	-29 (169,7)	KI [73,6; 184,4] p=<0,001
<i>Interaktion p=0,97</i>					
ACCORD II					
12 Wochen					
Stufe II	AB (n=79)	1599 (500)	1668 (460)	69 (240)	97 (38,2)
	Pbo (n=113)	1691 (464)	1664 (480)	-28 (287)	KI [22,2; 171,8] p= 0,01
Stufe III	AB (n=96)	961 (312)	1039 (351)	78 (215,6)	65 (31,1)
	Pbo (n=67)	1050 (302)	1063 (333)	13 (180,1)	KI [4; 126] p= 0,01
<i>Interaktion p=0,45</i>					
ATTAIN					
12 Wochen					
Stufe II	AB (n=184)	1690 (490)	1742 (500)	53 (257,7)	96 (26,9)
	Pbo (n=178)	1680 (462)	1637 (462)	-43 (253,5)	KI [43,3; 148,7] p=< 0,001
Stufe III	AB (n=84)	1119 (357)	1193 (386)	74 (210,8)	102 (32,5)
	Pbo (n=92)	1138 (291)	1110 (311)	-28 (220,6)	KI [38,3; 165,8] p= 0,001
<i>Interaktion p=0,91</i>					
24 Wochen					
Stufe II	AB (n=184)	1690 (490)	1737 (510)	47 (267)	119 (28,3)
	Pbo (n=178)	1680 (462)	1608 (477)	-72 (264)	KI [63,6; 174,4] p=<0,001
Stufe III	AB (n=84)	1119 (357)	1204 (391)	85 (220)	121 (33,9)
	Pbo (n=92)	1138 (291)	1103 (302)	-36 (187)	KI [54,5; 187,5] p=<0,001
<i>Interaktion p=0,97</i>					

¹ Analyse basiert auf einem adjustierten ANCOVA-Modell

AB: Acclidiniumbromid, Pbo: Placebo, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall

Tabelle 4.3-208 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für den 12 und 24 Wochen nach Baseline (Messung des Ausgangswertes) ermittelten CFB Trough-FEV₁. Sowohl Patienten mit mittelschwerer als auch schwerer COPD, die mit Acclidiniumbromid behandelt wurden, zeigten eine Verbesserung des Trough-FEV₁. Es liegt kein Hinweis auf Interaktion bezüglich des Schweregrades vor ($p > 0,2$).

Im Folgenden werden Meta-Analysen für den nach 12 Wochen ermittelten CFB Trough-FEV₁ nach Schweregraden durchgeführt. Für den Zeitpunkt 24 Wochen war dies nicht möglich, da ATTAIN die einzige Studie ist, die zu diesem Zeitpunkt Daten erhob.

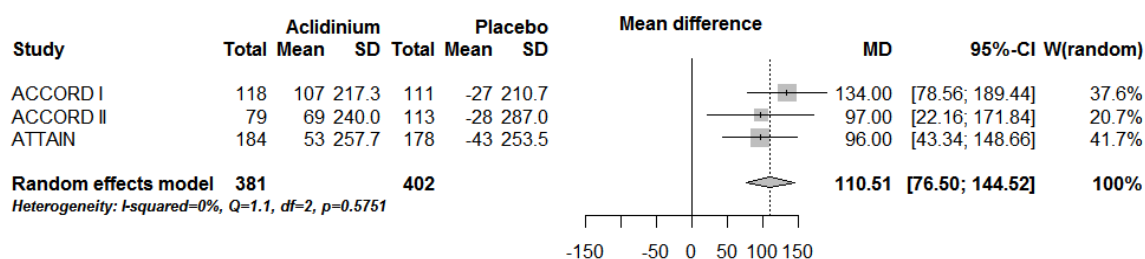


Abbildung 97: Meta-Analyse für mittelschwere COPD für CFB Trough-FEV₁ (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

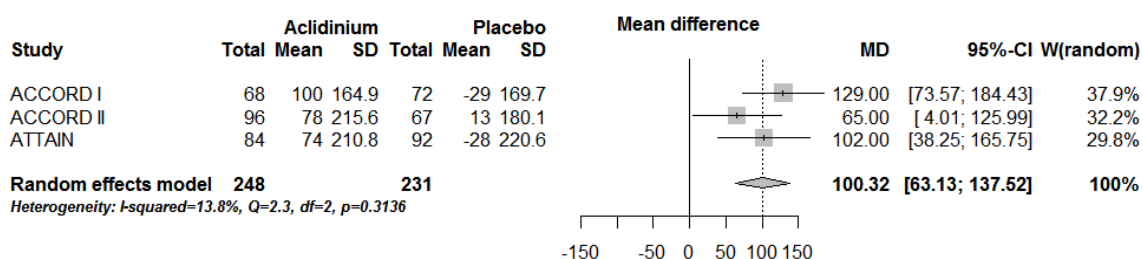


Abbildung 98: Meta-Analyse für schwere COPD für CFB Trough-FEV₁ (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Abbildung 97 zeigt, dass eine Behandlung mit Acclidiniumbromid im Vergleich mit Placebo und in Bezug auf die Erhöhung des Trough-FEV₁ nach 12 Wochen Behandlung für Patienten mit mittelschwerer COPD wirksamer ist (Differenz 110,5 [KI 95% 76,5; 144,5]). Ähnliche Ergebnisse wurden für Patienten mit schwerer COPD ermittelt (Abbildung 98). Eine Behandlung mit Acclidiniumbromid bewirkte bei Patienten mit schwerer COPD nach 12

Wochen einen höheren Trough-FEV₁ (Differenz 100,3 [KI 95% 63,1; 137,5]). Aus beiden Analysen ergab sich keine signifikante Studien-Heterogenität (mittelschwere COPD: Q=1,1; p=0,58; I²=0%; schwere COPD: Q=2,3; p=0,31; I²=13,8%).

4.3.2.1.3.17.2 CFB Peak-FEV₁Tabelle 4.3-209: Ergebnisse der Subgruppenanalyse für CFB Peak-FEV₁ aus RCT für indirekte Vergleiche

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
ACCORD I					
12 Wochen					
Stufe II	AB (n=118)	1536 (480)	1822 (514)	286 (260,7)	214 (33,9) KI [147,5; 280,5] p< 0,001
	Pbo (n=111)	1630 (557)	1702 (575)	72 (252,9)	
Stufe III	AB (n=68)	976 (271)	1212 (353)	236 (189,7)	178 (32,5) KI [114,2; 241,8] p< 0,001
	Pbo (n=72)	983 (314)	1042 (412)	58 (195,2)	
<i>Interaktion p=0,36</i>					
ACCORD II					
12 Wochen					
Stufe II	AB (n=79)	1599 (500)	1849 (490)	250 (257,8)	180 (41) KI [99,6; 260,4] p< 0,001
	Pbo (n=113)	1691 (464)	1761 (504)	70 (208,3)	
Stufe III	AB (n=96)	961 (312)	1155 (387)	194 (235,2)	72 (33,9) KI [5,5; 138,5] p= 0,03
	Pbo (n=67)	1050 (302)	1172 (422)	122 (196,4)	
<i>Interaktion p=0,04</i>					
ATTAIN					
12 Wochen					
Stufe II	AB (n=184)	1690 (490)	1911 (571)	222 (271,3)	178 (28,3) KI [122,6; 233,4] p< 0,001
	Pbo (n=178)	1680 (462)	1724 (490)	44 (266,8)	
Stufe III	AB (n=84)	1119 (357)	1373 (416)	253 (247,5)	194 (38,2) [119,2; 268,9] p< 0,001
	Pbo (n=92)	1050 (302)	1197 (311)	59 (259)	
<i>Interaktion p=0,69</i>					
24 Wochen					
Stufe II	AB (n=184)	1690 (490)	1926 (541)	237 (275)	196 (28,3) KI [140,6; 251,4] p=<0,001
	Pbo (n=178)	1680 (462)	1721 (497)	41 (272)	
Stufe III	AB (n=84)	1119 (357)	1367 (429)	248 (266)	204 (41,0)

Pbo (n=92)	1138 (291)	1183 (304)	44 (199)	KI [123,6; 284,4] p<0,001 <i>Interaktion p=0,80</i>
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell AB: Acclidiniumbromid, Pbo: Placebo, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall				

Tabelle 4.3-209 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für den 12 und 24 Wochen nach Baseline ermittelten CFB Peak-FEV₁. In allen Studien zeigten sowohl Patienten mit mittelschwerer als auch schwerer COPD nach der Behandlung mit Acclidiniumbromid im Vergleich mit der Placebogruppe eine verbesserte Spirometrie. Der Hinweis auf Interaktion in der ACCORD II-Studie (p=0,04) wird aufgrund der gleichgerichteten Effekte verworfen.

Dies bestätigt auch die beiden folgenden Meta-Analysen auch Schweregrad für den CFB Peak-FEV₁ nach 12 Wochen. Da nur die ATTAIN-Studie Daten nach 24 Wochen lieferte, war für diesen Zeitpunkt keine Meta-Analyse möglich.

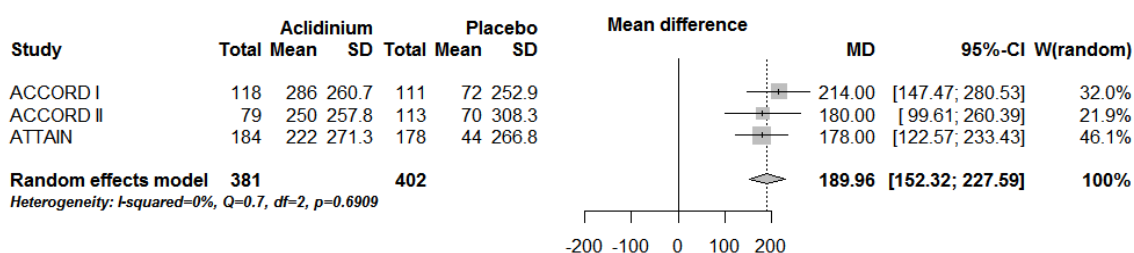


Abbildung 99: Meta-Analyse für mittelschwere COPD für FEV₁-Peak (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

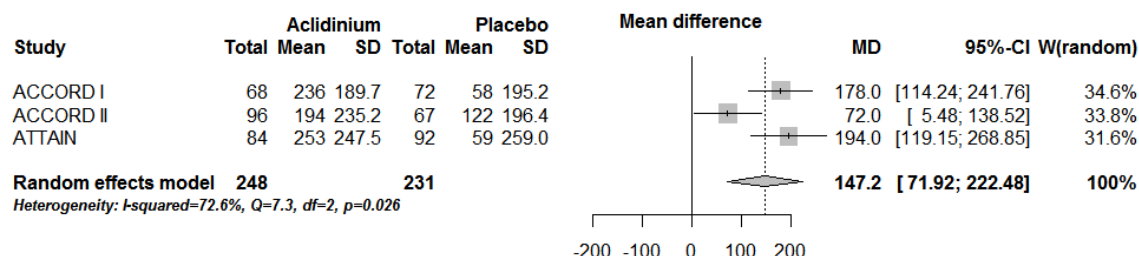


Abbildung 100: Meta-Analyse für schwere COPD für FEV₁-Peak (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Abbildung 1 zeigt, dass eine Behandlung mit Acclidiniumbromid von Patienten mit mittelschwerer COPD im Vergleich mit einem Placebo im Bezug auf die Erhöhung des FEV1-Peak nach 12 Wochen wirksamer ist (Differenz 189,9 [KI 95% 152,3; 227,6]). Die Ergebnisse für Patienten mit schwerer COPD waren ähnlich (Differenz 147,2 [KI 95% 71,9; 222,5]) (Abbildung 110). Die Analyse von Patienten mit mittelschwerer COPD zeigte keine signifikante Studien-Heterogenität ($Q=0,7$; $p=0,69$; $I^2=0\%$). Die Analyse von Patienten mit schwerer COPD zeigte jedoch eine signifikante Studien-Heterogenität ($Q=7,3$; $p=0,03$; $I^2=72,6\%$). Diese Heterogenität ist auf den Ausreißer in der ACCORD II-Studie zurückzuführen. Da es sich um einen signifikanten und den mittelschweren Patienten ähnlichen Gesamteffektschätzer handelt, entstehen keine Konsequenzen für die Nutzenbewertung.

4.3.2.1.3.17.3 TDI Focal Score

Tabelle 4.3-210: Ergebnisse der Subgruppenanalyse für den TDI-Score aus RCT für indirekte Vergleiche

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung
		BDI, Mean (SD)	TDI, Mean (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
ACCORD I				
12 Wochen				
Stufe II	AB (n=108)	6,60 (2,07)	1,29 (2,7)	0,55 (0,37) KI [-0,17; 1,27] p= 0,14
	Pbo (n=102)	6,99 (2,11)	0,74 (2,6)	
Stufe III	AB (n=60)	5,60 (2,04)	1,73 (3,2)	1,59 (0,57) KI [0,47; 2,72] p= 0,01 <i>Interaktion p=0,21</i>
	Pbo (n=57)	5,87 (2,10)	0,14 (3,1)	
ACCORD II				
12 Wochen				
Stufe II	AB (n=64)	6,52 (2,02)	1,77 (3,4)	1,51 (0,60) KI [0,34; 2,69] p= 0,01
	Pbo (n=91)	6,49 (2,14)	0,25 (4,0)	
Stufe III	AB (n=76)	5,65 (1,79)	0,68 (2,6)	0,58 (0,43) KI [-0,26; 1,41] p= 0,18 <i>Interaktion p=0,12*</i>
	Pbo (n=55)	5,60 (2,12)	0,11 (2,2)	
ATTAIN				
12 Wochen				
Stufe II	AB (n=176)	6,98 (2,17)	1,55 (2,8)	0,86 (0,29) KI [0,28; 1,43] p=0,003
	Pbo (n=167)	7,00 (1,98)	0,70 (2,7)	
Stufe III	AB (n=82)	6,22 (1,91)	1,56 (2,9)	0,75 (0,46) KI [-0,15; 1,64] p= 0,10 <i>Interaktion p=0,81</i>
	Pbo (n=86)	6,18 (1,96)	0,81 (3,0)	
24 Wochen				
Stufe II	AB (n=179)	6,98 (2,17)	1,83 (3,05)	0,86 (0,33) KI [0,23; 1,50] p=0,009
	Pbo (n=168)	7,00 (1,98)	0,96 (2,67)	
Stufe III	AB (n=82)	6,22 (1,91)	1,45 (3,51)	0,67 (0,55) KI [-0,40; 1,75] p=0,223
	Pbo (n=86)	6,18 (1,96)	0,78 (3,15)	

<i>Interaktion p=0,86</i>
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion AB: Acclidiniumbromid, Pbo: Placebo, SD: Standardabweichung, SE: Standardfehler, TDI: Transition Dyspnea Index, BDI: Baseline Dyspnea Index, KI: Konfidenzintervall

Tabelle 4.3-210 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für den 12 und 24 Wochen nach Baseline gemessenen TDI-Score. In allen Studien zeigten mit Acclidiniumbromid behandelte Patienten (unabhängig vom Schweregrad) verbesserte TDI-Scores. In der ACCORD II-Studie wurde ein Hinweis auf Interaktion bzgl. des Schweregrades festgestellt (p=0,12). Da in der Meta-Analyse (Abbildung 101, Abbildung 102) aller drei Studien ein signifikanter Gesamteffektschätzer für beide Schweregrade ermittelt wurde, folgen keine Auswirkungen auf die Nutzenbewertung.

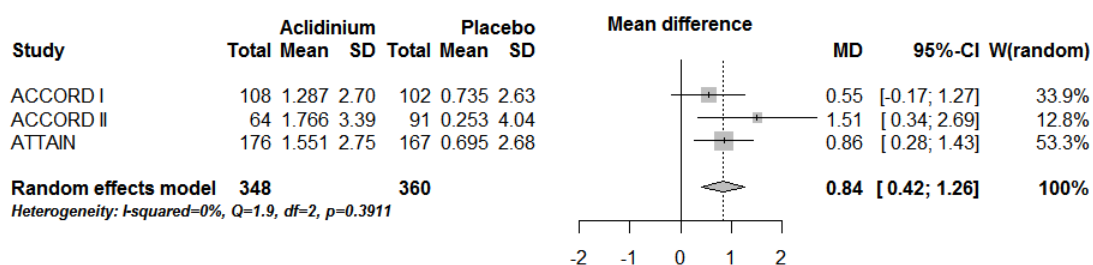


Abbildung 101: Meta-Analyse für mittelschwere COPD für den TDI Score (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

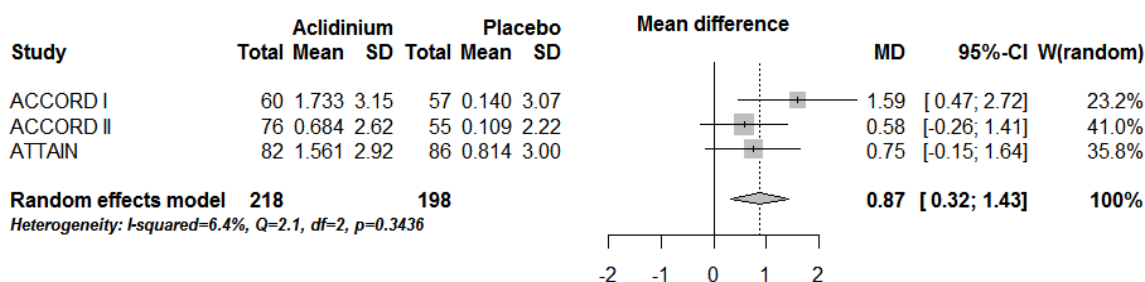


Abbildung 102: Meta-Analyse für schwere COPD für den TDI Score (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Abbildung 101 zeigt, dass eine Behandlung mit Acclidiniumbromid im Vergleich mit einem Placebo und in Bezug auf die Verbesserung des TDI Focal Score nach 12 Wochen Behandlung für Patienten mit mittelschwerer COPD wirksamer ist (Differenz 0,84 [KI 95% 0,42; 1,26]). Es existiert keine signifikante Heterogenität der Studien (Q=1,9; p=0,39; I²=0%).

Die Ergebnisse waren denen von Patienten mit schweren COPD-Symptomen ähnlich. Für Patienten mit schwerer COPD ist die Behandlung mit Acridiniumbromid im Vergleich mit einem Placebo ebenfalls wirksamer (Differenz 0,87 [KI 95% 0,32; 1,43]) (Abbildung 102). Es wurde wiederum keine Studien-Heterogenität erkannt ($Q=2,1$; $p=0,34$; $I^2=6,4\%$).

4.3.2.1.3.17.4 TDI-Responder

Tabelle 4.3-211: Ergebnisse der Subgruppenanalyse für TDI-Responder aus RCT für indirekte Vergleiche

Studie	Studienarm (Patienten)	Follow-Up	Veränderung
		N responder (%)	Gruppendifferenz ¹ (SE) KI [] p-Wert
ACCORD I			
12 Wochen			
Stufe II	AB (n=108)	45 (41,7%)	1,25 (0,28) KI [0,72; 2,19] p= 0,42
	Pbo (n=102)	37 (36,3%)	
Stufe III	AB (n=60)	35 (58,3%)	3,92 (0,40) KI [1,79; 8,57] p< 0,001
	Pbo (n=57)	15 (26,3%)	
<i>Interaktion p=0,08*</i>			
ACCORD II			
12 Wochen			
Stufe II	AB (n=64)	37 (57,8%)	2,30 (0,33) KI [1,20; 4,41] p= 0,01
	Pbo (n=91)	34 (37,4%)	
Stufe III	AB (n=76)	33 (43,4%)	1,87 (0,38) KI [0,89; 3,91] p= 0,10
	Pbo (n=55)	16 (29,1%)	
<i>Interaktion p=0,63</i>			
ATTAIN			
12 Wochen			
Stufe II	AB (n=176)	98 (55,7%)	1,92 (0,22) KI [1,25; 2,95] p= 0,003
	Pbo (n=167)	66 (39,5%)	
Stufe III	AB (n=82)	48 (58,5%)	1,52 (0,31) KI [0,82; 2,79] p= 0,18
	Pbo (n=85)	41 (58,2%)	
<i>Interaktion p=0,77</i>			
24 Wochen			
Stufe II	AB (n=176)	109 (60,9%)	2,37 (0,22) KI [1,54; 3,65] p=<0,001
	Pbo (n=167)	68 (40,5%)	
Stufe III	AB (n=82)	46 (56,1%)	1,37 (0,31)

Pbo (n=85)	41 (47,7%)	KI [0,75; 2,52] p=0,310 <i>Interaktion p=0,23</i>
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion Responder = Anzahl der Patienten mit einer Reduzierung des TDI Focal Scores von mindestens einer Einheiten im Vergleich zum Ausgangswert, SE=Standardfehler, TDI: Transition Dyspnea Index, KI=Konfidenzintervall		

Tabelle 4.3-211 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für die Anzahl der TDI-Responder, 12 und 24 Wochen nach Baseline. In allen Studien zeigten mit Acridiniumbromid behandelte Patienten (unabhängig vom Schweregrad) verbesserte TDI-Scores. Der einzige Hinweis auf Interaktion in der ACCORD I-Studie (0,08) hat aufgrund der signifikanten Ergebnisse der Meta-Analyse nach Schweregraden (Abbildung 1 und 2) keine Auswirkungen auf die Nutzenbewertung.

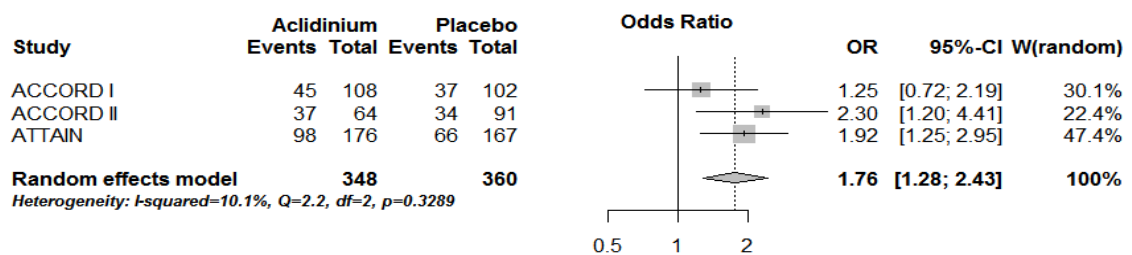


Abbildung 103: Meta-Analyse für mittelschwere COPD für TDI-Responder (12 Wochen) aus RCT; Acridiniumbromid im Vergleich zu Placebo

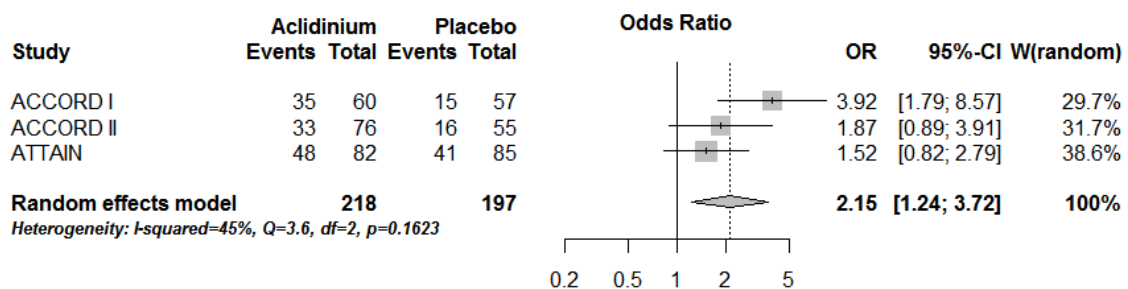


Abbildung 104: Meta-Analyse für schwere COPD für TDI-Responder (12 Wochen) aus RCT; Acridiniumbromid im Vergleich zu Placebo

Für Patienten mit mittelschwerer COPD erhöht Acridiniumbromid im Vergleich mit einem Placebo die Wahrscheinlichkeit der Response auf den TDI Focal Score (Differenz 1,76 [KI 95% 1,28; 2,43]) (Abbildung 103). Die Ergebnisse für Patienten mit schwerer COPD waren ähnlich (Differenz 2,15 [KI 95% 1,24; 3,72]) (Abbildung 104). In der letzteren Analyse ergaben sich jedoch extrem große Konfidenzintervalle. In keiner der beiden Analysen wurde eine Studien-Heterogenität erkannt (mittelschwere COPD: $Q=2,2$; $p=0,33$; $I^2=10,1\%$; schwere COPD: $Q=3,6$; $p=0,16$; $I^2=45\%$). Da nur drei Studien zur Berücksichtigung zur Verfügung standen, sollten diese Testergebnisse mit Vorsicht interpretiert werden.

4.3.2.1.3.17.5 CFB EXACT-Respirationssymptome

Tabelle 4.3-212: Ergebnisse der Subgruppenanalyse für EXACT-Respirationssymptome (12 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten	Baseline	Follow-up	Veränderung	Gruppendifferenz ¹ (SE) KI [] p-wert
		(ITT)	Mean (SD)	Mean (SD)	CFB (SD)	
ATTAIN						
Atemnot						
Stufe II	AB	183	6,3 (3,5)	5,3 (3,7)	-1,0 (2,6)	-1,26 (0,27)
	Placebo	176	6,0 (3,5)	6,2 (3,8)	0,3 (2, 4)	KI [-1,78; -0,73] p=<0,001
Stufe III	AB	84	7,8 (3,1)	6,8 (3,7)	-1,1 (2,6)	-0,74 (0,40)
	Placebo	92	7,7 (3,4)	7,3 (3,5)	-0,3 (2,3)	KI [-1,52; 0,05] p=0,064
<i>Interaktion p=0,27</i>						
Engegefühl im Brustbereich						
Stufe II	AB	183	3,3 (1,9)	2,6 (1,9)	-0,7 (1,4)	-0,73 (0,15)
	Placebo	176	3,0 (1,9)	3,0 (2,2)	-0,0 (1,5)	KI [-1,02; -0,44] p=<0,001
Stufe III	Acidinium	84	3,5 (1,9)	3,1 (2,1)	-0,4 (1,7)	-0,09 (0,27)
	Placebo	92	3,7 (1,9)	3,3 (2, 1)	-0,5 (1,6)	KI [-0,60; 0,43] p=0,739
<i>Interaktion p=0,04**</i>						
Husten und Sputum						
Stufe II	AB	183	3,9 (1,8)	3,1 (1,8)	-0,8 (1,4)	-0,68 (0,15)
	Placebo	176	3,7 (2,0)	3,5 (1,9)	-0,2 (1,4)	KI [-0,96; -0,39] p=<0,001
Stufe III	Acidinium	84	3,9 (1,7)	3,4 (1,8)	-0,5 (1,2)	-0,02 (0,19)
	Placebo	92	4,2 (1,7)	3,7 (1,7)	-0,5 (1,4)	KI [-0,39; 0,35] p=0,916
<i>Interaktion p=0,04**</i>						
Gesamtwert						
Stufe II	AB	183	13,5 (6,6)	10,9 (6,7)	-2,6 (4,8)	-2,66 (0,50)
	Placebo	176	12,6 (6,5)	12,7 (7,0)	0,1 (4,5)	KI [-3,64; -1,68] p=<0,001
Stufe III	AB	84	15,3 (5,8)	13,3 (6,6)	-2,0 (4,7)	-0,85 (0,72)
	Placebo	92	15,6 (6,4)	14,4 (6,4)	-1,2 (4,8)	KI [-2,27; 0,57] p=0,238
<i>Interaktion p=0,06*</i>						
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion, **Beleg für Interaktion N = Anzahl, AB = Acidiniumbromid, CFB= Veränderung gegenüber dem Ausgangswert, ITT = Intention to						

Treat, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall
--

Tabelle 4.3-212 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für die 12 Wochen nach Baseline registrierten EXACT-Respirationssymptome. Werte für die verschiedenen Kategorien und den Gesamtwert werden separat dargestellt. Nach 12 Wochen resultierte die Behandlung mit Aclidiniumbromid für alle Kategorien und COPD-Schweregrade in weniger Symptomen – diese Beobachtung ist jedoch nicht in allen Gruppen gleich stark ausgeprägt. Die Ergebnisse der Interaktionstests belegen eine deutlichere Verbesserung bei den mittelschweren gegenüber den schweren Patienten. Da nur eine Studie über dieses Outcome berichtete, wurde keine Meta-Analyse durchgeführt.

Tabelle 4.3-213: Ergebnisse der Subgruppenanalyse für EXACT-Respirationssymptome (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-wert
ATTAIN						
Atemnot						
Stufe II	AB	183	6,3 (3,5)	5,0 (3,8)	-1,3 (2,9)	-1,45 (0,30)
	Placebo	176	6,0 (3,5)	6,0 (4,0)	0,2 (2,6)	KI [-2,03; -0,86] p=<0,001
Stufe III	AB	84	7,8 (3,1)	6,9 (3,7)	-0,9 (2,9)	-0,58 (0,44)
	Placebo	92	7,7 (3,4)	7,4 (3,4)	-0,3 (2,4)	KI [-1,45; 0,29] p=0,187
<i>Interaktion p=0,09*</i>						
Engegefühl im Brustbereich						
Stufe II	AB	183	3,3 (1,9)	2,4 (1,9)	-1,0 (1,5)	-0,96 (0,16)
	Placebo	176	3,0 (1,9)	3,0 (2,3)	0,0 (1,7)	KI [-1,27; -0,66] p=<0,001
Stufe III	AB	84	3,5 (1,9)	3,2 (2,1)	-0,3 (1,8)	0,10 (0,28)
	Placebo	92	3,7 (1,9)	3,2 (2,2)	-0,4 (1,6)	KI [-0,44; 0,65] p=0,721
<i>Interaktion p=0,0008**</i>						
Husten und Sputum						
Stufe II	AB	183	3,9 (1,8)	2,9 (1,8)	-1,0 (1,4)	-0,83 (0,15)
	Placebo	176	3,7 (2,0)	3,4 (1,9)	-0,2 (1,4)	KI [-1,12; -0,54] p=<0,001
Stufe III	AB	84	3,9 (1,7)	3,4 (1,8)	-0,6 (1,3)	-0,06 (0,20)
	Placebo	92	4,2 (1,7)	3,7 (1,7)	-0,5 (1,5)	KI [-0,46; 0,34] p=0,764
<i>Interaktion p=0,01**</i>						
Gesamtwert						
Stufe II	AB	183	13,5 (6,6)	10,2 (6,8)	-3,3 (5,1)	-3,24 (0,53)
	Placebo	176	12,6 (6,5)	12,5 (7,3)	-0,0 (4,9)	KI [-4,28; -2,19] p=<0,001
Stufe III	AB	84	15,3 (5,8)	13,5 (6,6)	-1,8 (5,2)	-0,54 (0,81)
	Placebo	92	15,6 (6,4)	14,3 (6,5)	-1,2 (5,1)	KI [-2,12; 1,05] p=0,505
<i>Interaktion p=0,01**</i>						
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion, ** Beleg für Interaktion N = Anzahl, AB = Acridiniumbromid, CFB= Veränderung gegenüber dem Ausgangswert, ITT = Intention to Treat, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall						

Tabelle 4.3-213 zeigt die Studienergebnisse pro COPD-Schweregrad für die 24 Wochen nach Baseline registrierten EXACT-Respirationssymptome. Daten zu den verschiedenen Kategorien und dem Gesamtscore sind separat aufgeführt. Die Behandlung mit Acclidiniumbromid führte bei Patienten mit mittelschwerer COPD zu einer signifikanten Reduzierung der EXACT-Respirationssymptome. Bei Patienten mit schwerer COPD zeigt die Behandlung mit Acclidiniumbromid im Vergleich zu Placebo weniger deutliche Ergebnisse für die Kategorien des EXACT-Fragebogens. Dies wird durch die Ergebnisse der Interaktionstests bestätigt.

4.3.2.1.3.17.6 CFB Nächtliche COPD-Symptome

Tabelle 4.3-214 Ergebnisse der Subgruppenanalyse für nächtliche COPD-Symptome (12 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
			Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-wert
ACCORD I						
Häufigkeit von nächtlichem Husten (pro Woche)						
Stufe II	AcB	114	1,96 (1,56)	1,65 (1,55)	-0,29 (1,31)	-0,30 (0,17)
	Placebo	107	2,03 (1,40)	2,08 (1,85)	0,01 (1,33)	KI [-0,63; 0,03] p=0,078
Stufe III	AB	63	1,82 (1,69)	1,33 (1,50)	-0,52 (1,25)	-0,73 (0,23)
	Placebo	67	2,16 (1,66)	2,34 (2,07)	0,21 (1,35)	KI [-1,18; -0,28] p=0,002
<i>Interaktion p=0,06*</i>						
Häufigkeit von nächtlichem Auswurf (pro Woche)						
Stufe II	AB	114	1,36 (1,47)	0,94 (1,27)	-0,38 (0,93)	-0,39 (0,13)
	Placebo	107	1,29 (1,33)	1,29 (1,51)	-0,00 (1,02)	KI [-0,64; -0,14] p=0,003
Stufe III	AB	63	1,29 (1,67)	0,97 (1,41)	-0,35 (0,89)	-0,44 (0,16)
	Placebo	67	1,47 (1,62)	1,57 (1,79)	0,09 (0,91)	KI [-0,75; -0,13] p=0,006
<i>Interaktion p=0,57</i>						
Häufigkeit von nächtlicher Atemnot (pro Woche)						
Stufe II	AB	114	1,32 (1,18)	0,97 (1,07)	-0,35 (0,99)	-0,15 (0,13)
	Placebo	107	1,30 (1,10)	1,11 (1,23)	-0,20 (0,85)	KI [-0,40; 0,10] p=0,249
Stufe III	AB	63	1,53 (1,40)	0,97 (1,18)	-0,58 (1,32)	-0,53 (0,23)
	Placebo	67	1,61 (1,33)	1,53 (1,38)	-0,05 (1,01)	KI [-0,99; -0,07] p=0,021
<i>Interaktion p=0,05*</i>						

Auftreten von nächtlichem Keuchen (pro Woche)						
Stufe II	AB	114	1,24 (1,40)	0,71 (0,93)	-0,50 (1,10)	-0,49 (0,14)
	Placebo	107	1,20 (1,40)	1,21 (1,61)	-0,01 (1,15)	KI [-0,77; -0,21] p=<0,001
Stufe III	AB	63	1,26 (1,55)	0,70 (0,96)	-0,57 (1,55)	-0,56 (0,28)
	Placebo	67	1,54 (1,56)	1,58 (1,82)	-0,01 (1,15)	KI [-1,11; -0,01] p=0,046
<i>Interaktion p=0,38</i>						
Schwere und Grad der Einschränkung durch Atemprobleme im Schlaf (pro Woche)						
Stufe II	AB	114	0,86 (0,76)	0,66 (0,75)	-0,19 (0,54)	-0,11 (0,07)
	Placebo	107	0,79 (0,72)	0,69 (0,80)	-0,08 (0,58)	KI [-0,24; 0,02] p=0,116
Stufe III	AB	63	0,84 (0,80)	0,55 (0,73)	-0,33 (0,63)	-0,30 (0,11)
	Placebo	67	0,88 (0,71)	0,83 (0,82)	-0,03 (0,60)	KI [-0,52; -0,08] p=0,006
<i>Interaktion p=0,08*</i>						
Schwere und Grad der Einschränkung durch Atemnot (pro Woche)						
Stufe II	AB	114	1,65 (0,90)	1,30 (0,93)	-0,32 (0,80)	-0,10 (0,11)
	Placebo	107	1,67 (0,90)	1,48 (1,01)	-0,22 (0,68)	KI [-0,32; 0,12] p=0,363
Stufe III	AB	63	1,84 (0,95)	1,24 (0,90)	-0,63 (0,94)	-0,47 (0,17)
	Placebo	67	2,03 (0,93)	1,81 (1,04)	-0,16 (0,71)	KI [-0,81; -0,13] p=0,006
<i>Interaktion p=0,01**</i>						
Schwere und Grad der Einschränkung durch Husten (pro Woche)						
Stufe II	AB	114	1,42 (0,97)	1,20 (0,93)	-0,19 (0,74)	-0,01 (0,09)
	Placebo	107	1,43 (0,84)	1,30 (1,01)	-0,18 (0,81)	KI [-0,19; 0,17] p=0,912
Stufe III	AB	63	1,34 (0,99)	1,03 (0,84)	-0,33 (0,80)	-0,34 (0,14)
	Placebo	67	1,57 (0,94)	1,53 (0,97)	0,00 (0,72)	KI [-0,62; -0,06] p=0,015
<i>Interaktion p=0,01**</i>						
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion, ** Beleg für Interaktion N = Anzahl, AB = Acridiniumbromid, CFB= Veränderung gegenüber dem Ausgangswert, ITT = Intention to Treat, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall						

Tabelle 4.3-214 zeigt die Subgruppenergebnisse nach COPD-Schweregrad für die nach 12 Wochen registrierten nächtlichen Symptome in der ACCORD I Studie. In der Studie führte die Behandlung von Patienten mit mittelschwerer COPD mit Acridiniumbromid in vier der sieben Kategorien zu einer stärkeren Reduzierung der nächtlichen COPD-Symptome als bei Patienten mit schwerer COPD. Die entsprechenden Interaktionswerte sind in der Tabelle gekennzeichnet. In den übrigen Kategorien ist die Verbesserung in beiden Subgruppen gleichstark ausgeprägt.

Tabelle 4.3-215: Ergebnisse der Subgruppenanalyse für nächtliche COPD-Symptome (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-wert
ATTAIN						
Auswurf (letzte Nacht)						
Stufe II	AB	183	0,59 (0,38)	0,43 (0,43)	-0,16 (0,37)	-0,10 (0,04)
	Placebo	175	0,53 (0,38)	0,46 (0,44)	-0,06 (0,36)	KI [-0,18; -0,02] p=0,012
Stufe III	AB	84	0,56 (0,36)	0,43 (0,40)	-0,13 (0,38)	-0,08 (0,06)
	Placebo	92	0,63 (0,34)	0,59 (0,42)	-0,05 (0,42)	KI [-0,20; 0,04] p=0,182
<i>Interaktion p=0,60</i>						
Brustenge oder Kongestion (letzte Nacht)						
Stufe II	AB	182	0,28 (0,32)	0,20 (0,33)	-0,09 (0,30)	-0,09 (0,03)
	Placebo	175	0,25 (0,31)	0,26 (0,37)	0,01 (0,32)	KI [-0,16; -0,03] p=0,003
Stufe III	AB	84	0,27 (0,30)	0,27 (0,37)	0,00 (0,32)	0,0 (0,04)
	Placebo	92	0,30 (0,31)	0,31 (0,40)	0,00 (0,32)	KI [-0,09; 0,08] p=1,0
<i>Interaktion p=0,22</i>						
Husten (letzte Nacht)						
Stufe II	AB	183	0,65 (0,38)	0,47 (0,43)	-0,18 (0,36)	-0,13 (0,04)
	Placebo	176	0,64 (0,36)	0,58 (0,41)	-0,06 (0,30)	KI [-0,20; -0,05] p=0,001
Stufe III	AB	84	0,60 (0,37)	0,50 (0,43)	-0,10 (0,31)	-0,02 (0,04)
	Placebo	92	0,69 (0,34)	0,60 (0,42)	-0,09 (0,37)	KI [-0,10; 0,07] p=0,617
<i>Interaktion p=0,15*</i>						
Gefühl von Atemlosigkeit oder Atemnot (letzte Nacht)						
Stufe II	AB	183	0,34 (0,35)	0,26 (0,38)	-0,08 (0,33)	-0,12 (0,03)
	Placebo	175	0,28 (0,33)	0,33 (0,41)	0,04 (0,35)	KI [-0,18; -0,06] p<0,001
Stufe III	AB	84	0,41 (0,34)	0,38 (0,42)	-0,03 (0,38)	-0,03 (0,06)
	Placebo	92	0,46 (0,35)	0,46 (0,44)	0,00 (0,37)	KI [-0,15; 0,09] p=0,617
<i>Interaktion p=0,36</i>						
Keuchen (letzte Nacht)						
Stufe II	AB	183	0,35 (0,38)	0,24 (0,38)	-0,11 (0,36)	-0,07 (0,04)
	Placebo	175	0,29 (0,35)	0,25 (0,38)	-0,04 (0,32)	KI [-0,15; 0,0] p=0,080

Stufe III	AB	84	0,43 (0,39)	0,36 (0,43)	-0,07 (0,37)	0,01 (0,06)
	Placebo	92	0,36 (0,38)	0,29 (0,40)	-0,08 (0,35)	KI [-0,11; 0,13] p=0,868
<i>Interaktion p=0,14*</i>						
Jegliches Symptom (letzte Nacht)						
Stufe II	AB	183	0,83 (0,29)	0,67 (0,40)	-0,17 (0,36)	-0,11 (0,04)
	Placebo	175	0,83 (0,28)	0,77 (0,36)	-0,06 (0,28)	KI [-0,19; -0,03] p=0,006
Stufe III	AB	84	0,85 (0,24)	0,79 (0,34)	-0,06 (0,29)	-0,04 (0,05)
	Placebo	92	0,87 (0,24)	0,85 (0,31)	-0,02 (0,30)	KI [-0,13; 0,05] p=0,424
<i>Interaktion p=0,26</i>						
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion N = Anzahl, AB = Acridiniumbromid, CFB= Veränderung gegenüber dem Ausgangswert, ITT = Intention to Treat, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall						

Die Ergebnisse der nach 24 Wochen registrierten nächtlichen Symptome werden in Tabelle 4.3-215 dargestellt. Die Ergebnisse für die unterschiedlichen Kategorien sind nach COPD-Schweregrad eingeteilt. Nur die ATTAIN-Studie berichtete Daten für dieses Outcome nach 24 Wochen. In allen Symptom-Kategorien führte die Behandlung mit Acridiniumbromid in beiden Subgruppen zu einer mindestens numerischen Reduzierung der nächtlichen COPD-Symptome. Es existieren zwei Hinweise auf Interaktion, die die Gesamtbewertung des Scores jedoch nicht beeinflussen.

4.3.2.1.3.17.7 CFB Morgendliche COPD-Symptome

Tabelle 4.3-216: Ergebnisse der Subgruppenanalyse für morgendliche COPD-Symptome (12 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
			Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-wert
ACCORD I						
Grad der Einschränkung bei morgendlichen Aktivitäten						
Stufe II	AB	114	1,31 (0,87)	1,13 (0,80)	-0,15 (0,65)	-0,10 (0,08)
	Placebo	107	1,30 (0,84)	1,24 (0,88)	-0,05 (0,54)	KI [-0,26; 0,06] p=0,211
Stufe III	AB	63	1,53 (0,88)	1,08 (0,84)	-0,44 (0,86)	-0,43 (0,16)
	Placebo	67	1,63 (0,91)	1,54 (1,03)	-0,01 (0,59)	KI [-0,74; -0,12] p=0,007
<i>Interaktion p=0,01**</i>						

Schwere der Atemnot in der ersten Stunde nach dem Aufstehen am Morgen						
Stufe II	AB	114	1,42 (0,93)	1,20 (0,85)	-0,20 (0,70)	-0,07 (0,09)
	Placebo	107	1,43 (0,87)	1,31 (0,97)	-0,13 (0,56)	KI [-0,25; 0,11] p=0,437
Stufe III	AB	63	1,69 (0,89)	1,17 (0,87)	-0,50 (0,89)	-0,45 (0,16)
	Placebo	67	1,80 (0,98)	1,67 (1,09)	-0,05 (0,67)	KI [-0,76; -0,15] p=0,005
<i>Interaktion p=0,01**</i>						
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion, ** Beleg für Interaktion N = Anzahl, AB = Acridiniumbromid, CFB= Veränderung gegenüber dem Ausgangswert, ITT = Intention to Treat, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall						

4.3.2.1.3.17.8 CFB Morgendliche COPD-Symptome

Tabelle 4.3-216 zeigt die Subgruppenergebnisse nach COPD-Schweregrad für die nach 12 Wochen registrierten morgendlichen Symptome in der ACCORD I-Studie. In der Studie führte die Behandlung von Patienten mit schwerer COPD unter Acridiniumbromid in zwei der drei Kategorien zu einer stärkeren Reduzierung der morgendlichen COPD-Symptome als bei Patienten mit COPD-Stufe II.

Tabelle 4.3-217: Ergebnisse der Subgruppenanalyse für morgendliche COPD-Symptome (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	Gruppendifferenz ¹ (SE) KI [] p-wert
		N	Mean (SD)	Mean (SD)	CFB (SD)	
ATTAIN						
Auswurf (am Morgen)						
Stufe II	AB	182	0,64 (0,38)	0,56 (0,43)	-0,09 (0,29)	-0,12 (0,03) KI [-0,18; -0,06] p=<0,001
	Placebo	175	0,61 (0,38)	0,63 (0,41)	0,03 (0,35)	
Stufe III	AB	84	0,66 (0,38)	0,64 (0,44)	-0,02 (0,31)	-0,07 (0,04) KI [-0,16; 0,02] p=0,080
	Placebo	92	0,70 (0,34)	0,75 (0,36)	0,05 (0,35)	
<i>Interaktion p=0,65</i>						
Brustenge oder Kongestion (am Morgen)						
Stufe II	AB	182	0,28 (0,32)	0,23 (0,36)	-0,05 (0,30)	-0,11 (0,03) KI [-0,17; -0,04] p=<0,001
	Placebo	175	0,25 (0,32)	0,30 (0,40)	0,05 (0,32)	
Stufe III	AB	84	0,33 (0,35)	0,36 (0,40)	0,02 (0,35)	-0,01 (0,05) KI [-0,11; 0,09] p=0,841
	Placebo	92	0,30 (0,31)	0,33 (0,41)	0,03 (0,33)	
<i>Interaktion p=0,09*</i>						
Husten (am Morgen)						
Stufe II	AB	182	0,69 (0,38)	0,59 (0,42)	-0,10 (0,29)	-0,11 (0,03) KI [-0,17; -0,05] p=<0,001
	Placebo	175	0,70 (0,35)	0,71 (0,37)	0,01 (0,26)	
Stufe III	AB	84	0,69 (0,36)	0,65 (0,41)	-0,04 (0,26)	-0,01 (0,05) KI [-0,10; 0,08] p=0,841
	Placebo	92	0,72 (0,34)	0,69 (0,38)	-0,03 (0,34)	
<i>Interaktion p=0,05*</i>						
Gefühl von Atemlosigkeit oder Atemnot (am Morgen)						
Stufe II	AB	182	0,36 (0,37)	0,33 (0,42)	-0,03 (0,35)	-0,10 (0,04) KI [-0,17; -0,03] p=0,012
	Placebo	175	0,35 (0,36)	0,42 (0,43)	0,07 (0,35)	
Stufe III	AB	84	0,54 (0,35)	0,49 (0,44)	-0,05 (0,38)	-0,02 (0,06) KI [-0,14; 0,09] p=0,739
	Placebo	92	0,53 (0,36)	0,51 (0,45)	-0,02 (0,37)	
<i>Interaktion p=0,25</i>						
Keuchen (am Morgen)						
Stufe II	AB	182	0,34 (0,37)	0,26 (0,40)	-0,08 (0,35)	-0,10 (0,04) KI [-0,17; -0,02] p=0,012
	Placebo	175	0,28 (0,35)	0,30 (0,41)	0,01 (0,32)	

Stufe III	AB	84	0,46 (0,40)	0,39 (0,43)	-0,07 (0,38)	0,0 (0,06)
	Placebo	92	0,38 (0,39)	0,31 (0,42)	-0,07 (0,37)	KI [-0,12; 0,12] p=1,0
<i>Interaktion p=0,08*</i>						
Jegliches Symptom (am Morgen)						
Stufe II	AB	182	0,87 (0,25)	0,80 (0,32)	-0,08 (0,27)	-0,08 (0,03)
	Placebo	175	0,89 (0,23)	0,89 (0,27)	0,01 (0,22)	KI [-0,14; -0,02] p=0,008
Stufe III	AB	84	0,94 (0,18)	0,92 (0,22)	-0,03 (0,23)	-0,04 (0,03)
	Placebo	92	0,91 (0,23)	0,92 (0,22)	0,02 (0,22)	KI [-0,10; 0,02] p=0,182
<i>Interaktion p=0,18*</i>						
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion N = Anzahl, AB = Aclidiniumbromid, CFB= Veränderung gegenüber dem Ausgangswert, ITT = Intention to Treat, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall						

Tabelle 4.3-217 zeigt die Ergebnisse für die 24 Wochen nach Baseline registrierten morgendlichen COPD-Symptome pro COPD-Schweregrad. Die Ergebnisse werden für alle Kategorien und die beiden COPD-Schweregrade separat dargestellt. Es wird eine bedeutendere Symptomreduzierung bei den Stufe II-Patienten beobachtet. Dies spiegelt sich auch in den Interaktionswerten wieder.

4.3.2.1.3.17.9 Exazerbationsrate

Tabelle 4.3-218: Ergebnisse der Subgruppenanalyse für die Exazerbationsrate aus RCT für indirekte Vergleiche

Studie	Studienarm (Patienten)	Anzahl ² Personenjahre	Anzahl Exazerbationen	Rate	Rate ratio ¹ (SE) KI [] p-Wert
ACCORD I					
Stufe II	AB (n=118)	26	7	0,265	0,48 (1,60) KI [0,19; 1,19] p= 0,11
	Pbo (n=111)	23	13	0,509	
Stufe III	AB (n=68)	68	15	0,244	0,44 (1,82) KI [0,14; 1,44] p= 0,18
	Pbo (n=72)	72	15	0,533	
<i>Interaktion p=0,75</i>					
ACCORD II					
Stufe II	AB (n=79)	17	7	0,415	1,10 (1,66) KI [0,41; 2,95] p= 0,85
	Pbo (n=113)	24	9	0,366	
Stufe III	AB (n=96)	20	12	0,564	0,84 (1,53) [0,36; 1,94] p= 0,68
	Pbo (n=67)	14	10	0,711	
<i>Interaktion p=0,48</i>					
ATTAIN					
Stufe II	AB (n=184)	89	22	0,305	0,63 (1,32) KI [0,37; 1,09] p= 0,10
	Pbo (n=178)	82	32	0,481	
Stufe III	AB (n=84)	40	16	0,564	0,70 (1,38) KI [0,37; 1,32] p= 0,27
	Pbo (n=92)	42	24	0,744	
<i>Interaktion p=0,72</i>					
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell					
² Zur Berechnung Anzahl der Personenjahre wurde für Patienten mit mittelschwerer und schwerer COPD die Gesamt-Abbruchraten zugrunde gelegt (d. h. wenn die Gesamt-Abbruchrate 20% betrug, wurde davon ausgegangen, dass 20% der Patienten in beiden Gruppen diese Studie nicht abschließen würden)					
AB: Aclidiniumbromid, Pbo: Placebo, SE: Standardfehler, KI: Konfidenzintervall					

Tabelle 4.3-218 zeigt die Exazerbationsdaten zu Aclidiniumbromid pro COPD-Schweregrad. Die Exazerbationsrate war bei allen mit Aclidiniumbromid behandelten Patienten (unabhängig vom Schweregrad) im Vergleich zur der Placebogruppe generell niedriger (besser). Ein Hinweis auf Interaktion lag nicht vor ($p > 0,2$). Dies ist auch in den Meta-Analysen (Abbildung 105, Abbildung 106) deutlich erkennbar.

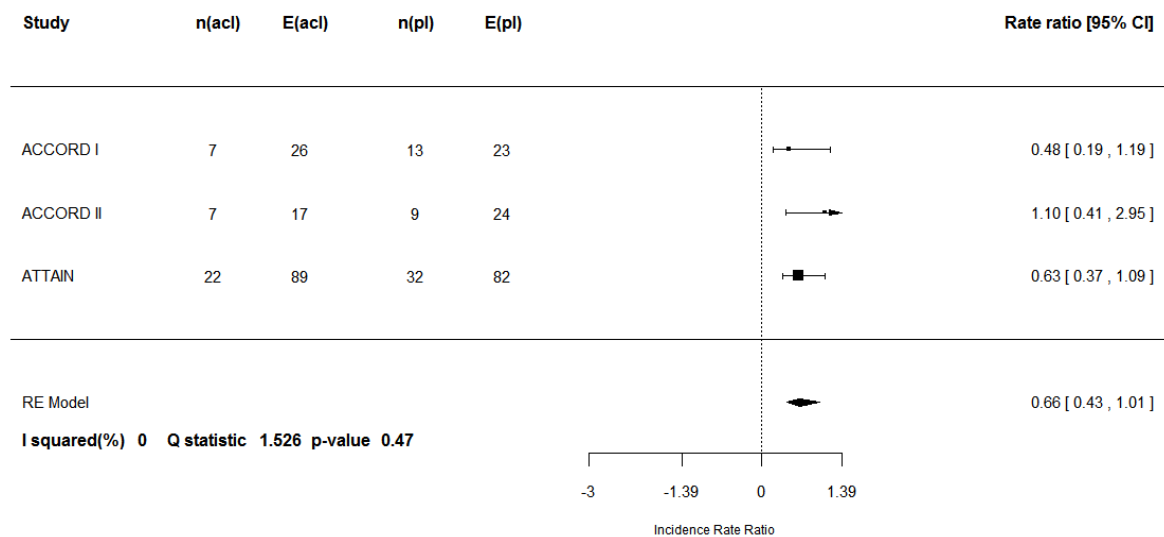


Abbildung 105: Meta-Analyse für mittelschwere COPD für Exazerbationsrate aus RCT; Aclidiniumbromid im Vergleich zu Placebo

Abbildung 105 zeigt, dass Aclidiniumbromid für Patienten mit mittelschwerer COPD bei der Reduzierung der Exazerbationsrate nicht signifikant wirksamer als ein Placebo ist (gepooltes Ratenverhältnis 0,66 [KI 95% 0,43; 1,01]). Der Heterogenitätstest ($Q=1,53$; $p=0,47$; $I^2=0$) zeigt, dass es keine Heterogenität unter den Studien gibt.

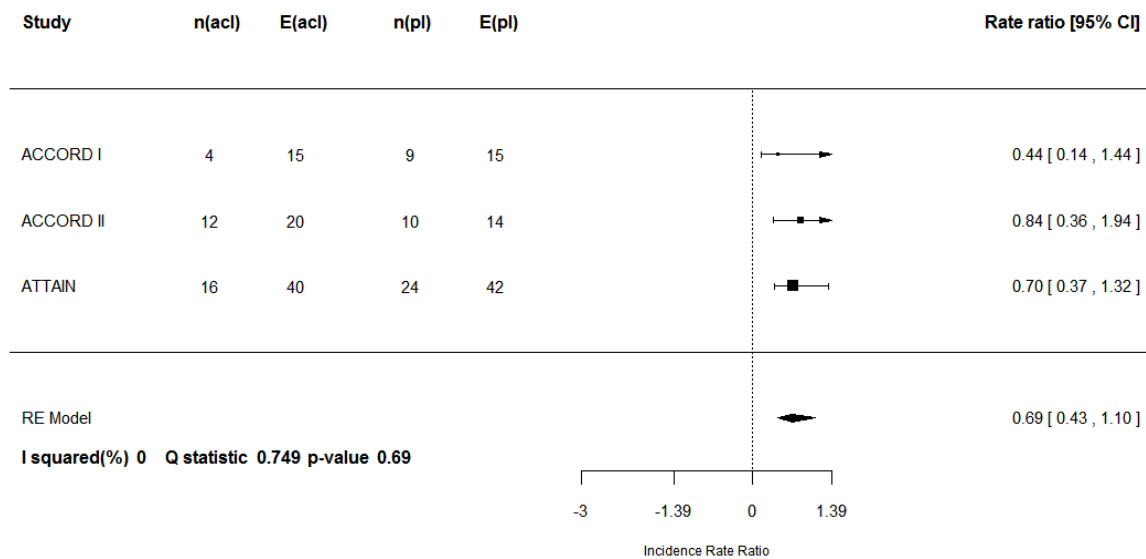


Abbildung 106: Meta-Analyse für mittelschwere COPD für Exazerbationsrate aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Abbildung 106 zeigt, dass Acclidiniumbromid auch für Patienten mit schwerer COPD bei der Reduzierung der Exazerbationsrate nicht signifikant wirksamer als ein Placebo ist (gepooltes Ratenverhältnis 0,69 [KI 95% 0,43; 1,10]). Der Ausreißertest zeigte keine signifikante Heterogenität ($Q=0,75$; $p=0,69$; $I^2=0$).

4.3.2.1.3.17.10 Krankenhausaufenthalte aufgrund von Exazerbationen

Tabelle 4.3-219: Ergebnis der Subgruppenanalyse für Krankenhausaufenthalte aufgrund von Exazerbationen (12 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Follow-up	Gruppenunterschied ¹
		N	N (%)	OR (SE) KI [] p-wert
ACCORD I				
Stufe II	Aclidiniumbromid	118	0 (0%)	-
	Placebo	111	1 (0,9%)	
Stufe III	Aclidiniumbromid	68	2 (4,4%)	-
	Placebo	73	0 (0%)	
ACCORD II				
Stufe II	Aclidiniumbromid	79	0 (0%)	-
	Placebo	113	2 (1,8%)	
Stufe III	Aclidiniumbromid	96	1 (1,0%)	0,34 (1,24) KI [0,03-3,85] p= 0,39
	Placebo	67	2 (3,0%)	
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell N = Anzahl, ITT = Intention to Treat, SE = Standardfehler, KI = Konfidenzintervall, - = Berechnung einer Odds Ratio ist aufgrund weniger oder keiner Ereignisse in den Behandlungsarmen nicht informativ				

Tabelle 4.3-219 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für die 12 Wochen nach Baseline registrierten Krankenhausaufenthalte aufgrund von Exazerbationen. Die Anzahl von Ereignissen war in allen Behandlungsgruppen sehr gering oder sogar Null. Da die statistische Aussagekraft einer Meta-Analyse wäre unzureichend, daher wurde keine Meta-Analyse durchgeführt. Aus dem gleichen Grund wurde kein Test auf Interaktion durchgeführt. Es ergeben sich keine Konsequenzen für die Nutzenbewertung.

Tabelle 4.3-220: Ergebnis der Subgruppenanalyse für Krankenhausaufenthalte aufgrund von Exazerbationen (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Follow-up	Gruppenunterschied ¹
		N	N (%)	OR (SE) KI [] p-wert
ATTAIN				
Stufe II	Aclidiniumbromid	184	1 (0,5%)	0,19 (1,10) KI [0,02; 1,63] p=0,130
	Placebo	178	5 (2,8%)	
Stufe III	Aclidiniumbromid	84	1 (1,2%)	0,21 (1,11) KI [0,02; 1,83] p=0,158
	Placebo	92	5 (8,7%)	

¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion
N = Anzahl, ITT = Intention to Treat, SE = Standardfehler, KI = Konfidenzintervall

Die Studienergebnisse pro COPD-Schweregrad für die 24 Wochen nach Baseline registrierten Krankenhausaufenthalte aufgrund von Exazerbationen sind in Tabelle 4.3-220 dargestellt. Nur die ATTAIN-Studie lieferte Daten über dieses Outcome pro COPD-Schweregrad. Die Behandlung mit Aclidiniumbromid führte zu weniger Krankenhausaufenthalten aufgrund von Exazerbationen. Da die Anzahl der Ereignisse jedoch sehr gering war, sollten diese Ergebnisse mit Vorsicht interpretiert werden.

4.3.2.1.3.17.11 SGRQ-Score

Tabelle 4.3-221: Ergebnisse der Subgruppenanalyse für den SGRQ-Score aus RCT für indirekte Vergleiche

Studie	Studienarm (Patienten)	Baseline	Follow-Up	Veränderung	
		Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-Wert
ACCORD I					
12 Wochen					
Stufe II	AB (n=117)	46,4 (17,9)	41,4 (18,4)	-4,78 (10,9)	-2,77 (1,4)
	Pbo (n=108)	41,7 (16,1)	39,4 (16,7)	-2,01 (10,4)	KI [-5,56; 0,01] p= 0,01
Stufe III	AB (n=68)	50,5 (17,2)	46,5 (16,2)	-4,06 (11,7)	-2,52 (2,01)
	Pbo (n=71)	50,1 (15,4)	48,3 (17,7)	-1,53 (12,0)	KI [-6,47; 1,42] p= 0,21
<i>Interaktion p=0,72</i>					
ACCORD II					
12 Wochen					
Stufe II	AB (n=78)	46,2 (17,7)	41,5 (19,6)	-4,85 (13,9)	0,44 (2,2)
	Pbo (n=111)	47,2 (17,6)	42,2 (20,4)	-5,29 (16,5)	KI [-3,9; 4,8] p= 0,84
Stufe III	AB (n=92)	54,1 (15,6)	47,9 (17,6)	-6,28 (15,6)	-3,53 (2,31)
	Pbo (n=65)	52,8 (16,3)	50,6 (18,4)	-2,75 (13,2)	KI [-8,05; 0,99] p= 0,13
<i>Interaktion p=0,39</i>					
ATTAIN					
12 Wochen					
Stufe II	AB (n=184)	45,6 (18,4)	40,2 (16,9)	-5,38 (11,5)	-3,46 (1,2)
	Pbo (n=177)	42,6 (16,4)	40,7 (16,3)	-1,93 (11,3)	KI [-5,8; -1,1] p=0,004
Stufe III	AB (n=84)	51,1 (17,5)	46,2 (17,1)	-4,92 (10,7)	-2,32 (1,65)
	Pbo (n=91)	49,3 (16,3)	46,7 (15,4)	-2,60 (11,1)	KI [-5,56; 0,92] p= 0,16
<i>Interaktion p=0,64</i>					
24 Wochen					
Stufe II	AB (n=184)	45,6 (18,4)	38,4 (17,4)	-7,16 (13,9)	-5,92 (1,45)
	Pbo (n=177)	42,6 (16,4)	41,4 (17,1)	-1,24 (11,3)	KI [-8,76; -3,07] p=<0,001
Stufe III	AB (n=84)	51,1 (17,5)	45,8 (18,0)	-5,32 (14,4)	-2,58 (2,23)
	Pbo (n=91)	49,3 (16,3)	46,7 (14,8)	-2,74 (11,1)	KI [-6,95; 1,78] p=0,247

<i>Interaktion p=0,17*</i>
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion AB: Acclidiniumbromid, Pbo: Placebo, SD: Standardabweichung, SE: Standardfehler, CFB: Change from Baseline, KI: Konfidenzintervall

Tabelle 4.3-208 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für den 12 und 24 Wochen nach Baseline erhobenen SGRQ-Score. Alle mit Acclidiniumbromid behandelten Patienten wiesen unabhängig vom COPD-Schweregrad eine verbesserte Lebensqualität auf (definiert als geringerer SGRQ-Score). Nach 12 Wochen gab es keinen Hinweis auf Interaktion ($p > 0,2$). Nach 24 Wochen führte die Behandlung mit Acclidiniumbromid bei beiden Schweregrad-Gruppen zu einer Verbesserung der Lebensqualität. Diese war bei den Patienten mit COPD-Stufe II im Gegensatz zu Stufe III signifikant ausgeprägt. Der Hinweis auf Interaktion ($p = 0,17$) beeinflusst aufgrund des nach 12 Wochen ermittelten homogenen Effektes bzgl. der Lebensqualität nicht die Nutzenbewertung.

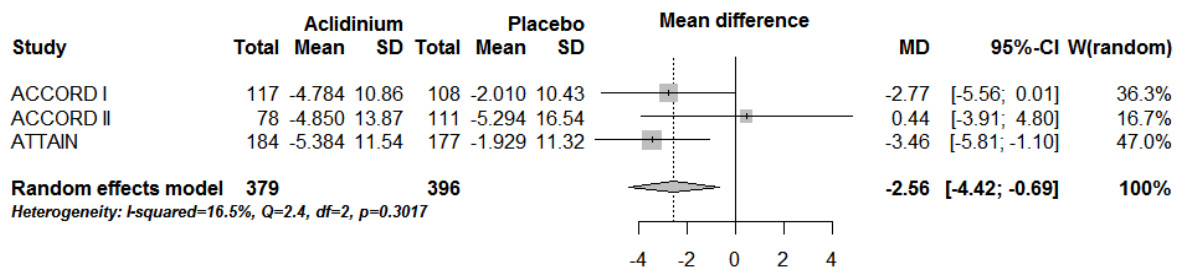


Abbildung 107: Meta-Analyse für mittelschwere COPD für den SGRQ-Wert (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

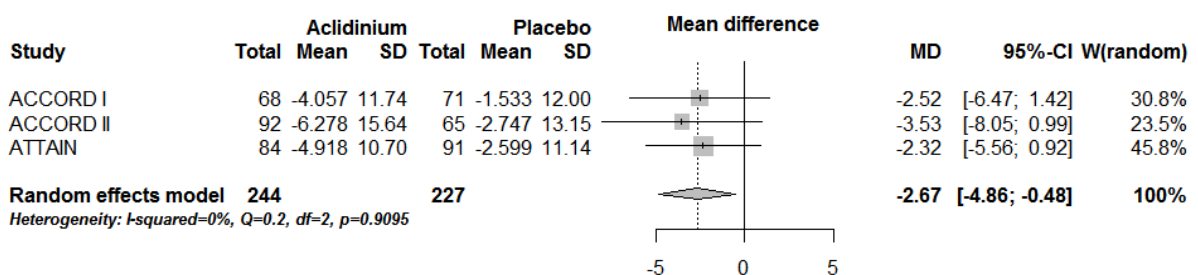


Abbildung 108: Meta-Analyse für schwere COPD für den SGRQ-Wert (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Für Patienten mit mittelschwerer COPD war die Behandlung mit Aclidiniumbromid in Bezug auf die Verbesserung der Lebensqualität nach 12 Wochen wirksamer als mit Placebo (Abbildung 107; Differenz -2,56 [-4,42; -0,69]). ACCORD II zeigt im Vergleich mit den Studien ACCORD I und ATTAIN gegensätzliche Behandlungseffekte. Für dieses Outcome wurden bei Patienten mit mittelschwerer COPD keine signifikante Studien-Heterogenität erkannt ($Q=2,4$; $p=0,30$; $I^2=16,5\%$). Die Meta-Analyse unter exklusiver Berücksichtigung von Patienten mit schwerer COPD ergab im Vergleich mit Patienten mit mittelschwerer COPD ähnliche Ergebnisse (Abbildung 108; Differenz -2,67 [-4,68; -0,48]). Es wurde keine signifikante Studien-Heterogenität festgestellt ($Q=0,2$; $p=0,91$; $I^2=0\%$).

4.3.2.1.3.17.12 SGRQ-Responder

Tabelle 4.3-222: Ergebnisse der Subgruppenanalyse für SGRQ-Responder aus RCT für indirekte Vergleiche

Studie	Studienarm (Patienten)	Follow-Up	Gruppenunterschied
		N Responder (%)	OR ¹ (SE) KI [p-Wert
ACCORD I			
12 Wochen			
Stufe II	AB (n=117)	52 (44,4%)	1,21 (0,27) KI [0,71; 2,06] p= 0,48
	Pbo (n=108)	43 (39,8%)	
Stufe III	AB (n=68)	28 (41,2%)	1,78 (0,36) KI [0,88; 3,62]p= 0,11
	Pbo (n=71)	20 (28,2%)	
<i>Interaktion p=0,24</i>			
ACCORD II			
12 Wochen			
Stufe II	AB (n=78)	33 (42,3%)	1,08 (0,30) KI [0,60; 1,94] p= 0,81
	Pbo (n=111)	45 (40,5%)	
Stufe III	AB (n=92)	44 (47,8%)	1,67 (0,33) KI [0,87; 3,21] p= 0,12
	Pbo (n=65)	23 (35,4%)	
<i>Interaktion p=0,42</i>			
ATTAIN			
12 Wochen			
Stufe II	AB (n=184)	98 (53,3%)	1,96 (0,21) KI [1,29; 2,99] p= 0,002
	Pbo (n=177)	65 (36,7%)	
Stufe III	AB (n=84)	40 (47,6%)	1,16 (0,30) KI [0,64; 2,10] p= 0,63
	Pbo (n=91)	40 (44,0%)	
<i>Interaktion p=0,16*</i>			
24 Wochen			
Stufe II	AB (n=184)	108 (58,7%)	2,57 (0,22) KI [1,68; 3,93] p=<0,001
	Pbo (n=177)	63 (35,6%)	
Stufe III	AB (n=84)	44 (52,4%)	1,28 (0,30) KI [0,71; 2,33] p=0,4107
	Pbo (n=91)	42 (46,2%)	
<i>Interaktion p=0,06*</i>			

¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion
 AB: Acclidiniumbromid, Pbo: Placebo, N=Anzahl der Patienten, Responder = Anzahl der Patienten mit einer Reduzierung des SGRQ-Gesamtwerts von mindestens 4 Einheiten im Vergleich zum Ausgangswert, SE=Standardfehler, KI=Konfidenzintervall

Tabelle 4.3-222 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad nach 12 und 24 Wochen für die nach Baseline ermittelte SGRQ-Responder. Alle mit Acclidiniumbromid behandelten Patienten wiesen unabhängig vom COPD-Schweregrad eine verbesserte Lebensqualität auf (definiert als höhere Anzahl von SGRQ-Resondern). Nach 12 Wochen gab es keinen Hinweis auf Interaktion ($p > 0,2$). Nach 24 Wochen führte die Behandlung mit Acclidiniumbromid bei beiden Schweregrad-Gruppen (II und III) zu einer Verbesserung der Lebensqualität. Diese war bei den Patienten mit COPD-Stufe II im Gegensatz zu Stufe III signifikant ausgeprägt. Der Hinweis auf Interaktion ($p = 0,06$) beeinflusst aufgrund den nach 12 Wochen mit Hilfe der Meta-Analyse (Abbildung 109 und Abbildung 110) über alle Studien ermittelten homogenen Effekt bzgl. der Lebensqualität nicht die Nutzenbewertung.

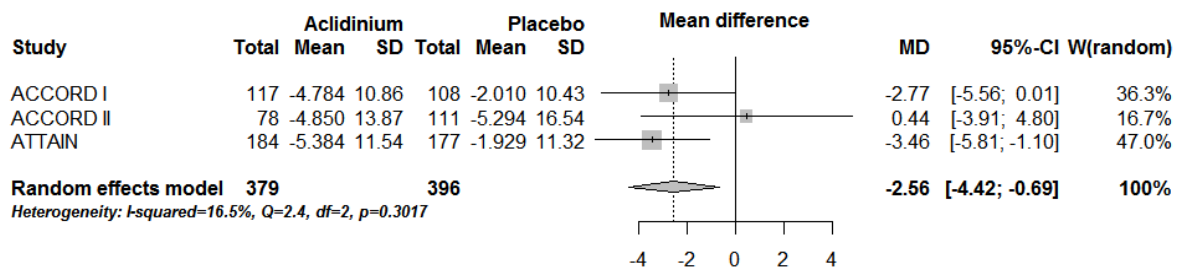


Abbildung 109: Meta-Analyse für mittelschwere COPD für den SGRQ-Wert (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

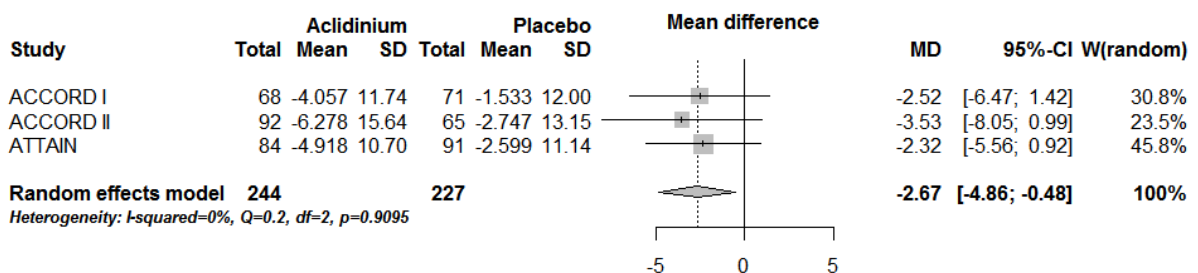


Abbildung 110: Meta-Analyse für schwere COPD für den SGRQ-Wert (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Für Patienten mit mittelschwerer COPD war die Behandlung mit Aclidiniumbromid in Bezug auf die Verbesserung der Lebensqualität nach 12 Wochen wirksamer als mit Placebo (Abbildung 1; Differenz -2,56 [-4,42; -0,69]) (Abbildung 109). ACCORD II zeigt im Vergleich mit den Studien ACCORD I und ATTAIN gegensätzliche Behandlungseffekte. Für dieses Outcome wurden bei Patienten mit mittelschwerer COPD keine signifikante Studien-Heterogenität erkannt ($Q=2,4$; $p=0,30$; $I^2=16,5\%$). Die Meta-Analyse unter exklusiver Berücksichtigung von Patienten mit schwerer COPD ergab im Vergleich mit Patienten mit mittelschwerer COPD ähnliche Ergebnisse (Abbildung 110; Differenz -2,67 [-4,68; -0,48]). Es wurde keine signifikante Studien-Heterogenität erkannt ($Q=0,2$; $p=0,91$; $I^2=0\%$). Der Hinweis auf Interaktion ($p=0,16$) bei der ATTAIN-Studie beeinflusst aufgrund des signifikanten Gesamteffektschätzers der Meta-Analysen nicht die Nutzenbewertung.

4.3.2.1.3.17.13 Unerwünschte Ereignisse

Tabelle 4.3-223: Ergebnisse der Subgruppenanalyse für Unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 24 Wochen)

Studie	Studienarm	Patienten	Follow-up	Gruppenunterschied ¹
		N	N (%)	OR (SE) KI [] p-wert
ACCORD I				
Stufe II	Aclidiniumbromid	118	49 (41,5%)	0,58 (0,27) KI [0,34; 0,98] p=0,043
	Placebo	111	61 (55,0%)	
Stufe III	Aclidiniumbromid	68	34 (50,0%)	1,09 (0,34) [0,56; 2,10] p= 0,807
	Placebo	73	35 (47,9%)	
<i>Interaktion p=0,12*</i>				
ACCORD II				
Stufe II	Aclidiniumbromid	79	43 (54,4%)	1,35 (0,29) KI [0,76; 2,41] p=0,305
	Placebo	113	53 (46,9%)	
Stufe III	Aclidiniumbromid	96	46 (47,9%)	0,79 (0,32) KI [0,42; 1,48] p= 0,465
	Placebo	67	36 (53,7%)	
<i>Interaktion p=0,26</i>				
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, * Hinweis auf Interaktion N=Anzahl der Patienten, SE=Standardfehler, KI=Konfidenzintervall				

Tabelle 4.3-223 zeigt die Patienten mit mittelschwerer und schwerer COPD, die innerhalb von 12 Wochen nach Baseline unerwünschte Ereignisse berichteten. In der ACCOD I-Studie

berichten die mittelschweren COPD-Patienten weniger, die schweren COPD-Patienten mehr Fälle als Placebo. Dieser deskriptiv beobachtete Unterschied wird durch das Ergebnis des Interaktionstests ($p=0,12$; Hinweis auf Interaktion) bestätigt. In der ACCORD II-Studie verhält es sich umgekehrt, das Ergebnis der Interaktionstests ($p=0,26$) lässt jedoch auf eine zufällige Abweichung schließen.

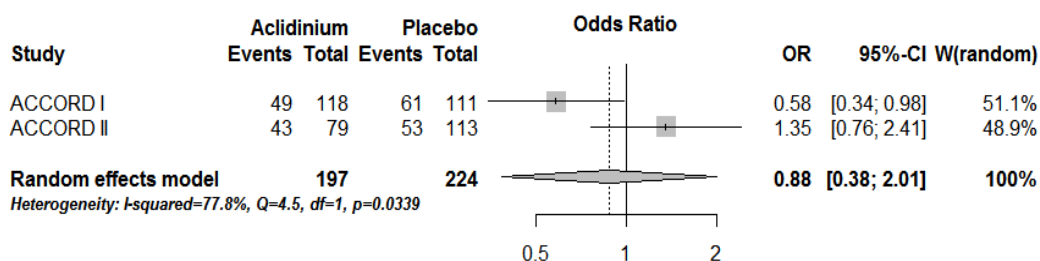


Abbildung 111: Meta-Analyse für mittelschwere COPD für unerwünschte Ereignisse (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

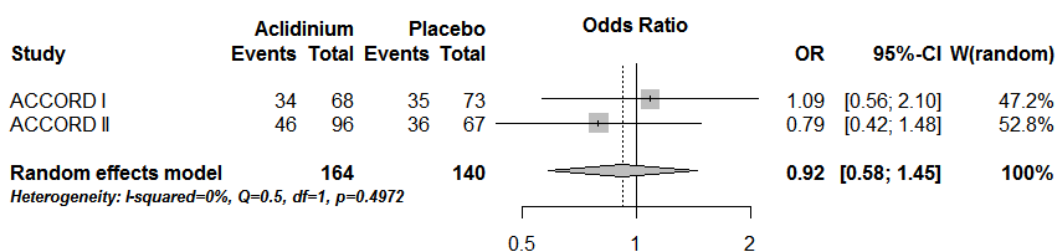


Abbildung 112: Meta-Analyse für schwere COPD für unerwünschte Ereignisse (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Abbildung 111 zeigt, dass sich Patienten mit mittelschwerer COPD unter Acclidiniumbromid in Bezug auf unerwünschte Ereignisse nicht signifikant von Placebo-Patienten unterscheiden (gepoolte OR 0,88 [KI 95% 0,38; 2,01]). Die Punktschätzer der beiden Studienergebnisse unterschieden sich deutlich. Für dieses Outcome existierte entsprechend eine signifikante Studien-Heterogenität ($Q=4,5$; $p=0,03$; $I^2=77,8\%$).

Abbildung 112 zeigt ebenfalls, dass zwischen den mit Acclidiniumbromid und Placebo behandelten für Patienten mit schwerer COPD in Bezug auf unerwünschte Ereignisse kein Unterschied besteht (gepoolte OR 0,92 [KI 95% 0,58; 1,45]). Die Punktschätzer der einzelnen Studienergebnisse unterschieden sich jedoch nicht so deutlich wie bei der Meta-Analyse der mittelschweren COPD-Patienten. Dementsprechend wurde keine signifikante Studien-Heterogenität nachgewiesen ($Q=0,5$; $p=0,50$; $I^2=0\%$).

Auch mit Hilfe der Meta-Analysen lässt sich keine systematische Effektbeeinflussung durch den Schweregrad erkennen. Es können keine Auswirkungen auf die Nutzenbewertung abgeleitet werden.

Tabelle 4.3-224: Ergebnisse der Subgruppenanalyse für unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Gruppenunterschied ¹
		N	N (%)	OR (SE) KI [] p-wert
ATTAIN				
Stufe II	Acidinium	184	99 (53,8%)	0,93 (0,21) KI [0,61; 1,41] p=0,729
	Placebo	178	99 (55,6%)	
Stufe III	Acidinium	84	45 (53,6%)	0,78 (0,31) KI [0,43; 1,41] p=0,406 <i>Interaktion p=0,57</i>
	Placebo	92	55 (59,8%)	
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell N=Anzahl der Patienten, SE=Standardfehler, KI=Konfidenzintervall				

Tabelle 4.3-224 zeigt die Studienergebnisse pro COPD-Schweregrad für die 24 Wochen nach Baseline registrierten Patienten mit unerwünschten Ereignissen. Die Anzahl der unerwünschten Ereignisse war bei der Behandlung mit Acidiniumbromid und Placebo von Patienten mit sowohl mittelschwerer als auch schwerer COPD ähnlich. Es lag kein Hinweis auf Interaktion vor.

4.3.2.1.3.17.14 Schwere unerwünschte Ereignisse

Tabelle 4.3-225: Ergebnisse der Subgruppenanalyse für schwere unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 12 Wochen)

Studie	Studienarm	Patienten (ITT)		Gruppenunterschied ¹
		N	N (%)	
ACCORD I				
Stufe II	Aclidiniumbromid	118	3 (2,5%)	0,94 (0,83) [0,19; 4,75] p=0,94
	Placebo	111	3 (2,7%)	
Stufe III	Aclidiniumbromid	68	3 (4,4%)	3,32 (1,17) [0,34; 32,75] p=0,30 <i>Interaktion p=0,36</i>
	Placebo	73	1 (1,4%)	
ACCORD II				
Stufe II	Aclidiniumbromid	79	0 (0%)	0,19 (1,08) [0,02; 1,61] p=0,10
	Placebo	113	7 (6,2%)	
Stufe III	Aclidiniumbromid	96	8 (8,3%)	1,43 (0,63) [0,41; 4,96] p=0,57 <i>Interaktion p=0,99</i>
	Placebo	67	4 (6,0%)	
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell N=Anzahl der Patienten, SE=Standardfehler, KI=Konfidenzintervall, ITT=Intention-to-treat-Population, OR=Odds Ratio				

Tabelle 4.3-225 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für die 12 Wochen nach Baseline registrierten Patienten mit schweren unerwünschten Ereignissen. Die Behandlung mit Aclidiniumbromid führte im Vergleich zu Placebo in beiden Studien und Schweregraden zu ähnlichen schweren unerwünschten Ereignissen.

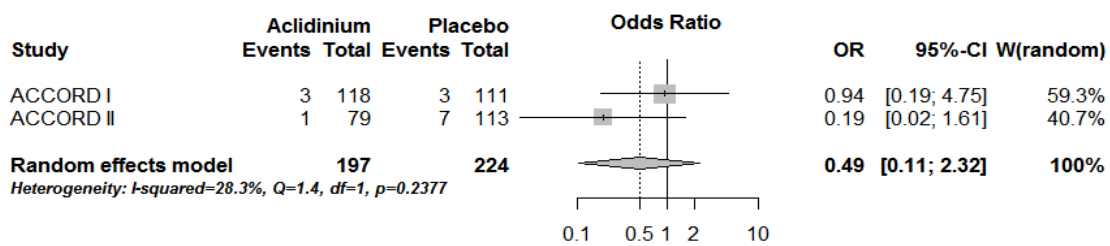


Abbildung 113: Meta-Analyse für mittelschwere COPD für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Acridiniumbromid im Vergleich zu Placebo

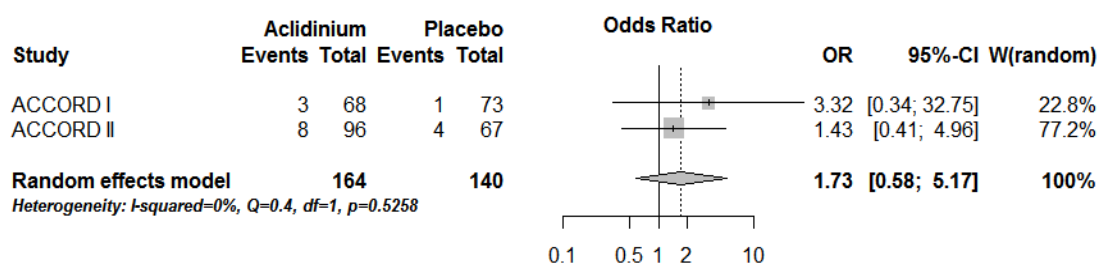


Abbildung 114: Meta-Analyse für schwere COPD für schwere unerwünschte Ereignisse (12 Wochen) aus RCT; Acridiniumbromid im Vergleich zu Placebo

Auch die Meta-Analyse (Abbildung 113) zeigt, dass sich die Anzahl der Patienten mit mittelschwerer COPD unter Acridiniumbromid in Bezug auf schwere unerwünschte Ereignisse nicht von den Patienten unter Placebo unterscheiden (gepoolte OR 0,49 [KI 95% 0,11; 2,32]). Die Anzahl von Ereignissen war jedoch sehr gering, was die statistische Aussagekraft begrenzt. ACCORD II berichtete keine schweren unerwünschten Ereignisse in der Behandlungsgruppe mit Acridiniumbromid. Da die Durchführung einer Meta-Analyse ohne Ereignisse nicht möglich ist, wurde in den Behandlungsarmen ohne Ereignisse ein Kontinuitäts-Korrekturfaktor von 1 eingefügt. Dies beeinflusst jedoch nicht die Schlussfolgerung der Analyse.

Die Meta-Analyse der Patienten mit schwerer COPD (Abbildung 114) zeigt ein ähnliches Ergebnis wie das der Patienten mit mittelschwerer COPD (gepoolte OR 1,73 [KI 95% 0,58; 5,17]). Keine der beiden Analysen zeigte eine signifikante Studien-Heterogenität für dieses Outcome (mittelschwere COPD: Q=1,4; p=0,24; I²=28,3%; schwere COPD: Q=0,4; p=0,53; I²=0%).

Tabelle 4.3-226: Ergebnisse der Subgruppenanalyse für schwere unerwünschte Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Gruppenunterschied ¹
		N	N (%)	OR (SE) KI [] p-wert
ATTAIN				
Stufe II	Aclidinium	184	9 (4.9%)	0.97 (0.48) [0.37; 2.49] p=0.943
	Placebo	178	9 (5.1%)	
Stufe III	Aclidinium	84	5 (6.0%)	0.91 (0.63) [0.27; 3.09] p=0.876
	Placebo	92	6 (6.5%)	
<i>Interaktion p=0,87</i>				
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell N=Anzahl der Patienten, SE=Standardfehler, KI=Konfidenzintervall, ITT=Intention-to-treat-Population, OR=Odds Ratio				

Tabelle 4.3-226 zeigt die Studienergebnisse pro COPD-Schweregrad für die 24 Wochen nach Baseline registrierten Patienten mit schweren unerwünschten Ereignissen. Nur die ATTAIN-Studie berichtete über dieses Outcome pro COPD-Schweregrad. Die Behandlung mit Aclidiniumbromid führte im Vergleich zu Placebo in beiden COPD-Schweregraden zu ähnlichen schweren unerwünschten Ereignissen.

4.3.2.1.3.17.15 Studienabbruch aufgrund unerwünschter Ereignisse

Tabelle 4.3-227: Ergebnisse der Subgruppenanalyse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 12 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Gruppenunterschied ¹
		N	N (%)	OR (SE) KI [] p-wert
ACCORD I				
Stufe II	Aclidiniumbromid	118	3 (2,54%)	0,39 (0,70) KI [0,10; 1,54] p= 0,18
	Placebo	111	7 (6,31%)	
Stufe III	Aclidiniumbromid	68	4 (5,88%)	0,85 (0,69) KI [0,22; 3,31] p= 0,81 <i>Interaktion p=0,35</i>
	Placebo	73	5 (6,85%)	
ACCORD II				
Stufe II	Aclidiniumbromid	79	5 (6,33%)	1,46 (0,65) KI [0,41; 5,22] p= 0,56
	Placebo	113	5 (4,42%)	
Stufe III	Aclidiniumbromid	96	8 (8,33%)	1,94 (0,70) KI [0,50; 7,60] p= 0,34 <i>Interaktion p=0,78</i>
	Placebo	67	3 (4,48%)	
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell N=Anzahl der Patienten, SE=Standardfehler, KI=Konfidenzintervall, ITT=Intention-to-treat-Population, OR=Odds Ratio				

Tabelle 4.3-227 zeigt die einzelnen Studienergebnisse pro COPD-Schweregrad für die 12 Wochen nach Baseline registrierten Studienabbrüche aufgrund von unerwünschten Ereignissen. Die Studienabbrüche in den Sub- und Vergleichsgruppen waren in beiden Studien nicht unterschiedlich. Dies bestätigen auch die beiden folgenden Meta-Analysen.

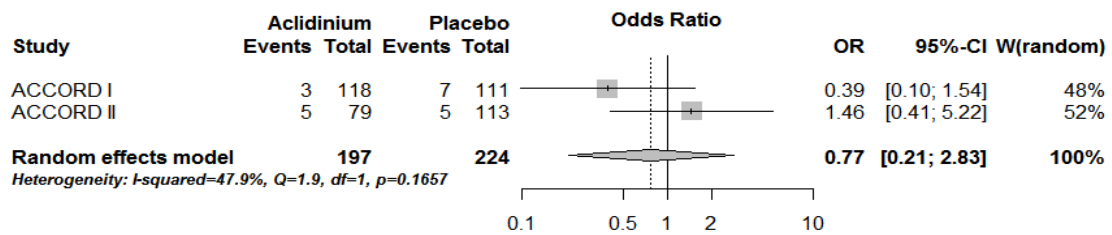


Abbildung 115: Meta-Analyse für mittelschwere COPD für Ausschlüsse aufgrund unerwünschter Ereignisse (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

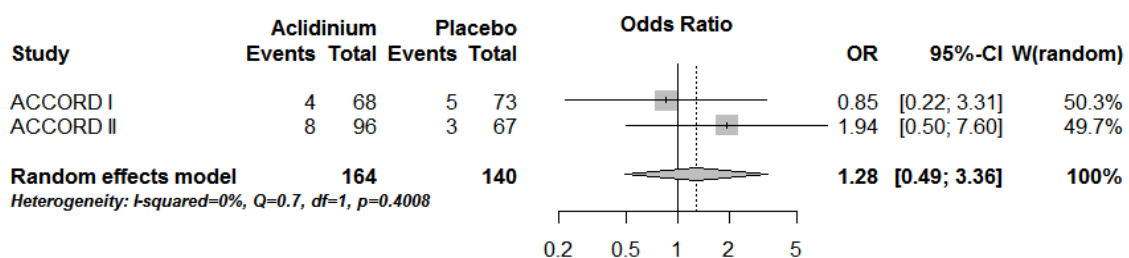


Abbildung 116: Meta-Analyse für schwere COPD für Ausschlüsse aufgrund unerwünschter Ereignisse (12 Wochen) aus RCT; Acclidiniumbromid im Vergleich zu Placebo

Die Meta-Analyse (Abbildung 115) zeigt, dass sich die Anzahl der Patienten mit mittelschwerer COPD unter Acclidiniumbromid in Bezug auf Studienabbrüche aufgrund unerwünschter Ereignisse nicht von den Patienten unter Placebo unterscheiden (gepoolte OR 0,77 [KI 95% 0,21; 2,83]). Für dieses Outcome gab es keine signifikante Heterogenität in der Studie ($Q=1,9$; $p=0,17$; $I^2=47,9%$).

Abbildung 116 zeigt, dass die Ergebnisse für Patienten mit schwerer COPD in Bezug auf Studienabbrüche aufgrund unerwünschter Ereignisse unter Acclidiniumbromid und Placebo ebenfalls ähnlich waren (gepoolte OR 1,28 [KI 95% 0,49; 3,36]). Es wurde keine Studien-Heterogenität erkannt ($Q=0,7$; $p=0,40$; $I^2=0%$).

Tabelle 4.3-228: Ergebnisse der Subgruppenanalyse für Studienabbrüche aufgrund unerwünschter Ereignisse aus RCT für indirekte Vergleiche (Studien mit einer Dauer von 24 Wochen)

Studie	Studienarm	Patienten (ITT)	Follow-up	Gruppenunterschied ¹
		N	N (%)	OR (SE) KI [] p-wert
ATTAIN				
Stufe II	Aclidiniumbromid	184	3 (1,6%)	0,28 (0,67) KI [0,08; 1,03] p=0,055
	Placebo	178	10 (5,6%)	
Stufe III	Aclidiniumbromid	84	5 (6,0%)	5,76 (1,11) KI [0,66; 50,35] p=0,114
	Placebo	92	1 (1,1%)	
<i>Interaktion p=0,02**</i>				
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, ** Beleg für Interaktion N = Anzahl, ITT = Intention to Treat, SE = Standardfehler, KI = Konfidenzintervall, OR=Odds Ratio				

Tabelle 4.3-228 zeigt die Studienergebnisse pro COPD-Schweregrad für die 24 Wochen nach Baseline registrierten Studienabbrüchen aufgrund unerwünschter Ereignisse. Bei Patienten mit mittelschwerer COPD führte die Behandlung mit Aclidiniumbromid nach 24 Wochen zu einer geringeren Anzahl von Studienabbrüchen aufgrund unerwünschter Ereignisse. Bei Patienten mit schwerer COPD führte die Behandlung mit Aclidiniumbromid nach 24 Wochen im Vergleich zu Placebo zu einer ähnlichen Anzahl von Studienabbrüchen aufgrund unerwünschter Ereignisse. Aufgrund der geringen Anzahl der Ereignisse weist das Konfidenzintervall und somit das Ergebnis eine erhebliche Unsicherheit auf. Für den Interaktionstest gilt dies entsprechend, sodass sein Ergebnis keine Auswirkungen auf die Nutzenbewertung hat.

4.3.2.1.3.17.16 Zusammenfassung der Subgruppenanalysen nach Schweregrad

Wie in den vorangegangenen Tabellen dargestellt, zeigten sich für einen Großteil der verfügbaren Daten keine unterschiedlichen Effekte in den Subgruppen. Es ergaben sich einige Fallkonstellationen, die Hinweise (p-Wert zwischen 0,05 und 0,2) oder Belege ($p < 0,05$) auf unterschiedliche Effekte erbrachten. Bei den Endpunkten SGRQ (Responder) und TDI (Responder) konnten diese nach Durchführung der Meta-Analyse und Feststellung eines signifikanten Gesamteffektschätzers verworfen werden. Die Subgruppenergebnisse der unerwünschten Arzneimittelwirkungen sind aufgrund der meist geringen Anzahl von Ereignissen und Heterogenität der Ergebnisse nicht eindeutig zu interpretieren und haben keinen Einfluss auf die Nutzenbewertung. Bei den Symptom-Fragebögen (EXACT, morgendliche und nächtliche Symptome) war die Aufklärung der Interaktionen mit Hilfe von Meta-Analysen nicht möglich. Die zum Teil auftretenden deutlicheren Effekte einer Subgruppe waren nicht systematisch einem COPD-Schweregrad zuzuordnen. Ableitungen bzgl. der Nutzenbewertung sind deshalb nicht möglich. Aufgrund dessen und aufgrund der fehlenden Subgruppendaten der Tiotropiumbromid-Studien wird die Zusatznutzenbewertung nicht beeinflusst.

Es war nicht möglich Subgruppenanalysen für die Outcomes „kardiovaskuläre Morbidität und Mortalität“ und „Letalität und Gesamtsterblichkeit“ durchzuführen, da die Anzahl der Ereignisse sehr begrenzt war. Außerdem wäre die statistische Aussagekraft einer Meta-Analyse zum Aufzeigen eines bedeutsamen Ergebnisses unzureichend.

4.3.2.1.3.17.17 A priori geplante Subgruppenanalysen

Tabelle 4.3-229: Ergebnisse der a priori geplanten Subgruppenanalyse für CFB Trough-FEV1 (12 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten	Baseline	Follow-up	Veränderung	
		(ITT)	Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-wert
ACCORD II						
COPD-Schweregrad						
I & II	AB	79	1599 (500)	-	61 (25)	89 (-) KI [30; 150] p=0,0058
	Pbo	115	1698 (468)	-	-29 (21)	
III & IV	AB	96	961 (312)	-	71 (20)	56 (-) KI [-10, 120] p=0,0733
	Pbo	67	1050 (302)	-	15 (24)	

¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, AB = Acridiniumbromid, Pbo = Placebo, ITT = Intention to Treat, N = Anzahl, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall

Tabelle 4.3-230: Ergebnisse der a priori gelanten Subgruppenanalyse für CFB Trough-FEV1 (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-wert
ATTAIN						
Alter						
< 60 Jahre	AB	96	1603 (583)	1688 (618)	85 (326)	158 (36)
	Pbo	102	1549 (529)	1483 (556)	-67 (315)	KI [87; 229] p<0,0001
60 – 70 Jahre	AB	108	1518 (503)	1549 (481)	31 (259)	87 (34)
	Pbo	121	1490 (494)	1439 (503)	-51 (203)	KI [21; 153] p=0,0095
> 70 Jahre	AB	65	1350 (434)	1443 (467)	93 (200)	146 (48)
	Pbo	50	1423 (378)	1367 (363)	-56 (157)	KI [52; 240] p=0,0024
Geschlecht						
Männlich	AB	182	1653 (526)	1728 (532)	75 (278)	149 (26)
	Pbo	189	1592 (493)	1527 (501)	-64 (234)	KI [97; 201] p<0,0001
Weiblich	AB	87	1203 (370)	1249 (383)	46 (264)	81 (39)
	Pbo	84	1293 (416)	1250 (450)	-43 (266)	[4; 157] p=0,0382
COPD-Schweregrad						
Stufe II	AB	184	1690 (490)	1737 (510)	47 (267)	123 (25)
	Pbo	178	1680 (462)	1608 (477)	-72 (264)	KI [73; 172] p<0,0001
Stufe III	AB	84	1119 (357)	1204 (391)	85 (220)	125 (36)
	Pbo	92	1138 (291)	1103 (302)	-36 (187)	KI [54; 195] p=0,0006
Raucherstatus						
Raucher	AB	148	1500 (500)	1579 (522)	79 (294)	163 (30)
	Pbo	144	1467 (464)	1390 (482)	-77 (267)	KI [105; 221] p<0,0001
Ex-Raucher	AB	121	1517 (557)	1566 (557)	49 (246)	87 (32)
	Pbo	129	1537 (516)	1501 (519)	-36 (214)	KI [24; 150] p=0,0066
Reversibilität						
Ja	AB	81	1541 (422)	1619 (446)	78 (265)	140 (38)
	Pbo	81	1385 (452)	1344 (452)	-41 (221)	KI [66; 214] p=0,0002
Nein	AB	187	1498 (562)	1549 (570)	51 (248)	116 (25)
	Pbo	189	1543 (492)	1475 (498)	-68 (249)	KI [68; 165] p<0,0001

BMI						
Untergewicht bis normal (<25 kg/m ²)	AB	97	1386 (479)	1440 (470)	54 (182)	126 (35)
	Pbo	113	1409 (485)	1344 (486)	-65 (222)	KI [58; 195] p=0,0003
Übergewicht (25-30 kg/m ²)	AB	99	1499 (524)	1599 (563)	100 (305)	163 (37)
	Pbo	92	1594 (488)	1522 (543)	-72 (232)	KI [91; 235] p<0,0001
Adipös (≥ 30 kg/m ²)	AB	73	1681 (544)	1716 (550)	35 (324)	75 (43)
	Pbo	68	1524 (479)	1498 (446)	-26 (292)	KI [-9; 159] p=0,0788
Verwendung von ICS						
Ja	AB	128	1447 (477)	1501 (469)	54 (272)	99 (31)
	Pbo	145	1458 (491)	1412 (492)	-45 (250)	KI [38; 159] p=0,0014
Nein	AB	141	1562 (561)	1639 (587)	76 (275)	156 (31)
	Pbo	128	1547 (486)	1476 (513)	-72 (237)	KI [95; 217] p=<0,0001
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, AB = Aclidiniumbromid, Pbo = Placebo, ITT = Intention to Treat, N = Anzahl, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall						

Tabelle 4.3-231: Ergebnisse der a priori gelanten Subgruppenanalyse für CFB Peak-FEV1 (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Baseline	Follow-up	Veränderung	
		N	Mean (SD)	Mean (SD)	CFB (SD)	Gruppendifferenz ¹ (SE) KI [] p-wert
ATTAIN						
Alter						
< 60 Jahre	AB	96	1603 (583)	1912 (659)	309 (345)	276 (39)
	Pbo	102	1549 (529)	1588 (560)	38 (323)	KI [199; 352] p<0,0001
60 – 70 Jahre	AB	108	1518 (503)	1732 (495)	213 (263)	182 (36)
	Pbo	121	1490 (494)	1527 (520)	37 (202)	KI [110; 253] p<0,0001
> 70 Jahre	AB	65	1350 (434)	1559 (486)	209 (226)	150 (52)
	Pbo	50	1423 (378)	1492 (424)	69 (179)	KI [39; 242] p=0,0068
Geschlecht						
Männlich	AB	182	1653 (526)	1919 (568)	265 (297)	226 (29)
	Pbo	189	1592 (493)	1641 (519)	50 (247)	KI [170; 282] p<0,0001
Weiblich	AB	87	1203 (370)	1410 (403)	207 (272)	174 (42)
	Pbo	84	1293 (416)	1323 (451)	30 (257)	KI [91; 256] p<0,0001
COPD-Schweregrad						
Stufe II	AB	184	1690 (490)	1926 (541)	237 (275)	200 (27)
	Pbo	178	1680 (462)	1721 (497)	41 (272)	KI [148; 252] p<0,0001
Stufe III	AB	84	1119 (357)	1367 (429)	248 (266)	212 (38)
	Pbo	92	1138 (291)	1183 (304)	44 (199)	KI [137; 288] p<0,0001
Raucherstatus						
Raucher	AB	148	1500 (500)	1770 (538)	270 (295)	260 (32)
	Pbo	144	1467 (464)	1484 (488)	18 (269)	KI [197; 323] p<0,0001
Ex-Raucher	AB	121	1517 (557)	1735 (612)	218 (282)	151 (35)
	Pbo	129	1537 (516)	1609 (546)	73 (225)	KI [83; 219] p<0,0001
Reversibilität						
Ja	AB	81	1541 (422)	1885 (467)	344 (257)	285 (40)
	Pbo	81	1385 (452)	1463 (470)	78 (247)	KI [207; 362] p<0,0001
Nein	AB	187	1498 (562)	1693 (602)	195 (267)	169 (26)
	Pbo	189	1543 (492)	1570 (522)	27 (249)	KI [118; 220] p<0,0001

BMI						
Untergewicht bis normal ($<25 \text{ kg/m}^2$)	AB	97	1386 (479)	1621 (501)	235 (222)	210 (38)
	Pbo	113	1409 (485)	1448 (509)	39 (228)	KI [135; 284] p<0,0001
Übergewicht ($25\text{-}30 \text{ kg/m}^2$)	AB	99	1499 (524)	1777 (607)	278 (323)	252 (40)
	Pbo	92	1594 (488)	1617 (548)	23 (220)	KI [174; 330] p<0,0001
Adipös ($\geq 30 \text{ kg/m}^2$)	AB	73	1681 (544)	1900 (578)	219 (320)	147 (46)
	Pbo	68	1524 (479)	1602 (478)	79 (316)	KI [57; 238] p=0,0015
Verwendung von ICS						
Ja	AB	128	1447 (477)	1685 (500)	238 (295)	184 (33)
	Pbo	145	1458 (491)	1513 (512)	55 (254)	KI [119; 250] p<0,0001
Nein	AB	141	1562 (561)	1817 (624)	255 (286)	234 (34)
	Pbo	128	1547 (486)	1578 (527)	31 (246)	KI [168; 300] p<0,0001
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, AB = Aclidiniumbromid, Pbo = Placebo, ITT = Intention to Treat, N = Anzahl, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall						

Tabelle 4.3-232: Ergebnisse der a priori gelanten Subgruppenanalyse für SGRQ-Responder (24 Wochen) aus RCT für indirekte Vergleiche

Studie	Studienarm	Patienten (ITT)	Follow-up	Gruppenunterschied ¹
		N	N Responder (%)	OR (SE) KI [] p-wert
ATTAIN				
Alter				
< 60 Jahre	AB	96	57 (59,4%)	1,91 (0,568) KI [1,068; 3,434] p=0,0292
	Pbo	101	44 (44,0%)	
60 – 70 Jahre	AB	108	61 (56,5%)	2,11 (0,584) KI [1,229; 3,630] p=0,0068
	Pbo	121	45 (37,2%)	
> 70 Jahre	AB	65	28 (43,1%)	1,16 (0,458) KI [0,532; 2,513] p=0,7144
	Pbo	50	18 (36,0%)	
Geschlecht				
Männlich	AB	182	88 (48,4%)	1,24 (0,268) KI [0,816; 1,897] p=0,3097
	Pbo	188	79 (42,3%)	
Weiblich	AB	87	58 (66,7%)	3,93 (1,303) KI [2,048; 7,525] p<0,0001
	Pbo	84	28 (33,3%)	
COPD-Schweregrad				
Stufe II	AB	184	104 (56,5%)	1,92 (0,422) KI [1,246; 2,952] p=0,0031
	Pbo	177	70 (39,6%)	
Stufe III	AB	84	41 (48,8%)	1,50 (0,474) KI [0,809; 2,788] p=0,1971
	Pbo	91	35 (38,5%)	
Raucherstatus				
Raucher	AB	148	83 (56,1%)	1,83 (0,446) KI [1,138; 2,951] p=0,0128
	Pbo	143	59 (41,3%)	
Ex-Raucher	AB	121	63 (52,1%)	1,65 (0,436) KI [0,987; 1,775] p=0,0562
	Pbo	129	48 (37,5%)	
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, AB = Acridiniumbromid, Pbo = Placebo, ITT = Intention to Treat, N = Anzahl, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall				

Tabelle 4.3-233: Ergebnisse der a priori gelanten Subgruppenanalyse für TDI-Responder (24 Wochen) aus RCT für indirekte Vergleiche nach 24 Wochen

Studie	Studienarm (Patienten)	Patienten (ITT)	Follow-up	Gruppenunterschied ¹
		N	N responder (%)	OR (SE) KI [] p-wert
ATTAIN				
Alter				
< 60 Jahre	AB	94	54 (57,5%)	1,42 (0,424) KI [0,793; 2,55] p=0,2371
	Pbo	99	49 (50,0%)	
60 – 70 Jahre	AB	106	62 (58,5%)	2,08 (0,58) KI [1,207; 3,595] p=0,0084
	Pbo	114	47 (41,6%)	
> 70 Jahre	AB	62	33 (53,2%)	1,56 (0,62) KI [0,712; 3,397] p=0,2676
	Pbo	46	21 (45,7%)	
Geschlecht				
Männlich	AB	177	107 (60,5%)	1,65 (0,358) KI [1,077; 2,522] p=0,0212
	Pbo	180	88 (49,4%)	
Weiblich	AB	85	42 (49,4%)	1,78 (0,58) KI [0,942; 3,375] p=0,0755
	Pbo	79	29 (36,7%)	
COPD-Schweregrad				
Stufe II	AB	179	109 (60,9%)	1,93 (0,427) KI [1,249; 2,975] p=0,003
	Pbo	168	76 (45,2%)	
Stufe III	AB	82	40 (48,8%)	1,34 (0,425) KI [0,721; 2,498] p=0,3542
	Pbo	88	39 (45,4%)	
Raucherstatus				
Raucher	AB	145	83 (57,2%)	1,86 (0,458) KI [1,147; 3,014] p=0,0119
	Pbo	135	59 (43,7%)	
Ex-Raucher	AB	117	66 (56,4%)	1,49 (0,393) KI [0,889; 2,499] p=0,1298
	Pbo	124	58 (47,5%)	
¹ Analyse basiert auf einem adjustierten ANCOVA-Modell, AB = Acridiniumbromid, Pbo = Placebo, ITT = Intention to Treat, N = Anzahl, SD = Standardabweichung, SE = Standardfehler, KI = Konfidenzintervall				

Im Folgenden werden die Ergebnisse der a priori Subgruppenanalysen – nach potentiellen Effektmodifikatoren geliedert – zusammengefasst.

ATTAIN

Alter:

Bezüglich der primären und sekundären Endpunkte war der Umfang des Behandlungseffektes bei Acridiniumbromid und Placebo in allen Altersgruppen vergleichbar (<60, 60 Jahre bis 69 Jahre und ≥ 70 Jahre). Es wurden keine Anzeichen für eine geringere Wirksamkeit von Acridiniumbromid mit zunehmendem Alter beobachtet.

Geschlecht:

Bezüglich der bronchodilatatorischen Endpunkte (Veränderung des morgendlichen FEV1-Trough und des FEV1-Peak in Woche 24 gegenüber dem Ausgangswert) und der Anzahl (%) an Patienten, die eine klinisch relevante Verbesserung des TDI von mindestens 1 Einheit bis Woche 24 erreicht haben, war der Umfang des Behandlungseffektes zwischen Acridiniumbromid und Placebo bei Männern und Frauen vergleichbar. Bezüglich der Anzahl (%) an Patienten, die eine klinisch relevante Verbesserung des SGRQ von mindestens 4 Einheiten bis Woche 24 erreicht haben, war der beobachtete Behandlungseffekt von Acridiniumbromid bei einer Dosis von 400 µg bei Frauen größer als bei Männern: Die Odds Ratios (KI 95 %) des Vergleichs zwischen Acridiniumbromid und Placebo betrug bei den weiblichen Patienten 3,93 (KI [2,05; 7,53]) und bei den männlichen Patienten 1,24 (KI [0,82; 1,90]).

COPD-Schweregrad:

Bezüglich der primären und sekundären Endpunkte war der Umfang des Behandlungseffekt bei Acridiniumbromid und Placebo unter den verschiedenen COPD-Schweregraden vergleichbar (Leicht/Mittelschwer, Schwer/Sehr schwer).

Raucherstatus:

Bezüglich der Veränderung des morgendlichen FEV1-Wertes (Trough) und der sekundären Endpunkte (SGRQ und TDI), war der Umfang des Behandlungseffektes zwischen Acridiniumbromid und Placebo unter den Patientengruppen mit unterschiedlichem Raucherstatus (aktiver Raucher, Ex-Raucher) vergleichbar. Bezüglich des sekundären Endpunktes über die Veränderung des morgendlichen FEV1-Peak in Woche 24 gegenüber dem Ausgangswert, war der Umfang des Behandlungseffektes bei Acridiniumbromid in der Subgruppe der aktiven Raucher (-0,012 l bzw. 0,247 l) höher als in der Subgruppe mit den Ex-Rauchern (-0,060 l bzw. 0,21). Dieser größere Behandlungseffekt auf die FEV1-Peaks in der Subgruppe der aktiven Raucher ist hauptsächlich auf den niedrigen FEV1-Peak in der Placebo-Gruppe der aktiven Raucher im Vergleich zu den Ex-Rauchern zurückzuführen.

Reversibilität²⁶:

Bezüglich der Veränderung zwischen dem morgendlichen FEV1-Wert (Trough) gegenüber dem Ausgangswert, war der Umfang des Behandlungseffektes bei Acridiniumbromid und Placebo unter den beiden Untergruppen basierend auf ihrer Reversibilität (Ja, Nein) vergleichbar. Bezüglich des sekundären Endpunktes über die Veränderung des morgendlichen FEV1-Peak in Woche 24 gegenüber dem Ausgangswert, war der beobachtete Behandlungseffekt zwischen Acridiniumbromid-Patienten mit Reversibilität (0,285) größer als in der Subgruppe mit Acridiniumbromid-Patienten ohne Reversibilität (0,169 l).

Gebrauch von inhalierbaren Corticosteroiden:

Bezüglich der primären und sekundären Endpunkte war der Umfang des Behandlungseffektes basierend auf der Bronchodilatation zwischen Acridiniumbromid und Placebo unter den verschiedenen Subgruppen basierend auf deren Gebrauch von inhalierbaren Corticosteroiden (Ja, Nein) vergleichbar.

BMI:

Bezüglich der primären und sekundären Endpunkte war der Umfang des Behandlungseffektes basierend auf der Bronchodilatation zwischen Acridiniumbromid und Placebo unter den verschiedenen BMI-Untergruppen (Untergewicht bis Normalgewicht, Übergewicht, Adipositas) vergleichbar.

ACCORD II

Die Ergebnisse deuten darauf hin, dass die Verzerrungen bei der Randomisierung eines höheren Prozentsatzes von Patienten mit einem höheren COPD-Schweregrad in den Acridiniumbromid-Behandlungsgruppen als in der Placebo-Gruppe zu einer geringeren Wirksamkeit beim primären Endpunkt (FEV1-Trough in Woche 12) der Acridiniumbromid-Behandlungsgruppen im Vergleich zur Placebo-Gruppe führte. Diesem Umstand wird mit Sensitivitätsanalysen Rechnung getragen.

Zusammenfassend sind aufgrund der Ergebnisse der a priori Subgruppenanalysen ist keine Modifikation der Nutzen- und Zusatznutzenbewertung notwendig.

²⁶ Reversibel ist wie folgt definiert: Reversibilität mit Bronchodilatoren $\geq 12\%$ und Veränderung des präbronchodilatatorischen FEV1-Wertes $\geq 0,200$ L

4.3.2.1.3.18 Zusammenfassung der Ergebnisse aus indirekten Vergleichen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.1.3.18 und 4.3.2.4) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Im Kapitel 4.3.2.1.3 wurden Meta-Analysen und Netzwerk-Meta-Analysen (NMA) für Acclidiniumbromid 400µg BID und die zweckmäßige Vergleichstherapie Tiotropiumbromid 18µg und 5µg OD durchgeführt. Alle Studien verglichen die jeweilige Intervention mit Placebo. Drei Studien untersuchten Acclidiniumbromid: ACCORD I, ACCORD II und ATTAIN. Vierzehn Studien verglichen Tiotropiumbromid 18µg mit Placebo: Beeh 2006, Brusasco 2003, Casaburi 2000, Casaburi 2002, SAFE, Covelli 2005, INHANCE, Donohue 2002, SPRUCE, Johansson 2008, SAFE Portugal, Niewoehner 2005, UPLIFT und Verkindre 2006. Drei Studien untersuchten die Wirkung von Tiotropiumbromid 5µg: Bateman 2010b, Bateman 2010a und Voshaar 2008. Die in der folgenden **Fehler! Ungültiger Eigenverweis auf Textmarke.** dargestellten Endpunkte wurden in Woche 12 und/ oder 24 analysiert:

Tabelle 4.3-234: Im direkten Vergleich herangezogene Endpunkte

Mortalität
COPD-assoziierte kardiovaskuläre Morbidität und Mortalität
COPD-bedingte Letalität
Gesamtmortalität
Morbidität (COPD-Symptome)
Husten, Auswurf, Atemnot (EXACT -Fragebogen)
Nächtliche und morgendliche Symptome
Dyspnoe (TDI Focal Score und TDI Responder)
Exazerbationen (Exazerbationsrate und Hospitalisierungen aufgrund von Exazerbationen)
Trough-FEV ₁ -Wert
Peak-FEV ₁ -Wert
Gesundheitsbezogene Lebensqualität
Lebensqualität (SGRQ-Score und SGRQ Responder)
Unerwünschte Arzneimittelwirkungen
Unerwünschte Ereignisse (AEs)
Schwerwiegende unerwünschte Ereignisse (SAEs)
Studienabbruch aufgrund unerwünschter Ereignisse

Angaben zu durchgeführten Sensitivitätsanalysen

Zur Untersuchung des Einflusses der ACCORD-II-Studie auf die Meta-Analyse der Aclidiniumbromid-Studien wurde eine Sensitivitätsanalyse unter Ausschluss der ACCORD-II-Studie durchgeführt. In dieser Studie waren die Baseline-Charakteristika der Patienten nur eingeschränkt vergleichbar.

Zwei Studien mit Tiotropiumbromid 18µg (Freeman 2007 (SPRUCE) Johansson 2008) schlossen im Gegensatz zu allen anderen Studien nur sehr wenig mit Patienten mit schwerer COPD ein. Sofern zutreffend, wurden Sensitivitätsanalysen unter Ausschluss dieser Studien durchgeführt (Endpunkte: FEV₁-Trough, unerwünschte Ereignisse, schwere unerwünschte Ereignisse, Studienabbrüche aufgrund eines unerwünschten Ereignisses). Bestand die Heterogenität der Studien weiterhin, wurden Studien mit einem hohen Verzerrungspotenzial von der Analyse ausgeschlossen. Konnte die Heterogenität auch auf diese Weise nicht erklärt werden, wurden mit den Studien, die verschiedene Tiotropiumbromid-Dosen untersuchten (d. h. 5µg und 18µg), separate Sensitivitätsanalysen durchgeführt.

In der NMA wurden primär Studien ohne begleitende Verwendung eines LABA berücksichtigt. In Szenario-Analysen wurden Studien, die die begleitende Verwendung eines LABA gestatteten, in das bestehende Netzwerk aufgenommen und ein zweites Netzwerk („alle Studien“) erstellt.

Die Ergebnisse der Meta-Analyse und der NMA pro Endpunkt werden im Folgenden zusammengefasst.

Trough-FEV₁

In Woche 12 zeigten die Meta-Analysen, dass Aclidiniumbromid im Vergleich zu Placebo wirksamer ist (mittlere Differenz: 104,79 ml [KI 95 % 80,2; 129,4]. Tiotropiumbromid 5µg/18µg ist im Vergleich zu Placebo ebenfalls wirksamer (mittlere Differenz: 112,9 ml [KI 95% 94,8; 131,0]). Es bestand eine signifikante Heterogenität der Tiotropium-Studien, deren Ursachen mit Hilfe von Sensitivitätsanalysen weiter analysiert wurden.

Aclidiniumbromid: Die Sensitivitätsanalyse ohne die ACCORD-II-Studie aus der Meta-Analyse ergab keine wesentlichen unterschiedlichen gepoolten Effektschätzer. Dies war auch bei allen anderen Endpunkten der Fall, für die diese Sensitivitätsanalyse durchgeführt wurde.

Tiotropiumbromid: Die Sensitivitätsanalyse ohne die Studien Johansson 2008 und Freeman 2007 (SPRUCE) ergab ähnliche Ergebnisse wie die Analyse des gesamten Tiotropium-Studienpools (mittlere Differenz 117,9 ml [KI 95 % 99,5; 136,3]). Schließt man die Tiotropiumbromid-Studien mit hohem Verzerrungsrisiko aus, besteht keine Heterogenität mehr ($p < 0,31$) und die gepoolte mittlere Differenz beträgt 104,2 ml (95 % CI: 90,1; 117,6).

Nach 24 Wochen waren Tiotropiumbromid 18µg und Tiotropiumbromid 5µg ebenfalls signifikant wirksamer als Placebo (mittlere Differenz: 110,8 ml [KI 95 % 98,0; 123,7].

Aufgrund der signifikanten Studien-Heterogenität ($p = 0,06$), wurde die Sensitivitätsanalyse unter Ausschluss der Studien mit hohem Verzerrungspotenzial durchgeführt. Der Ausschluss dieser Studien hatte kaum Auswirkungen auf den gepoolten Effektschätzer, die Studien-Heterogenität bestand jedoch nicht mehr ($p=0,30$).

Nur die ATTAIN-Studie berichtet über Ergebnisse für Aclidiniumbromid nach 24 Wochen, weshalb für Aclidiniumbromid keine Meta-Analyse durchgeführt werden konnte. Da ATTAIN die einzige Studie zu Aclidiniumbromid ist, die über Ergebnisse nach 24 Wochen berichtet, gilt dies für alle Endpunkte.

Die NMA zeigte auf der Basis von Studien ohne begleitende LABA-Therapie, dass Aclidiniumbromid und Tiotropiumbromid $5\mu\text{g}/18\mu\text{g}$ den Trough-FEV₁ (CFB) nach 12 Wochen (mittlere Differenz: -20 ml [95 %-KI: -48 ; 7]) und 24 Wochen (mittlere Differenz: 23 ml [95 %-KI: -21 ; 68]) in gleicher Weise verbessern (die 95 %-Konfidenzintervalle schließen die Null mit ein).

Eine Analyse aller Nachweise (unter Einschluss der Studien mit und ohne begleitende LABA-Therapie) zeigte ein vergleichbares Ergebnis des indirekten Vergleichs zwischen Aclidiniumbromid und Tiotropiumbromid (die 95 %-Konfidenzintervalle schließen die Null mit ein). Zu beiden Zeitpunkten (12 und 24 Wochen) waren die Ergebnisse der Meta-Analyse und der NMA konsistent. Der Nutzen von Aclidiniumbromid und Tiotropiumbromid in Bezug auf den Endpunkt Trough-FEV₁ ist deshalb gleichwertig.

Peak-FEV₁

Nach 12 Wochen zeigte die Meta-Analyse, dass Aclidiniumbromid und Tiotropiumbromid 5 oder $18\mu\text{g}$ jeweils den Peak-FEV₁ im Vergleich zu Placebo verbesserten (mittlere Differenz: $172,9$ ml [KI 95 % $134,0$; $211,8$] bzw. $192,5$ ml [KI 95 % $132,8$; $252,3$]). In beiden Analysen bestand eine signifikante Studien-Heterogenität. Die Sensitivitätsanalyse unter Ausschluss der ACCORD-II-Studie zeigte ähnliche Ergebnisse und erklärte die Heterogenität. Alle Tiotropiumbromid-Studien berichteten über Patienten mit mittelschwerer bis schwerer COPD und hatten ein geringes Verzerrungspotenzial. Die Durchführung einer Sensitivitätsanalyse auf der Basis der Dosis erklärte die Studien-Heterogenität nicht. Der Gesamteffektschätzer der Meta-Analyse wird trotz vorliegender Heterogenität berücksichtigt, da Tiotropiumbromid in allen Studien den Peak-FEV₁ signifikant verbesserte (gleiche Effektrichtung).

Nach 24 Wochen verbesserte Tiotropiumbromid $18\mu\text{g}$ den Peak-FEV₁ im Vergleich zu Placebo mit einem mittleren Unterschied von $218,8$ ml [KI 95 % $157,5$; $280,1$]. Allerdings lag eine signifikante Studien-Heterogenität vor ($p < 0,15$). Weder das Verzerrungspotenzial noch die Inhalationsdosis erklären die Heterogenität. Die Effektrichtung stimmt jedoch in allen Studien überein.

Nach 12 Wochen zeigten beide Szenarios in der NMA (Analyse der Studien mit begleitender LABA-Therapie sowie die Analyse aller Studien), dass Aclidiniumbromid in Bezug auf die Verbesserung des Peak-FEV₁ mit Tiotropiumbromid vergleichbar ist (das 95 %-Konfidenzintervall schließt die Null mit ein; Unterschied 19 ml [95% KI: -18 , 57]).

Nach 24 Wochen zeigte die NMA mit Acclidiniumbromid und Tiotropiumbromid 18µg eine vergleichbare Verbesserung des Peak-FEV₁ (das 95 %-Konfidenzintervall schließt die Null mit ein; Differenz: -18ml [95 % KI: -81, 42]). Keine der Studien, die über den Peak-FEV₁ nach 24 Wochen berichteten, ließ eine begleitende LABA-Behandlung zu. Die Ergebnisse der Meta-Analyse stimmen für diesen Endpunkt zu beiden Zeitpunkten mit denen der NMA überein. Aus den Ergebnissen lässt sich schlussfolgern, dass der Nutzen von Acclidiniumbromid und Tiotropiumbromid bezüglich des Endpunktes CFB Peak-FEV₁ gleichwertig ist.

TDI Focal Score

Nach 12 Wochen verbesserte die Behandlung mit Acclidiniumbromid die Dyspnoe im Vergleich zu Placebo (mittlere Differenz des TDI Focal Scores: 0,96 [KI 95 % 0,64; 1,27]. Der Ausschluss der ACCORD-II-Studie aus der Meta-Analyse hatte keine wesentlichen Auswirkungen auf den gepoolten Effektschätzer. Ein ähnliches Ergebnis ergab der Vergleich zwischen Tiotropiumbromid 18µg und Placebo (mittlere TDI-Differenz: 0,82 [KI 95% 0,59; 1,05]). In beide Meta-Analysen lag für den Endpunkt TDI Focal Score Studien-Homogenität vor.

Nach 24 Wochen verbesserte Tiotropiumbromid 18µg den TDI Focal Score im Vergleich zu Placebo ebenfalls (mittlere Differenz: 0,91 [KI 95 % 0,67; 1,14]). Heterogenität zwischen den Studien bestand nicht.

Für dieses Ergebnis waren nur Studien verfügbar, die eine begleitende LABA-Therapie nicht gestatteten. Nach 12 (OR 1,2 (95 %-KI: 0,8, 1,6) und 24 Wochen (OR 1,0 (95%-KI: 0,7, 1,5) zeigte die NMA, dass eine Behandlung mit Acclidiniumbromid oder Tiotropiumbromid 18µg zu vergleichbaren Verbesserungen des TDI Focal Scores führt. Die 95 %-Konfidenzintervalle schließen zu beiden Zeitpunkten die Null mit ein. Die Ergebnisse der NMA und der Meta-Analyse waren zu beiden Zeitpunkten konsistent. Der Nutzen von Acclidiniumbromid und Tiotropiumbromid in Bezug auf den Endpunkt TDI Focal Score kann deshalb als gleichwertig angesehen werden.

TDI-Responder

Nach 12 Wochen war keine Meta-Analyse von Tiotropiumbromid 18µg möglich (nur eine Tiotropium-Studie berichtete diesen Endpunkt). Die Behandlung mit Acclidiniumbromid führte zu mehr TDI-Respondern als die Behandlung mit Placebo (OR = 1,93 [KI 95 % 1,54; 2,41]). Der Ausschluss der ACCORD-II-Studie aus der Meta-Analyse ergab beinahe den gleichen gepoolten Effektschätzer.

Nach 24 Wochen war der Anteil der TDI-Responder in der Tiotropiumbromid 18µg-Gruppe höher als in der Placebogruppe (OR 1,70 [95 %-KI 1,40; 2,06]). Es bestand keine Studien-Heterogenität.

Wie beim TDI Focal Score, waren auch für TDI-Responder nur Studien verfügbar, die eine begleitende LABA-Therapie nicht zuließen. Nach 12 und 24 Wochen sind der Prozentsatz der TDI-Responder mit Acridiniumbromid und mit Tiotropiumbromid 18µg ähnlich (die 95 %-Konfidenzintervalle schließen zu beiden Zeitpunkten die Eins mit ein). Die Ergebnisse der Meta-Analyse stimmen mit denen der NMA für diesen Endpunkt zu beiden Zeitpunkten überein. Der Nutzen von Acridiniumbromid und Tiotropiumbromid in Bezug auf den Endpunkt TDI Responder kann deshalb als gleichwertig angesehen werden.

EXACT-Respirationssymptome, nächtliche und morgendliche COPD-Symptome

Für EXACT-Respirationssymptome wurden keine Meta-Analyse und keine NMA durchgeführt, da nur eine Studie (ATTAIN) verfügbar war. Diese Studie zeigte, dass eine Behandlung mit Acridiniumbromid mit deutlich weniger COPD-Symptomen assoziiert ist als eine Placebo-Behandlung.

Für nächtliche und morgendliche COPD-Symptome wurde keine NMA durchgeführt, da nur 2 Studien verfügbar waren, die Acridiniumbromid mit Placebo zu zwei unterschiedlichen Erhebungszeitpunkten vergleichen (ACCORD I und ATTAIN). Die Studien zeigten, dass eine Behandlung mit Acridiniumbromid mit deutlich weniger nächtlichen und morgendlichen Symptomen assoziiert war als eine Placebo-Behandlung.

Exazerbationsrate

Die Ergebnisse der Meta-Analyse zeigen, dass beide aktive Behandlungen im Vergleich zu Placebo die Exazerbationsrate in vergleichbarer Weise reduzieren (Acridiniumbromid: Hazard Ratio: 0,69 [95 %-KI 0,53; 0,91]; Tiotropiumbromid: Hazard Ratio 0,68 [95 %-KI 0,59; 0,78]). Die Sensitivitätsanalyse unter Ausschluss der ACCORD-II-Studie zeigte ähnliche Ergebnisse (Hazard Ratio 0,64 [95 %-KI 0,47; 0,87]).

Die Meta-Analyse von Tiotropiumbromid und Placebo ergaben eine signifikante Heterogenität zwischen den Studien. Sensitivitätsanalysen zeigten, dass die Heterogenität nicht mit dem Verzerrungspotenzial oder der verwendeten Dosis zu erklären sind. Es ist möglich, dass die Heterogenität durch unterschiedliche Definitionen dieses Endpunktes in den berücksichtigten Studien verursacht wird.

Auf der Basis von Studien ohne begleitende Verwendung von LABAs ist Acridiniumbromid in Bezug auf Exazerbationen genauso wirksam wie Tiotropiumbromid (die Eins liegt innerhalb des 95 %-Konfidenzintervalls; Odds Ratio 1,2 [95 % KI: 0,7, 1,8]). Unter Berücksichtigung aller Studien wurden gleichwertige Ergebnisse erreicht.

Die Ergebnisse der Meta-Analyse und der NMA waren zu beiden Zeitpunkten konsistent. Acridiniumbromid und Tiotropiumbromid haben einen gleichwertigen Nutzen in Bezug auf den Endpunkt Exazerbationsrate.

Krankenhausaufenthalte aufgrund von Exazerbationen

Acclidiniumbromid führte zu einer vergleichbaren Anzahl stationärer Aufenthalte aufgrund von Exazerbationen wie Placebo (OR 0,78 [KI 95 % 0,14; 4,33]). Jedoch war die Zahl der Ereignisse in beiden Acclidiniumbromid-Studien sehr niedrig, was die statistische Aussagekraft begrenzte. Nach 12 Wochen war keine Meta-Analyse von Studien zu Tiotropiumbromid möglich.

Nach 24 Wochen reduzierte Tiotropiumbromid 18µg die Anzahl der aufgrund von Exazerbationen stationär behandelten Patienten im Vergleich zu Placebo (OR 0,72 [KI 95 % 0,54; 0,98]). Es bestand keine Studien-Heterogenität.

Die Durchführung einer NMA nach 12 Wochen war nicht möglich, da nach 12 Wochen zu diesem Endpunkt keine Daten zu Tiotropiumbromid 5µg und Tiotropiumbromid 18µg verfügbar waren.

Eine Analyse nach 24 Wochen der Studien ohne begleitende LABA-Therapie zeigte, dass Acclidiniumbromid und Tiotropiumbromid zu einer vergleichbaren Anzahl stationärer Einweisungen aufgrund von Exazerbationen führten, wobei sich eine Tendenz zugunsten von Acclidiniumbromid abzeichnete (OR 0,31 (95 % KI: 0,03, 1,1).

Bei Berücksichtigung aller Studien ließ sich eine Tendenz zu einer Überlegenheit von Acclidiniumbromid im Vergleich zu Tiotropiumbromid 18µg nach 24 Wochen beobachten (die Eins liegt nur knapp innerhalb des 95 %-Konfidenzintervalls; OR 0,31 [95 % KI: 0,03, 1,0]).

Die Ergebnisse der Meta-Analyse waren zu beiden Zeitpunkten mit der NMA konsistent. Der Nutzen von Acclidiniumbromid im Vergleich zu Tiotropiumbromid in Bezug auf den Endpunkt Krankenhausaufenthalte aufgrund von Exazerbationen kann deshalb als zumindest gleichwertig (mit Trend zur Überlegenheit) angesehen werden.

SGRQ-Score

Nach 12 Wochen verbesserte Acclidiniumbromid die Lebensqualität von COPD-Patienten im Vergleich zu Placebo signifikant (mittlere Differenz des SGRQ-Scores: -2,94 [KI 95% -4,52; -1,37]). Es gab keine Studien-Heterogenität und der Ausschluss der ACCORD II-Studie hatte keine signifikanten Auswirkungen auf das Ergebnis.

Tiotropiumbromid 18µg verbesserte die Lebensqualität nach 12 Wochen im Vergleich zu Placebo ebenfalls (mittlere Differenz: -2,82 [KI 95 % -5,36; -0,28]). Es bestand zwar eine signifikante Studien Heterogenität, die Effektrichtung war jedoch in allen drei Studien identisch.

Nach 24 Wochen verbesserte Tiotropiumbromid 5µg/18µg die Lebensqualität ebenfalls im Vergleich zu Placebo (mittlerer Unterschied: -2,51 [KI 95 % -3,01; -2,01]). Es bestand keine Studien-Heterogenität.

Die NMA zeigte auf der Basis von Studien ohne begleitende LABA-Therapie, dass die Verbesserungen des SGRQ-Gesamtwerts mit Acclidiniumbromid und Tiotropiumbromid nach 12 und 24 Wochen gleichwertig sind [Differenz 0,2 Punkte (95 %-KI: -3,9, 4,9)].

Auch bei Berücksichtigung aller Studien führen Acclidiniumbromid und Tiotropiumbromid zu vergleichbaren Verbesserungen des SGRQ-Gesamtwerts nach 24 Wochen [mittlere Differenz: -1,9 Punkte (95 %-KI: -4,2, 0,4)]. Die Ergebnisse der Meta-Analyse stimmten zu beiden Zeitpunkten mit der NMA überein. Der Nutzen von Acclidiniumbromid und Tiotropiumbromid auf die Lebensqualität ist demnach gleichwertig.

Prozentsatz der SGRQ-Responder

Nach 12 Wochen war die SGRQ-Responderrate mit Acclidiniumbromid höher als mit Placebo (OR = 1,63 [KI 95 % 1,27; 2,09]). Die Meta-Analyse der Tiotropiumbromid 18µg-Studie ergab ein ähnliches Ergebnis (OR = 1,78 [95 % CI: 1,19; 2,66]). Es wurde keine Studien-Heterogenität gefunden. Der Ausschluss der ACCORD-II-Studie aus der Meta-Analyse hatte keinen wesentlichen Einfluss auf die gepoolten Effektschätzer.

Nach 24 Wochen war die SGRQ-Responderrate mit Tiotropiumbromid ebenfalls höher als im Placeboarm (OR = 1,42 [95 % CI: 1,27; 1,59]).

Bei exklusiver Berücksichtigung der Studien ohne begleitende LABA-Therapie zeigte die NMA, dass die SGRQ-Responderraten mit Acclidiniumbromid und Tiotropiumbromid 5µg/18µg nach 12 gleichwertig sind [12 Wochen: OR 0,96 (95 %-KI: 0,6, 1,4)] Nach 24 Wochen bestand eine Tendenz zugunsten einer Überlegenheit von Acclidiniumbromid [OR 1,4 (95 %-KI: 0,9, 2,1)].

Bei Verwendung aller Studien führte Acclidiniumbromid nach 24 Wochen zu einer entsprechenden SGRQ-Responderrate wie Tiotropiumbromid 5µg/18µg. Die Ergebnisse der Meta-Analyse stimmten nach 12 und 24 Wochen mit denen der NMA überein. Der Nutzen von Acclidiniumbromid im Vergleich zur zweckmäßigen Vergleichstherapie Tiotropiumbromid auf den Endpunkt SGRQ-Responderrate ist also mindestens gleichwertig.

Unerwünschte Ereignisse

Nach 12 Wochen waren Acclidiniumbromid und Placebo in der Meta-Analyse mit einer ähnlichen Anzahl unerwünschter Ereignisse assoziiert (OR 0,88 [0,63; 1,25]). Mit Tiotropiumbromid wurden nach 12 Wochen weniger unerwünschte Ereignisse als mit Placebo beobachtet (OR 0,82 [KI 95 % 0,70; 0,96]). Es bestand keine Studien-Heterogenität. Ein

vergleichbares Ergebnis ergab die Sensitivitätsanalyse unter Ausschluss der Tiotropiumbromid-Studien mit wenig Patienten mit schwerer COPD (OR 0,82 [95 %-KI 0,68; 0,98]).

Nach 24 Wochen unterschied sich die Anzahl unerwünschter Ereignisse mit Tiotropiumbromid 18µg und Placebo nicht (OR 1,04 [KI 95 % 0,77; 1,39]). Meta-Analysen zu beiden Zeitpunkten ergaben keine Studien-Heterogenität

Sowohl auf der Grundlage von Studien ohne begleitende LABA-Therapie als auch bei Berücksichtigung aller Studien ist die Anzahl der unerwünschten Ereignisse mit Acclidiniumbromid und Tiotropiumbromid innerhalb von 12 und 24 Wochen gleichwertig. Allerdings berichteten keine der Studien, die eine begleitende LABA-Therapie zuließen, nach 24 Wochen über diesen Endpunkt. Die Ergebnisse der Meta-Analyse stimmten zu beiden Zeitpunkten mit denen der NMA überein. Der Nutzen von Acclidiniumbromid und der zweckmäßigen Begleittherapie Tiotropiumbromid im Hinblick auf den Endpunkt unerwünschte Ereignisse kann deshalb als gleichwertig angesehen werden.

Schwere unerwünschte Ereignisse

Nach 12 Wochen war die Zahl schwerer unerwünschter Ereignisse mit Acclidiniumbromid und Placebo (OR 0,88 [KI 95 % 0,42; 1,85]) sowie mit Tiotropiumbromid 5µg/18µg und Placebo vergleichbar (OR 1,19 [KI 95 % 0,69; 2,06]). Die Sensitivitätsanalyse unter Ausschluss der Studien mit wenigen schweren COPD-Patienten (Freeman (SPRUCE) und Johansson 2008) ergab ein ähnliches Ergebnis (OR 1,11 [KI 95 % 0,63; 1,96]).

Nach 24 Wochen war Tiotropiumbromid 18µg in Bezug auf schwere unerwünschte Ereignisse mit Placebo vergleichbar (OR = 1,03 [KI 95 % 0,83; 1,28]).

Bei Berücksichtigung der Studien, die eine begleitende LABA-Therapie nicht zuließen, war Acclidiniumbromid nach 12 Wochen mit Tiotropiumbromid gleichwertig [OR 1,0 (95 %KI: 0,3, 2,4)]. Das Szenario mit allen Studien lieferte ähnliche Ergebnisse. Nach 24 Wochen war Acclidiniumbromid in Szenarios mit und ohne Studien, die eine begleitende Behandlung mit LABAs zuließen, ebenfalls mit Tiotropiumbromid gleichwertig. Die Ergebnisse der Meta-Analyse stimmten zu beiden Zeitpunkten mit der NMA überein. Der Nutzen von Acclidiniumbromid und der zweckmäßigen Vergleichstherapie Tiotropiumbromid im Hinblick auf den Endpunkt schwere unerwünschte Ereignisse kann als deshalb gleichwertig angesehen werden.

Studienabbrüche aufgrund unerwünschter Ereignisse

Nach 12 Wochen führten Acclidiniumbromid und Placebo zu einer vergleichbaren Anzahl von Studienabbrüchen aufgrund unerwünschter Ereignisse (OR 0,96 [KI 95 % 0,31; 3,00]).

Tiotropiumbromid führte im Vergleich zu Placebo zu einer geringeren Anzahl von Studienabbrüchen aufgrund unerwünschter Ereignisse (OR 0,59 [KI 95 % 0,37; 0,94]).

Nach 24 Wochen wurde mit Tiotropiumbromid 18µg und Placebo eine vergleichbare Zahl von Studienabbrüchen aufgrund unerwünschter Ereignisse beobachtet (OR 0,49 [KI 95 % 0,19; 1,27]). Allerdings bestand eine signifikante Heterogenität und auch die Effektrichtung war in den drei analysierten Studien nicht einheitlich.

Nach 12 Wochen war die Zahl der Studienabbrüche aufgrund unerwünschter Ereignisse mit Acridiniumbromid und Tiotropiumbromid in Szenarios mit und ohne begleitende LABA-Therapie gleichwertig. Jedoch war die Anzahl der Studienabbrüche aufgrund unerwünschter Ereignisse in allen Studien sehr niedrig. Bei Berücksichtigung von Studien, die die Verwendung von LABAs nicht zuließen, war Acridiniumbromid nach 24 Wochen mit Placebo und Tiotropiumbromid ebenfalls gleichwertig. Nach 24 Wochen waren keine Studien verfügbar, die die Verwendung von LABAs gestatteten. Die Ergebnisse der Meta-Analyse stimmten zu beiden Zeitpunkten mit der NMA überein. Der Nutzen von Acridiniumbromid und der zweckmäßigen Vergleichstherapie Tiotropiumbromid im Hinblick auf den Endpunkt Studienabbrüche aufgrund unerwünschter Ereignisse ist deshalb gleichwertig.

Kardiovaskuläre Morbidität und Mortalität

Nach 12 Wochen zeigte Acridiniumbromid für diesen Endpunkt im Vergleich z Placebo ähnliche Ergebnisse (OR 1,26 [KI 95 % 0,33; 4,80]).

Es wurde keine NMA durchgeführt, da die Ergebnisse der Meta-Analyse als Folge der unzureichenden statistischen Aussagekraft bereits große Konfidenzintervalle aufwiesen.

Letalität und Gesamtmortalität

Die berichtete Anzahl von Ereignissen war in allen Behandlungen und Studien sehr niedrig, sodass die statistische Aussagekraft einer Meta-Analyse oder NMA für ein aussagekräftiges Ergebnis unzureichend sein wird. Daher wurde keine Meta-Analyse oder NMA durchgeführt.

Zusammenfassung der Ergebnisse

Auf der Grundlage der Ergebnisse der Meta-Analysen und den Daten aus der ATTAIN-Studie zum Vergleich von Acridiniumbromid und Placebo nach 24 Wochen, wird geschlussfolgert, dass sowohl Acridiniumbromid als auch Tiotropiumbromid 5µg bzw. 18µg im Vergleich zu Placebo nach 12 und 24 Wochen die Lungenfunktion verbessern und die Lebensqualität von COPD-Patienten erhöhen. Beide Behandlungen reduzieren außerdem die Exazerbationsrate im Vergleich zu Placebo. Die Ergebnisse der Meta-Analysen und der ATTAIN-Studie nach 24 Wochen zeigen außerdem ein ähnliches Sicherheitsprofil von Acridiniumbromid und

Placebo, sowie von Tiotropiumbromid und Placebo nach 12 und 24 Wochen. Zwischen den Ergebnissen nach 12 und 24 Wochen wurden keine erkennbaren Unterschiede festgestellt. Die NMA zeigte, dass die Wirksamkeit und Sicherheit von Aclidiniumbromid und Tiotropiumbromid 5µg/18µg gleichwertig sind. Aclidiniumbromid und Tiotropiumbromid sind über einen Zeitraum von mindestens 24 Wochen wirksamer als und genau so sicher wie Placebo. Beim Endpunkt Krankenhausaufenthalte aufgrund von Exazerbationen gab es nach 24 Wochen eine Tendenz zu einer Überlegenheit von Aclidiniumbromid im Vergleich zu Tiotropiumbromid (die Eins lag nur knapp innerhalb des 95 %-Konfidenzintervalls). Die Ergebnisse der Meta-Analyse stimmten bei allen Endpunkten und zu beiden Zeitpunkten mit der NMA überein. Der Nutzen von Aclidiniumbromid und der zweckmäßigen Vergleichstherapie Tiotropiumbromid kann deshalb als gleichwertig angesehen werden.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

Nicht randomisierte vergleichende Studien werden nicht vorgelegt. Dieser Abschnitt bleibt leer.

4.3.2.3 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

<< Angaben des pharmazeutischen Unternehmers >>

4.3.2.3.1 Charakteristika der nicht randomisierten vergleichenden Studien

Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-235: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Hinweise auf ergebnisgesteuerte Berichterstattung	Andere Aspekte, die Verzerrungen verursachen können
			Patient	Behandler		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

4.3.2.3.2 Ergebnisse aus nicht randomisierten vergleichenden Studien

4.3.2.3.2.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-236: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-G. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-G.

Tabelle 4.3-237: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Ver- zerrungspotenzial be- einflussende Punkte
<Studie 1>	ja / nein / unklar	ja / nein / unklar	ja / nein / unklar	ja / nein

1) Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar.

<< Angaben des pharmazeutischen Unternehmers >>

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.4.

4.3.2.4 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.1.3.18 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.4.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.1.3.18 aufgeführt sind. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung. Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

<< Angaben des pharmazeutischen Unternehmers >>

4.3.2.4.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungspotenzial.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2. und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

<< Angaben des pharmazeutischen Unternehmers >>

4.3.2.4.3 Ergebnisse aus weiteren Untersuchungen

4.3.2.4.3.1 <Endpunkt xxx> – weitere Untersuchungen

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4.3-238: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

<< Angaben des pharmazeutischen Unternehmers >>

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung.

<< Angaben des pharmazeutischen Unternehmers >>

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.4.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.4.

<< Angaben des pharmazeutischen Unternehmers >>

4.3.2.5 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.1.3.18 und 4.3.2.4) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Die im Dossier als Resultat der systematischen Übersicht dargelegten Nachweise zum Zusatznutzen von Acclidiniumbromid berücksichtigen Ergebnisse von zwei RCTs mit direktem Vergleich und 24 RCTs für den indirekten Vergleich. Die im Dossier dargestellten Ergebnisse der einzelnen RCTs stellen eine Evidenz der Stufe I dar. Alle vorgelegten Studien sind randomisierte, kontrollierte Studien, von denen bis auf zwei (verwendet für den indirekten Vergleich) alle doppelblind durchgeführt wurden. Die Qualität der Studie M/34273/23 (direkter Vergleich) wird durch die Heranziehung in der Kostenübernahmeentscheidung der dänischen Erstattungsbehörde „Medicintilskudsnaevnet“ bestätigt [67].

In den zwei RCTs für den direkten Vergleich liegen weder auf Studien- noch auf Endpunktebene Verzerrungspotentiale vor.

Bei den 24 RCTs für den indirekten Vergleich sind teilweise Verzerrungspotentiale festzustellen. Bei drei placebokontrollierten Tiotropiumbromid-Studien wurde auf Studienebene aufgrund von subjektiver Berichterstattung auf teilweise unklarer Datenbasis Verzerrungspotential festgestellt.

Auf Endpunktebene wurden bei durchschnittlich der Hälfte der Tiotropiumbromid-Studien Verzerrungspotentiale festgestellt. Dies ist in den meisten Fällen auf die ausschließliche Verfügbarkeit der Publikationen zurückzuführen, die häufig keine methodischen Details (wie zum Beispiel die Erzeugung der Randomisierung) beinhalten. Wenn keine gesicherte Aussage zu den einzelnen Verzerrungsaspekten getroffen werden kann, wird der Verzerrungsaspekt im Zweifelsfall als unklar oder vorhanden eingestuft. Dies führt zu einer eher konservativen Bewertung des Verzerrungspotentials.

Auf der Seite der Acclidiniumbromid-Studien wurde bei der ACCORD II-Studie ein Ungleichgewicht zwischen den Vergleichsgruppen bezüglich des Schweregrades festgestellt. Dies ist laut der Nutzendossievorlage²⁷ per se kein Grund für ein Verzerrungspotential, wurde aber aufgrund des analog zur konservativen Bewertung der Tiotropiumbromid-Studien geplanten Vorgehens weiter verfolgt. Das daraufhin hinzugezogene Zulassungsdossier, welches die primäre gepoolte Analyse mit Zustimmung der EMA ohne die ACCORD II-Studie durchführt, führte zur konservativen Annahme eines Verzerrungspotentials auf Endpunktebene.

²⁷ Abschnitt 4.2.4: „Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.“

Aufgrund des festgestellten Verzerrungspotentials wurden für jeden Endpunkt Sensitivitätsanalysen der Meta-Analysen durchgeführt. Ein verzerrender Einfluss der ACCORD II-Studie wird dabei widerlegt - der Ein- oder Ausschluss hat keinen Einfluss auf die Homogenität des Studienpools oder den Gesamteffektschätzer des jeweiligen Endpunktes.

Der innerhalb der Sensitivitätsanalyse durchgeführte Ausschluss der verzerrten Tiotropiumbromid-Studien führte meist zu einem homogeneren Studienpool, veränderte den Gesamteffektschätzer in den meisten Fällen jedoch nicht grundlegend.

Zusammenfassend besitzen die vorgelegten RCTs eine gute Studienqualität.

Den Beleg für den Zusatznutzen von Acclidiniumbromid liefern die zwei direkten Vergleichsstudien M/34273/23 und M/34273/39 hinsichtlich der Morbidität (COPD-Symptome) und der Lebensqualität (Behandlungszufriedenheit). Der indirekte Vergleich von 3 placebokontrollierten Acclidiniumbromid-Studien und 21 placebokontrollierten Tiotropiumbromid-Studien ergänzt die Ergebnisse des direkten Vergleiches um längerfristige Endpunkte, wie die krankheitsspezifische Lebensqualität, die Exazerbationsrate sowie das Nebenwirkungs- und kardiovaskuläre Sicherheitsprofil. Die Ergebnisse zeigen nicht nur eine Nichtunterlegenheit für Acclidiniumbromid, sondern deuten mindestens auf eine Gleichwertigkeit der beiden Wirkstoffe hin. In einigen Endpunkten - insbesondere bei längerer Studiendauer von 24 Wochen - ist eine numerisch positive Tendenz für Acclidiniumbromid erkennbar.

Der Zusatznutzen wurde mit Hilfe der Auswertung folgender patientenrelevanter Endpunkte bestimmt, deren Validität in Abschnitt 4.2.5.2 dargestellt ist:

- Morbidität (COPD-Symptome)
 - Husten, Auswurf, Atemnot (EXACT-Fragebogen)
 - morgendliche und nächtliche Symptome
 - Veränderung des forcierten expiratorischen Volumens in einer Sekunde (FEV₁ trough, peak, AUC)
 - Dyspnoe (TDI)
 - Exazerbationen
 - Hospitalisierungen aufgrund von Exazerbationen
- Gesundheitsbezogene Lebensqualität
 - Lebensqualität (SGRQ)
 - Behandlungszufriedenheit

- Unerwünschte Arzneimittelwirkungen
 - Unerwünschte Ereignisse
 - Schwerwiegende Ereignisse
 - Studienabbrüche aufgrund von unerwünschten Ereignissen
- Mortalität
 - COPD-assoziierte kardiovaskuläre Morbidität und Mortalität
 - COPD-bedingte Letalität und Gesamtmortalität

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Anhand der im Dossier analysierten patientenrelevanten Endpunkte lässt sich die Linderung von COPD-Symptomen, die Verringerung von Exazerbationen und Hospitalisierungen, eine

bedeutende Verbesserung der Lebensqualität und eine große Behandlungszufriedenheit mit dem Genuair®-Inhalator nachweisen.

Nach §5 (7) Nr. 2 des 5. Kapitels der Verfahrensordnung wurde gegenüber der zweckmäßigen Vergleichstherapie eine moderate und nicht nur eine geringfügige Verbesserung des therapierelevanten Nutzens erreicht. Dies betrifft vor allen Dingen die Verringerung nicht schwerwiegender Symptome. Die teilweise signifikante Überlegenheit von Acridiniumbromid im direkten Vergleich und die im indirekten Vergleich gezeigte Gleichwertigkeit bis numerische Überlegenheit wurden anhand folgender patientenrelevanter Endpunkte gezeigt:

Morbidität (COPD-Symptome)

- Die im EXACT-Fragebogen ermittelten Ausprägungen der COPD-Symptome (Husten, Auswurf, Atemnot) sowie die morgendlichen und abendlichen Symptome konnten ausgehend von der Baselineerhebung in den beiden direkten Vergleichsstudien (M/34273/23 und M/34273/39) gegenüber Tiotropiumbromid numerisch überlegen verringert werden. Die Einschränkung der körperlichen Aktivität durch COPD-Symptome wurde bei Acridiniumbromid-Patienten zudem signifikant ($p=0,04$) seltener als bei Tiotropiumbromid-Patienten beobachtet.
- Die über den Endpunkt FEV₁ AUC analysierte Entwicklung der Lungenfunktion im Tagesverlauf ist unter Acridiniumbromid in der zweiten Tageshälfte gegenüber Tiotropiumbromid signifikant (M/34273/23) bzw. numerisch (M/34273/39) überlegen. Durch die zweimal tägliche Gabe von Acridiniumbromid kann es auch zu einer leichteren Verarbeitung der durch die zirkadiane Rhythmik verursachten tageszeitlicher Schwankungen (belastende COPD-Symptome in der Nacht und am Morgen) durch den Patienten kommen. Numerisch bessere Trough-FEV₁-Werte zeigen die nachhaltige Verbesserung der Lungenfunktion. Die numerisch größeren maximalen Peak-FEV₁-Werte gegenüber Tiotropiumbromid zeigen die sofortige Wirkung der Acridiniumbromid-Behandlung ab dem ersten Tag. Die Auswertung hinsichtlich Trough- und Peak-FEV₁-Werte im indirekten Vergleich ergab eine Gleichwertigkeit der Acridiniumbromid- und Tiotropiumbromid-Behandlung.
- Die mittels des TDI ermittelte Dyspnoe, die Exazerbationsrate sowie die dadurch bedingten Hospitalisierungen konnten nur im indirekten Vergleich ausgewertet werden. Dieser ergab eine Gleichwertigkeit der Acridiniumbromid- und Tiotropiumbromid-Behandlung. Die Gleichwertigkeit stand bei der Hospitalisierungsrate an der Grenze zur signifikanten Überlegenheit zugunsten von Acridiniumbromid.

Gesundheitsbezogene Lebensqualität

- Die mittels des SGRQ über den Studienzeitraum ermittelte Verbesserung der Lebensqualität unter Acridiniumbromid ist gegenüber Placebo signifikant. Im Vergleich zu Tiotropiumbromid wird Acridiniumbromid als mindestens gleichwertig eingestuft.

- Die Behandlungszufriedenheit mit dem Genuair[®]-Inhalator ist signifikant höher als die mit dem HandiHaler[®] - 80% der Patienten der Studie M/34273/39 bevorzugen den Genuair[®]-Inhalator, u.a. bezüglich der einfachen Handhabung und des Feedbacksystems.

Unerwünschte Arzneimittelwirkungen

- Das Nebenwirkungsprofil von Acridiniumbromid unterscheidet sich sowohl im direkten als auch im indirekten Vergleich nicht von dem von Tiotropiumbromid. Die bei Acridiniumbromid nicht erforderliche Dosisanpassungen oder das Ergreifen von Vorsichtsmaßnahmen bei Patienten mit mittlerer bis schwerer Nierenfunktionsstörung birgt gegenüber Tiotropiumbromid einen bedeutenden Vorteil bezüglich der Therapebreite.

Mortalität

- Die Analyse der COPD-bedingten kardiovaskulären Morbidität und Mortalität sowie die COPD-assoziierte Letalität und Gesamtmortalität im direkten und indirekten Vergleich lassen auf ein vergleichbares Sicherheitsprofil von Acridiniumbromid und Tiotropiumbromid schließen.

In dem vorgelegten Dossier wird ein relevanter Zusatznutzen gezeigt und aufgrund der gegenüber der zweckmäßigen Vergleichstherapie erzielten moderaten und nicht nur geringfügigen Verbesserung des therapierelevanten Nutzens als „gering“ im Sinne der Verfo eingestuft.

Der im Dossier gezeigte Zusatznutzen von Acridiniumbromid gegenüber Tiotropiumbromid gilt für COPD-Patienten, die leitlinienkonform mit LAMAs behandelt werden. Dabei handelt es sich um Patienten mit COPD-Stufe II bis IV (Zielpopulation). Um den Anforderungen des G-BA bezüglich der Berücksichtigung der Schweregrade nachzukommen, wurden Subgruppenanalysen für die COPD-Stufen II und III²⁸ vorgenommen.

Die Darstellung der Ergebnisse bezüglich der Schweregrade zeigten im direkten Vergleich keine relevanten Wirksamkeitsunterschiede zwischen den Subgruppen und der gesamten Studienpopulation.

Im indirekten Vergleich konnten aufgrund von fehlenden Subgruppendaten der Tiotropiumbromid-Studien keine Subgruppenanalyse auf Ebene der Netzwerk-Meta-Analyse vorgenommen werden. Auf Ebene der Meta-Analyse war eine Subgruppenanalyse für die placebokontrollierten Acridiniumbromid-Studien möglich. Dabei wurden keine die Nutzenbewertung systematisch beeinflussenden Wirksamkeitsunterschiede festgestellt.

Der im vorliegenden Dossier gezeigte Zusatznutzen gilt daher für die gesamte Zielpopulation.

²⁸ Aufgrund der geringen Zahl von Stufe IV-Patienten war die Auswertung dieser Subgruppe nicht möglich.

Aufgrund der Ergebnisvollständigkeit der dargestellten Endpunkte kann der dargestellte Zusatznutzen von Acridiniumbromid mit großer Sicherheit als faktisch angesehen werden. Im direkten Vergleich (2 Studien) kann der Zusatznutzen von Acridiniumbromid gegenüber Tiotropiumbromid gezeigt werden. Dies entspricht der Aussagewahrscheinlichkeit eines Beleges. Die 24 im indirekten Vergleich herangezogenen RCTs zeigen nicht nur eine Nichtunterlegenheit für Acridiniumbromid, sondern deuteten mindestens auf eine Gleichwertigkeit der beiden Wirkstoffe hin. In einigen Endpunkten - insbesondere bei längerer Studiendauer von 24 Wochen - ist eine numerisch positive Tendenz für Acridiniumbromid erkennbar.

Für die systematische Identifizierung der dem Dossier zugrunde liegenden Ergebnisse wurden öffentlich zugängliche Quellen zu klinischen Studien, Berichte zu klinischen Studien und Literaturdatenbanken, sowie die beim pharmazeutischen Hersteller verfügbaren Informationen zu relevanten klinischen Studien voll umfänglich berücksichtigt. Insofern liegt Ergebnisvollständigkeit vor.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4.4-1: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit chronisch obstruktiver Lungenerkrankung (Stufe II-IV nach Nationaler VersorgungsLeitlinie)	gering

4.4.4 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht – Arzneimittel zur Behandlung eines seltenen Leidens

Für Arzneimittel, die zur Behandlung eines seltenen Leidens zugelassen sind, müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden (Abschnitt 4.3), solange der Umsatz des Arzneimittels mit der gesetzlichen Krankenversicherung zu Apothekenverkaufspreisen einschließlich Umsatzsteuer in den letzten 12 Monaten einen Betrag von 50 Millionen Euro nicht übersteigt.

Für diese Arzneimittel sind Angaben zu Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, im vorliegenden Abschnitt einzureichen. Dabei soll zunächst das Ausmaß des Zusatznutzens (in der Definition gemäß AM-NutzenV) begründet beschrieben werden. Unter Berücksichtigung des Ausmaßes des Zusatznutzens sollen Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, benannt werden. Weitere Abschnitte von Modul 4 können dabei nach Bedarf bearbeitet werden.

Beschreiben Sie begründet das Ausmaß des Zusatznutzens unter Verwendung der Kategorisierung gemäß AM-NutzenV. Berücksichtigen Sie bei den Aussagen zum Zusatznutzen nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen. Verweisen Sie, wo notwendig, auf weitere Abschnitte von Modul 4.

Trifft nicht zu.

Geben Sie auf Basis des Ausmaßes des Zusatznutzens in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4.4-2: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens – Arzneimittel zur Behandlung eines seltenen Leidens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Neben der Bewertung des Nutzens und Zusatznutzens mit Hilfe zweier RCTs mit einem direkten Vergleich des zu bewertenden Arzneimittels und der Vergleichstherapie wurde ein indirekter Vergleich durchgeführt. Die RCTs mit direktem Vergleich liefern valide Ergebnisse zur Verbesserung der Lungenfunktion und der COPD-Symptomatik bei einer Studienpopulation von 30 bzw. 329 Patienten über einen Studienzeitraum von 15 Tagen bzw. 6 Wochen. Zur Erhöhung der Aussagesicherheit wurde zusätzlich ein indirekter Vergleich vorgenommen. Dabei werden längerfristige Effekte demonstriert und weiterer Endpunkte an einer Patientenpopulation von insgesamt 636 Aclidiniumbromid-Patienten analysiert. Die für den indirekten Vergleich herangezogenen Studien eignen sich aufgrund ihrer klinischen, methodologischen und statistischen Homogenität.

Klinische Homogenität

Homogene Patientenpopulation: Die in den Studien eingeschlossenen Patienten sind aufgrund der in den einzelnen Studien definierten Ein- und Ausschlusskriterien vergleichbar. Bei offensichtlichen Unterschieden der Population (z.B. bzgl. des Schweregrades) werden Sensitivitätsanalysen zur Überprüfung der Aussagesicherheit der Ergebnisse durchgeführt.

Homogene Intervention und Kontrolle: Bei den in den Studien untersuchten Interventionen handelt es sich um inhalative Wirkstoffe. Die Vergleichstherapie stellt in jedem Fall Placebo dar.

Homogene Definition von Endpunkten: Die Operationalisierung der zu untersuchenden Endpunkte wird für jede Studie dargestellt. Bei Unterschieden in der Definition und deren vermutete Auswirkungen auf die Ergebnisse (Heterogenitätstest je Endpunkt) wurden Sensitivitätsanalysen zur Überprüfung der Aussagesicherheit der Ergebnisse durchgeführt.

Methodologische Homogenität

Homogene Studiencharakteristik: Hierzu wurden nur RCTs mit einer vergleichbaren Studiendauer herangezogen. Letztes wurde durch die Beschränkung des Studienpools laut Ein- und Ausschlusskriterien (vgl. 4.2.2) auf Studien mit einer Dauer von mindestens 12 Wochen garantiert.

Homogene Studienqualität: Die Beschreibung des Verzerrungspotentials je Studie und je Endpunkt machte mögliche Einflüsse auf die Aussagesicherheit der Ergebnisse sichtbar. Wurden unterschiedliche Verzerrungspotentiale als Ursache für die Abweichung der einzelnen Studienergebnisse voneinander vermutet, wurden Sensitivitätsanalysen durchgeführt. Ergab diese, dass das unterschiedliche Verzerrungspotential einen Einfluss auf

die Effektdifferenz besitzt, wurde bezüglich dieses Endpunktes keine Aussage zum Zusatznutzen getroffen.

Statistische Homogenität

Variabilität der Effekte der Intervention: Um die statistische Heterogenität der eingeschlossenen Studien zu überprüfen, wurden je betrachtetem Endpunkt Homogenitätstests durchgeführt. Der Grad der Heterogenität spiegelte sich im I^2 - und p-Wert, sowie in der Breite der Konfidenzintervalle wider. Lag diese über den in Abschnitt 4.2.5.3 festgelegten Grenzwerten, konnte das Ergebnis des indirekten Vergleiches für diese Endpunkte nur eingeschränkt interpretiert werden.

4.5.2 Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nichtrandomisierte vergleichende Studien (Abschnitt 4.3.2.1.3.18) oder weitere Untersuchungen (Abschnitt 4.3.2.4) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Trifft nicht zu.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Trifft nicht zu.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005²⁹, Molenberghs 2010³⁰). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006³¹) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006³²) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

²⁹ Burzykowski T (Ed.: *The evaluation of surrogate endpoints*. New York: Springer; 2005.

³⁰ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: *A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials*. *Stat Methods Med Res* 2010; 19(3): 205-236.

³¹ Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

³² Weir CJ, Walley RJ. *Statistical evaluation of biomarkers as surrogate endpoints: a literature review*. *Stat Med* 2006; 25(2): 183-203.

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Die Veränderung des forcierten Volumens in einer Sekunde (ermittelt als Trogh- und Peak- FEV_1 und FEV_1AUC) dient der Beschreibung der COPD-Symptomatik und stellt einen patientenrelevanten Endpunkt hinsichtlich der Beurteilung des Therapieansprechens auf Bronchodilatoren dar [1, 5] (vgl. 4.2.5.2). Unabhängig davon ist der FEV_1 ein Surrogat für weitere patientenrelevante Endpunkte. Dies soll im Folgenden dargestellt werden.

Die – wenn auch schwache bis mittlere – Korrelation von FEV_1 mit patientenrelevanten Endpunkten wie Lebensqualität, Exazerbationsfrequenz, COPD-Symptomen und verminderter körperliche Belastbarkeit ist in zahlreichen Studien³³ nachgewiesen (vgl. Tabelle 4.5-1).

Tabelle 4.5-1: Übersicht über die Aussagen zur Validierung/ Korrelation/ Relevanz von FEV_1

Erstautor und Jahr	Publikationsart	Endpunkt	Höhe der Korrelation
Westwood 2011 [68]	Systematischer Review, 36 Studien, Regressionsmodell	Lebensqualität 36 Studien 33 Studien TDI (8 Studien)	LAMA, LABA, Placebo: $r = -0,46, p < 0,01$ LAMA, LABA: $r = -0,35; p = 0,046$ LAMA, LABA, Placebo: $r = 0,56, p = 0,02$
Agusti 2010 [69]	Kohortenstudie mit 2164 Patienten, 3 Jahre	Leistungsfähigkeit Lebensqualität Dyspnoe Exazerbationen	6MWD: $r = 0,34, p < 0,001$ SGRQ: $r = 0,38, p < 0,001$ mMRC: $r = 0,36, p < 0,001$ Exazerbationen: $r = 0,21, p < 0,001$
Jones 2009 [70]	Literatur-Zusammenfassung	Lebensqualität Lebensqualität (FEV ₁ Abnahme)	Jones 1997: Salmeterol (283 Patienten) $R = 0,23, p < 0,00001$ Spencer 2004: fluticasone propionate (FP) $R = 0,24$
Jones 2006 [71]	Literaturübersicht	Mortalität und	schwache Korrelation (kein Wert angegeben) bei einer Senkung von

³³ Suche auf: <http://www.ncbi.nlm.nih.gov/pubmed>, Letzte Aktualisierung der Suche: 04.09.2012. Suchstring: ("Respiratory Function Tests" [MeSH]) AND (association OR correlation OR marker) AND outcomes AND COPD AND ("review"[Publication Type]); 11 Treffer (davon 5 relevant), ergänzend wurden in den als relevant eingestuften Treffern zitierte Publikationen herangezogen

		Gesundheitsstatus	mindestens 50%
Morr 2006 [72]	Literaturübersicht	Mortalität und Lebensqualität	schwacher Korrelation (kein Wert angegeben)
Cazzola 2008 [7]	Literaturübersicht	Dyspnoe, Aktivität und HRQoL	schwacher Korrelation (kein Wert angegeben)
Hoogendoorn 2010 [73]	Review, Regressionsmodell, 37 Studien	Exazerbations-frequenz Exazerbationsbezogene Symptome	Regressionskoeffizient= -0,014, p=0,053 Regressionskoeffizient= -0,009, p=0,19
r= Korrelation, 6MWD=6 minutes walk distance (6 Minuten Gehstest), mMRC= modifizierte Dyspnoeskala des Medical Research Council			

Die Relevanz der Veränderung des FEV₁ bei der Beschreibung weiterer patientenrelevanter Endpunkte im Versorgungsalltag wird auch anhand der folgenden zwei Beispiele deutlich.

Die Empfehlungen zur Behandlung der stabilen COPD der überarbeiteten GOLD Leitlinie [5] richtet sich nach einem Klassifikationsmodell, welches Lebensqualität, Exazerbationsrate, die Schwere der COPD-Symptome und den FEV₁-Wert enthält (Details vgl. Modul 3).

Die Verknüpfung des Endpunktes FEV₁ mit anderen patientenrelevanten Endpunkten wird auch durch die Anforderungen der Zulassungsbehörden bestätigt: Das ‘Committee for Medicinal Products for Human Use‘ (CHMP) der European Medicines Agency (EMA) gibt an, dass der primäre Endpunkt aus der Messung der Lungenfunktion (z. B. FEV₁) und weiterer Symptome (z. B. Exazerbationen, SGRQ) bestehen sollte [4].

Im vorliegenden Nutzendossier wird die Veränderung des FEV₁ primär als patientenrelevanter Endpunkt der Morbidität ausgewertet. Der Einfluss der verringerten Lungenfunktion auf die anderen Endpunkte aus den Kategorien Morbidität und Lebensqualität wurden nicht weiter dargestellt, da für diese Endpunkte eigene Messinstrumente existieren und diese in den Studien angewandt wurden.

4.6 Liste der eingeschlossenen Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Datenquellen auf.

Studien für den direkten Vergleich (Abschnitt 4.3.1)		
M/34273/23	Fuhr R (2010): Clinical Trial M/34273/23 - A multiple dose, double blind, double-dummy, two-week 3 way cross-over, placebo controlled clinical trial to assess the efficacy and safety of twice daily inhaled acclidinium bromide 400 µg compared to placebo and to an active comparator in patients with stable moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 22/04/2010. Confidential.	[32]
	Fuhr R, Magnussen H, Sarem K, Ribera Llovera A, Kirsten AM, Falques M, et al. (2012): Efficacy of Acclidinium Bromide 400 µg Twice Daily Compared With Placebo and Tiotropium in Patients With Moderate to Severe COPD. Chest; 141(3):745-52.	[31]
M/34273/39	Beier J (2011): Clinical Trial Protocol M/34273/39 - A multiple dose, double-blind, double-dummy, placebo controlled, parallel clinical trial to assess the efficacy and safety of twice daily inhaled acclidinium bromide 400 µg compared to placebo and to tiotropium bromide in patients with stable moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 09/2011. Confidential.	[33]
Studien für den indirekten Vergleich (Abschnitt 4.3.2.1)		
ACCORD I	Kerwin EM, D'Urzo AD, Gelb AF, Lakkis H, Garcia GE, Caracta CF, et al. (2012): Efficacy and safety of a 12-week treatment with twice-daily acclidinium bromide in COPD patients (ACCORD COPD I). Copd: Journal of Chronic Obstructive Pulmonary Disease; 9(2):90-101.	[41]
	Kerwin E (2011): Clinical Trial M/34273/33 - Efficacy and Safety of Acclidinium Bromide at Two Dose Levels (200 µg Twice Daily, 400 µg Twice Daily) vs. Placebo When Administered to Patients With Moderate to Severe Chronic Obstructive Pulmonary Disease (COPD). Stand: 21/01/2011. Confidential.	[40]
ATTAIN	Jones PW, Singh D, Bateman ED, Agusti A, Lamarca R, de Miquel G, et al. (2012): Efficacy and safety of twice-daily acclidinium bromide in COPD patients: The ATTAIN study. The European	[39]

	respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; ³⁴	
	Jones PW (2011): Clinical Trial M/34273/34 - Efficacy and safety of aclidinium bromide at two dose levels vs placebo when administered to patients with moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 12/04/2011. Confidential.	[38]
ACCORD II	Taglietti M (2011): Clinical Trial M/34273/38 Part A - A Randomized, Double-blind, Placebo-Controlled Study Evaluating the Efficacy, Safety, and Tolerability of 2 Doses of Aclidinium Bromide Compared With Placebo for 12 Weeks in Patients With Moderate to Severe, Stable Chronic Obstructive Pulmonary Disease Followed by a 40-Week Evaluation of the Higher Aclidinium Bromide Dose. Stand: 16/03/2011. Confidential.	[19]
SAFE	Chan CK, Maltais F, Sigouin C, Haddon JM, Ford GT, Group SS (2007): A randomized controlled trial to assess the efficacy of tiotropium in Canadian patients with chronic obstructive pulmonary disease. Canadian Respiratory Journal; 14(8):465-72.	[42]
SAFE Portugal	Moita J, Barbara C, Cardoso J, Costa R, Sousa M, Ruiz J, et al. (2008): Tiotropium improves FEV ₁ in patients with COPD irrespective of smoking status. Pulmonary Pharmacology & Therapeutics; 21(1):146-51.	[43]
TIPHON	Tonnel AB, Perez T, Grosbois JM, Verkindre C, Bravo ML, Brun M, et al. (2008): Effect of tiotropium on health-related quality of life as a primary efficacy endpoint in COPD. International Journal of COPD; 3(2):301-10.	[44]
UPLIFT	Tashkin DP, Celli B, Senn S, Burkhart D, Kesten S, Menjoge S, et al. (2008): A 4-year trial of tiotropium in chronic obstructive pulmonary disease. New England Journal of Medicine; 359(15):1543-54.	[45]
	Troosters T, Celli B, Lystig T, Kesten S, Mehra S, Tashkin DP, et al. (2010): Tiotropium as a first maintenance drug in COPD: secondary analysis of the UPLIFT trial. European Respiratory Journal; 36(1):65-73.	[49]
	Decramer M, Celli B, Kesten S, Lystig T, Mehra S, Tashkin DP, et al. (2009): Effect of tiotropium on outcomes in patients with moderate chronic obstructive pulmonary disease (UPLIFT): a prespecified subgroup analysis of a randomised controlled trial. Lancet; 374(9696):1171-8.	[74]
	Tashkin D, Celli B, Kesten S, Lystig T, Decramer M (2010): Effect	[46]

³⁴ Bei der bibliografischen Literatursuche wurde lediglich der Abstract gefunden. Durch eine daraufhin stattfindende freie Suche wurde der nur online publizierte Volltext („Epub ahead of print“) gefunden und in den Studienpool aufgenommen.

	of tiotropium in men and women with COPD: results of the 4-year UPLIFT trial. <i>Respiratory Medicine</i> ; 104(10):1495-504.	
	Tashkin DP, Celli BR, Decramer M, Lystig T, Liu D, Kesten S (2012): Efficacy of Tiotropium in COPD Patients with FEV ₁ >= 60% Participating in the UPLIFT[REGISTERED] Trial. <i>Copd: Journal of Chronic Obstructive Pulmonary Disease</i> ; 9(3):289-96.	[48]
	Tashkin DP, Celli B, Kesten S, Lystig T, Mehra S, Decramer M (2010): Long-term efficacy of tiotropium in relation to smoking status in the UPLIFT trial.[Erratum appears in <i>Eur Respir J</i> . 2010 May;35(5):1195]. <i>European Respiratory Journal</i> ; 35(2):287-94.	[47]
Vogelmeier 2008	Vogelmeier C, Kardos P, Harari S, Gans SJ, Stenglein S, Thirlwell J (2008): Formoterol mono- and combination therapy with tiotropium in patients with COPD: a 6-month study. <i>Respiratory Medicine</i> ; 102(11):1511-20.	[50]
SPRUCE	Freeman D, Lee A, Price D (2007): Efficacy and safety of tiotropium in COPD patients in primary care--the SPiRiva Usual CarE (SPRUCE) study. <i>Respiratory Research</i> ; 8:45.	[21]
Verkindre 2006	Verkindre C, Bart F, Aguilaniu B, Fortin F, Guerin JC, Le MC, et al. (2006): The effect of tiotropium on hyperinflation and exercise capacity in chronic obstructive pulmonary disease. <i>Respiration</i> ; 73(4):420-7.	[51]
Covelli 2005	Covelli H, Bhattacharya S, Cassino C, Conoscenti C, Kesten S (2005): Absence of electrocardiographic findings and improved function with once-daily tiotropium in patients with chronic obstructive pulmonary disease. <i>Pharmacotherapy: The Journal of Human Pharmacology & Drug Therapy</i> ; 25(12):1708-18.	[52]
Niewoehner 2005	Niewoehner DE, Rice K, Cote C, Paulson D, Cooper JA, Jr., Korducki L, et al. (2005): Prevention of exacerbations of chronic obstructive pulmonary disease with tiotropium, a once-daily inhaled anticholinergic bronchodilator: a randomized trial.[Summary for patients in <i>Ann Intern Med</i> . 2005 Sep 6;143(5):120; PMID: 16144886]. <i>Annals of Internal Medicine</i> ; 143(5):317-26.	[53]
	Rice KL, Leimer I, Kesten S, Niewoehner DE (2008): Responses to tiotropium in African-American and Caucasian patients with chronic obstructive pulmonary disease. <i>Translational Research: The Journal Of Laboratory & Clinical Medicine</i> ; 152(2):88-94.	[54]
Brusasco 2003	Brusasco V, Hodder R, Miravittles M, Korducki L, Towse L, Kesten S (2003): Health outcomes following treatment for six months with once daily tiotropium compared with twice daily salmeterol in patients with COPD.[Erratum appears in <i>Thorax</i> . 2005 Feb;60(2):105]. <i>Thorax</i> ; 58(5):399-404.	[55]
Casaburi 2000	Casaburi R, Briggs DD, Jr., Donohue JF, Serby CW, Menjoge SS, Witek TJ, Jr. (2000): The spirometric efficacy of once-daily dosing	[56]

	with tiotropium in stable COPD: a 13-week multicenter trial. The US Tiotropium Study Group. Chest; 118(5):1294-302.	
Donohue 2002	Donohue JF, van Noord JA, Bateman ED, Langley SJ, Lee A, Witek TJ, Jr., et al. (2002): A 6-month, placebo-controlled study comparing lung function and health status changes in COPD patients treated with tiotropium or salmeterol. Chest; 122(1):47-55.	[57]
Casaburi 2002	Casaburi R, Mahler DA, Jones PW, Wanner A, San PG, ZuWallack RL, et al. (2002): A long-term evaluation of once-daily inhaled tiotropium in chronic obstructive pulmonary disease. European Respiratory Journal; 19(2):217-24.	[58]
Bateman 2010a	Bateman E, Singh D, Smith D, Disse B, Towse L, Massey D, et al. (2010): Efficacy and safety of tiotropium Respimat SMI in COPD in two 1-year randomized studies. International Journal of COPD; 5:197-208.	[62]
INHANCE	Donohue JF, Fogarty C, Lotvall J, Mahler DA, Worth H, Yorgancioglu A, et al. (2010): Once-daily bronchodilators for chronic obstructive pulmonary disease: indacaterol versus tiotropium. American Journal of Respiratory & Critical Care Medicine; 182(2):155-62.	[59]
Voshaar 2008	Voshaar T, Lapidus R, Maleki-Yazdi R, Timmer W, Rubin E, Lowe L, et al. (2008): A randomized study of tiotropium Respimat Soft Mist inhaler vs. ipratropium pMDI in COPD. Respiratory Medicine; 102(1):32-41.	[63]
Bateman 2010b	Bateman ED, Tashkin D, Siafakas N, Dahl R, Towse L, Massey D, et al. (2010): A one-year trial of tiotropium Respimat plus usual therapy in COPD patients. Respiratory Medicine; 104(10):1460-72.	[64]
Beeh 2006	Beeh KM, Beier J, Buhl R, Stark-Lorenzen P, Gerken F, Metzendorf N, et al. (2006): [Efficacy of tiotropium bromide (Spiriva) in patients with chronic-obstructive pulmonary disease (COPD) of different severities]. [German]. Pneumologie; 60(6):341-6.	[65]
MISTRAL	Dusser D, Bravo ML, Iacono P (2006): The effect of tiotropium on exacerbations and airflow in patients with COPD.[Erratum appears in Eur Respir J. 2006 May;27(5):1076]. European Respiratory Journal; 27(3):547-55.	[60]
Johansson 2008	Johansson G, Lindberg A, Romberg K, Nordstrom L, Gerken F, Roquet A (2008): Bronchodilator efficacy of tiotropium in patients with mild to moderate COPD. Primary Care Respiratory Journal; 17(3):169-75.	[20]
Powrie 2007	Powrie DJ, Wilkinson TM, Donaldson GC, Jones P, Scrine K, Viel K, et al. (2007): Effect of tiotropium on sputum and serum inflammatory markers and exacerbations in COPD. European Respiratory Journal; 30(3):472-8.	[61]

4.7 Referenzliste

Benennen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie in Modul 4 angegeben haben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard).

1. Ärztliches Zentrum für Qualität in der Medizin (2012): Nationale VersorgungsLeitlinie COPD. Stand: 01/2012 [Zugriff: 23.03.2012]. URL: http://www.versorgungsleitlinien.de/themen/copd/pdf/nvl_copd_lang.pdf.
2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (2011): Allgemeine Methoden Version 4.0. Stand: 23.09.2011 [Zugriff: 30.01.2012]. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf.
3. Gemeinsamer Bundesausschuss (2012): Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV - Beratungsanforderung 2012-B-015 - Acclidiniumbromid zur Behandlung der COPD. Berlin: 21.06.2012.
4. Committee for Medicinal Products for Human Use (CHMP) (2012): Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD). Stand: 21.06.2012 [Zugriff: 10.08.2012]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/08/WC500130880.pdf.
5. Global Initiative for Chronic Obstructive Lung Disease (GOLD) (2011): Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease - revised 2011. Stand: 12/2011 [Zugriff: 04.04.2012]. URL: http://www.goldcopd.org/uploads/users/files/GOLD_Report_2011_Feb21.pdf.
6. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. (2005): Standardisation of spirometry. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; 26(2):319-38.
7. Cazzola M, MacNee W, Martinez FJ, Rabe KF, Franciosi LG, Barnes PJ, et al. (2008): Outcomes for COPD pharmacological trials: from lung function to biomarkers. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; 31(2):416-69.
8. Bundesversicherungsamt (2011): Kriterien des Bundesversicherungsamtes zur Evaluation strukturierter Behandlungsprogramme. Stand: 30.09.2011 [Zugriff: 22.03.2012]. URL: http://www.bundesversicherungsamt.de/nr_1046154/DE/DMP/Downloads/Downloads_Evaluation_gesamt,templateId=raw,property=publicationFile.pdf/Downloads_Evaluation_gesamt.pdf.
9. Donohue JF (2005): Minimal clinically important differences in COPD lung function. Copd; 2(1):111-24.
10. Gemeinsamer Bundesausschuss (2011): Zusammenfassende Dokumentation über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage IX – Festbetragsgruppenbildung und Anlage X – Aktualisierung von Vergleichsgrößen Beta2-sympathomimetische Antiasthmatica, Gruppe 8, in Stufe 2 nach § 35 Absatz 1 SGB V vom 23.Juni 2011.
11. Witek TJ, Jr., Mahler DA (2003): Minimal important difference of the transition dyspnoea index in a multinational clinical trial. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology; 21(2):267-72.

12. Sexton CC, Leidy NK, Notte SM, Jones P, Monz BU, Nelsen L, et al. (2011): Quantifying The Severity Of Respiratory Symptoms Of Chronic Obstructive Pulmonary Disease (COPD): Reliability And Validity Of A Patient Diary. *Am J Respir Crit Care Med*; 183:A1481.
13. United BioSource Corporation (2010): The Exacerbations of Chronic Pulmonary Disease Tool-Respiratory Symptoms (E-RS). Scoring Guide (Version 1.0). Confidential.
14. Agusti A, Hedner J, Marin JM, Barbe F, Cazzola M, Rennard S (2011): Night-time symptoms: a forgotten dimension of COPD. *European respiratory review : an official journal of the European Respiratory Society*; 20(121):183-94.
15. Partridge MR, Karlsson N, Small IR (2009): Patient insight into the impact of chronic obstructive pulmonary disease in the morning: an internet survey. *Current medical research and opinion*; 25(8):2043-8.
16. Jones PW (2005): St. George's Respiratory Questionnaire: MCID. *Copd*; 2(1):75-9.
17. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003): Measuring inconsistency in meta-analyses. *Bmj*; 327(7414):557-60.
18. Boehringer Ingelheim International GmbH (2004): A randomised, double-blind, double-dummy, placebo-controlled, crossover efficacy and safety comparison of 4-week treatment periods of two doses [5 µg (2 actuations of 2.5 µg) and 10 µg (2 actuations of 5 µg)] of tiotropium inhalation solution delivered by the Respimat inhaler, tiotropium inhalation powder capsule (18µg) delivered by the HandiHaler in patients with chronic obstructive pulmonary disease (COPD). Stand: 26.11.2004.
19. Taglietti M (2011): Clinical Trial M/34273/38 Part A - A Randomized, Double-blind, Placebo-Controlled Study Evaluating the Efficacy, Safety, and Tolerability of 2 Doses of Acclidinium Bromide Compared With Placebo for 12 Weeks in Patients With Moderate to Severe, Stable Chronic Obstructive Pulmonary Disease Followed by a 40-Week Evaluation of the Higher Acclidinium Bromide Dose. Stand: 16/03/2011. Confidential.
20. Johansson G, Lindberg A, Romberg K, Nordstrom L, Gerken F, Roquet A (2008): Bronchodilator efficacy of tiotropium in patients with mild to moderate COPD. *Primary Care Respiratory Journal*; 17(3):169-75.
21. Freeman D, Lee A, Price D (2007): Efficacy and safety of tiotropium in COPD patients in primary care--the SPiRiva Usual CarE (SPRUCE) study. *Respiratory Research*; 8:45.
22. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (2012): Tiotropiumbromid bei COPD - Abschlussbericht. Stand: 26.08.2012 [Zugriff: 22.08.2012]. URL: https://www.iqwig.de/download/A05-18_Abschlussbericht_Tiotropiumbromid-bei-COPD.pdf.
23. Cope S, Capkun-Niggli G, Gale R, Jardim JR, Jansen JP (2011): Comparative efficacy of indacaterol 150 µg and 300 µg versus fixed-dose combinations of formoterol+ budesonide or salmeterol+ fluticasone for the treatment of chronic obstructive pulmonary disease—a network meta-analysis. *International journal of chronic obstructive pulmonary disease*; 6:329.
24. Caldwell DM, Ades AE, Higgins JP (2005): Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *Bmj*; 331(7521):897-900.
25. Lu G, Ades AE (2004): Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in medicine*; 23(20):3105-24.

26. Jansen JP, Crawford B, Bergman G, Stam W (2008): Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*; 11(5):956-64.
27. Dempster AP (1997): The direct use of likelihood for significance testing. *Statistics and computing*; 7(4):247-52.
28. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002): Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*; 64(4):583-639.
29. Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000): WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and computing*; 10(4):325-37.
30. Dias S, Welton NJ, Sutton AJ, Ades AE (2012): NICE DSU technical support document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomized controlled trials. Stand: 04/2012 [Zugriff: 28.08.2012]. URL: <http://www.nicedsu.org.uk/TSD2%20General%20meta%20analysis.final.08.05.12.pdf>
31. Fuhr R, Magnussen H, Sarem K, Ribera Llovera A, Kirsten AM, Falques M, et al. (2012): Efficacy of Acclidinium Bromide 400 µg Twice Daily Compared With Placebo and Tiotropium in Patients With Moderate to Severe COPD. *Chest*; 141(3):745-52.
32. Fuhr R (2010): Clinical Trial M/34273/23 - A multiple dose, double blind, double-dummy, two-week 3 way cross-over, placebo controlled clinical trial to assess the efficacy and safety of twice daily inhaled acclidinium bromide 400 µg compared to placebo and to an active comparator in patients with stable moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 22/04/2010. Confidential.
33. Beier J (2012): Clinical Trial Protocol M/34273/39 - A multiple dose, double-blind, double-dummy, placebo controlled, parallel clinical trial to assess the efficacy and safety of twice daily inhaled acclidinium bromide 400 µg compared to placebo and to tiotropium bromide in patients with stable moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 31/08/2012. Confidential.
34. Global Initiative for Chronic Obstructive Lung Disease (GOLD) (2006): Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Stand: 2006 [Zugriff: 12.07.2012]. URL: http://www.who.int/respiratory/copd/GOLD_WR_06.pdf.
35. Welte T, Metzenauer P, Hartmann U (2008): Once versus twice daily formoterol via Novolizer for patients with moderate to severe COPD--a double-blind, randomised, controlled trial. *Pulm Pharmacol Ther*; 21(1):4-13.
36. Leidy NK, Wilcox TK, Jones PW, Murray L, Winnette R, Howard K, et al. (2010): Development of the EXAcerbations of Chronic Obstructive Pulmonary Disease Tool (EXACT): a patient-reported outcome (PRO) measure. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*; 13(8):965-75.
37. Leidy NK, Wilcox TK, Jones PW, Jones P, Roberts L, Powers JH, et al. (2011): Standardizing measurement of chronic obstructive pulmonary disease exacerbations: Reliability and validity of a patient-reported diary. *American journal of respiratory and critical care medicine*; 183(3):323-9.

38. Jones PW (2011): Clinical Trial M/34273/34 - Efficacy and safety of acclidinium bromide at two dose levels vs placebo when administered to patients with moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 12/04/2011. Confidential.
39. Jones PW, Singh D, Bateman ED, Agusti A, Lamarca R, de Miquel G, et al. (2012): Efficacy and safety of twice-daily acclidinium bromide in COPD patients: The ATTAIN study. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology;
40. Kerwin E (2011): Clinical Trial M/34273/33 - Efficacy and Safety of Acclidinium Bromide at Two Dose Levels (200 µg Twice Daily, 400 µg Twice Daily) vs. Placebo When Administered to Patients With Moderate to Severe Chronic Obstructive Pulmonary Disease (COPD). Stand: 21/01/2011. Confidential.
41. Kerwin EM, D'Urzo AD, Gelb AF, Lakkis H, Garcia GE, Caracta CF, et al. (2012): Efficacy and safety of a 12-week treatment with twice-daily acclidinium bromide in COPD patients (ACCORD COPD I). Copd: Journal of Chronic Obstructive Pulmonary Disease; 9(2):90-101.
42. Chan CK, Maltais F, Sigouin C, Haddon JM, Ford GT, Group SS (2007): A randomized controlled trial to assess the efficacy of tiotropium in Canadian patients with chronic obstructive pulmonary disease. Canadian Respiratory Journal; 14(8):465-72.
43. Moita J, Barbara C, Cardoso J, Costa R, Sousa M, Ruiz J, et al. (2008): Tiotropium improves FEV1 in patients with COPD irrespective of smoking status. Pulmonary Pharmacology & Therapeutics; 21(1):146-51.
44. Tonnel AB, Perez T, Grosbois JM, Verkindre C, Bravo ML, Brun M, et al. (2008): Effect of tiotropium on health-related quality of life as a primary efficacy endpoint in COPD. International Journal of COPD; 3(2):301-10.
45. Tashkin DP, Celli B, Senn S, Burkhart D, Kesten S, Menjoge S, et al. (2008): A 4-year trial of tiotropium in chronic obstructive pulmonary disease. New England Journal of Medicine; 359(15):1543-54.
46. Tashkin D, Celli B, Kesten S, Lystig T, Decramer M (2010): Effect of tiotropium in men and women with COPD: results of the 4-year UPLIFT trial. Respiratory Medicine; 104(10):1495-504.
47. Tashkin DP, Celli B, Burkhart D, Kesten S, Liu D, Mehra S, et al. (2009): Long-term efficacy of tiotropium in continuing smokers vs sustained ex-smokers in the UPLIFT trial [Abstract]. American Thoracic Society International Conference, May15 20, 2009, San Diego:A6175.
48. Tashkin DP, Celli BR, Decramer M, Lystig T, Liu D, Kesten S (2012): Efficacy of Tiotropium in COPD Patients with FEV1 \geq 60% Participating in the UPLIFT[REGISTERED] Trial. Copd: Journal of Chronic Obstructive Pulmonary Disease; 9(3):289-96.
49. Troosters T, Celli B, Lystig T, Kesten S, Mehra S, Tashkin DP, et al. (2010): Tiotropium as a first maintenance drug in COPD: secondary analysis of the UPLIFT trial. European Respiratory Journal; 36(1):65-73.
50. Vogelmeier C, Kardos P, Harari S, Gans SJ, Stenglein S, Thirlwell J (2008): Formoterol mono- and combination therapy with tiotropium in patients with COPD: a 6-month study. Respiratory Medicine; 102(11):1511-20.
51. Verkindre C, Bart F, Aguilaniu B, Fortin F, Guerin JC, Le MC, et al. (2006): The effect of tiotropium on hyperinflation and exercise capacity in chronic obstructive

- pulmonary disease. *Respiration; international review of thoracic diseases*; 73(4):420-7.
52. Covelli H, Bhattacharya S, Cassino C, Conoscenti C, Kesten S (2005): Absence of electrocardiographic findings and improved function with once-daily tiotropium in patients with chronic obstructive pulmonary disease. *Pharmacotherapy: The Journal of Human Pharmacology & Drug Therapy*; 25(12):1708-18.
 53. Niewoehner DE, Rice K, Cote C, Paulson D, Cooper JA, Jr., Korducki L, et al. (2005): Prevention of exacerbations of chronic obstructive pulmonary disease with tiotropium, a once-daily inhaled anticholinergic bronchodilator: a randomized trial.[Summary for patients in *Ann Intern Med*. 2005 Sep 6;143(5):I20; PMID: 16144886]. *Annals of Internal Medicine*; 143(5):317-26.
 54. Rice KL, Leimer I, Kesten S, Niewoehner DE (2008): Responses to tiotropium in African-American and Caucasian patients with chronic obstructive pulmonary disease. *Translational Research: The Journal Of Laboratory & Clinical Medicine*; 152(2):88-94.
 55. Brusasco V, Hodder R, Miravittles M, Korducki L, Towse L, Kesten S (2003): Health outcomes following treatment for six months with once daily tiotropium compared with twice daily salmeterol in patients with COPD.[Erratum appears in *Thorax*. 2005 Feb;60(2):105]. *Thorax*; 58(5):399-404.
 56. Casaburi R, Briggs DD, Jr., Donohue JF, Serby CW, Menjoge SS, Witek TJ, Jr. (2000): The spirometric efficacy of once-daily dosing with tiotropium in stable COPD: a 13-week multicenter trial. The US Tiotropium Study Group. *Chest*; 118(5):1294-302.
 57. Donohue JF, van Noord JA, Bateman ED, Langley SJ, Lee A, Witek TJ, Jr., et al. (2002): A 6-month, placebo-controlled study comparing lung function and health status changes in COPD patients treated with tiotropium or salmeterol. *Chest*; 122(1):47-55.
 58. Casaburi R, Mahler DA, Jones PW, Wanner A, San PG, ZuWallack RL, et al. (2002): A long-term evaluation of once-daily inhaled tiotropium in chronic obstructive pulmonary disease. *European Respiratory Journal*; 19(2):217-24.
 59. Donohue JF, Fogarty C, Lotvall J, Mahler DA, Worth H, Yorgancioglu A, et al. (2010): Once-daily bronchodilators for chronic obstructive pulmonary disease: indacaterol versus tiotropium. *American Journal of Respiratory & Critical Care Medicine*; 182(2):155-62.
 60. Dusser D, Bravo ML, Iacono P (2006): The effect of tiotropium on exacerbations and airflow in patients with COPD.[Erratum appears in *Eur Respir J*. 2006 May;27(5):1076]. *European Respiratory Journal*; 27(3):547-55.
 61. Powrie DJ, Wilkinson TM, Donaldson GC, Jones P, Scrine K, Viel K, et al. (2007): Effect of tiotropium on sputum and serum inflammatory markers and exacerbations in COPD. *European Respiratory Journal*; 30(3):472-8.
 62. Bateman E, Singh D, Smith D, Disse B, Towse L, Massey D, et al. (2010): Efficacy and safety of tiotropium Respimat SMI in COPD in two 1-year randomized studies. *International Journal of COPD*; 5:197-208.
 63. Voshaar T, Lapidus R, Maleki-Yazdi R, Timmer W, Rubin E, Lowe L, et al. (2008): A randomized study of tiotropium Respimat Soft Mist inhaler vs. ipratropium pMDI in COPD. *Respiratory Medicine*; 102(1):32-41.

64. Bateman ED, Tashkin D, Siafakas N, Dahl R, Towse L, Massey D, et al. (2010): A one-year trial of tiotropium Respimat plus usual therapy in COPD patients. *Respiratory Medicine*; 104(10):1460-72.
65. Beeh KM, Beier J, Buhl R, Stark-Lorenzen P, Gerken F, Metzdorf N, et al. (2006): [Efficacy of tiotropium bromide (Spiriva) in patients with chronic-obstructive pulmonary disease (COPD) of different severities]. [German]. *Pneumologie (Stuttgart, Germany)*; 60(6):341-6.
66. Almirall SA (2011): Common technical document - 2.5 Clinical overview. Stand: 25.07.2011. Confidential.
67. Medicintilskudsnævnet (Dänische Erstattungsbehörde) (2012): Medicintilskudsnævnets indstilling [Recommendation from Reimbursement Committee - Eklira Genuair]. Schreiben vom: 30. August 2012.
68. Westwood M, Bourbeau J, Jones PW, Cerulli A, Capkun-Niggli G, Worthy G (2011): Relationship between FEV1 change and patient-reported outcomes in randomised trials of inhaled bronchodilators for stable COPD: a systematic review. *Respir Res*; 12:40.
69. Agusti A, Calverley PM, Celli B, Coxson HO, Edwards LD, Lomas DA, et al. (2010): Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respir Res*; 11:122.
70. Jones PW (2009): Health status and the spiral of decline. *Copd*; 6(1):59-63.
71. Jones PW, Agusti AG (2006): Outcomes and markers in the assessment of chronic obstructive pulmonary disease. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*; 27(4):822-32.
72. Morr H (2006): [Therapeutic goals and outcome measurements in COPD patients]. (Therapieziele und Messungen des Therapieerfolgs bei COPD.). *Medizinische Klinik (Munich, Germany)*; 101(4):279-82.
73. Hoogendoorn M, Feenstra TL, Hoogenveen RT, Al M, Molken MR (2010): Association between lung function and exacerbation frequency in patients with COPD. *International journal of chronic obstructive pulmonary disease*; 5:435-44.
74. Decramer M, Celli B, Kesten S, Lystig T, Mehra S, Tashkin DP, et al. (2009): Effect of tiotropium on outcomes in patients with moderate chronic obstructive pulmonary disease (UPLIFT): a prespecified subgroup analysis of a randomised controlled trial. *Lancet*; 374(9696):1171-8.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.2 beschriebene bibliografische Literaturrecherche an, und zwar einzeln für jede Datenbank. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ³⁵] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

³⁵ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455.

Suchstrategie zur Identifizierung von RCTs mit direktem Vergleich von Acclidiniumbromid und Tiotropiumbromid (Abschnitt 4.3.1):

Datenbankname	MEDLINE
Suchoberfläche	PubMed
Datum der Suche	28.08.2012
Zeitsegment	Bis 2012 KW 35
Suchfilter	Filter für randomisierte kontrollierte Studien ³⁶

#	Suchbegriffe	Ergebnis
1	"(3R)-3-((hydroxy(di-2-thienyl)acetyl)oxy)-1-(3-phenoxypropyl)-1-azoniabicyclo(2.2.2)octane bromide"[Supplementary Concept] ³⁷	30
2	"aclidinium bromide" [all fields] OR aclidinium [all fields] OR Eklira [all fields]	44
3	#1 OR #2	45
4	"Tiotropium" [Supplementary Concept] ³⁸	549
5	"tiotropium bromide" [all fields] OR tiotropium [all fields] OR Spiriva [all fields]	814
6	#4 OR #5	814
7	#3 AND #6	13
8	"Pulmonary Disease, Chronic Obstructive "[Mesh]	18336
9	"Chronic Obstructive Pulmonary Disease" [text word]	29889
10	"COPD" [text word]	21042
11	#8 OR #9 OR #10	33785
12	#7 AND #11	12
13	randomized controlled trial [Publication Type]	329335
14	randomized [Title/Abstract]	269672
15	#13 OR #14	441990
16	#12 AND #15	3

³⁶In Wong et al. 2006 wird zur Suche nach RCT's in MEDLINE (Suchoberfläche Ovid) folgende Sucheinschränkung empfohlen: "randomized controlled trial [Publication Type] OR randomized [multiple posting] OR placebo [multiple posting]". In der verwendeten Suchoberfläche PubMed wird für [multiple posting] der Feldname [Title/Abstract] verwendet. Für die hier beschriebene Suche nach RCT's mit direktem Vergleich wird auf die Sucheinschränkung 'placebo [Title/Abstract]' verzichtet.

³⁷ Eingeschlossen sind (Entry Terms): Acclidinium bromide.

³⁸ Eingeschlossen sind (Entry Terms): Spiriva, tiotropium bromide, 7-((hydroxybis(2-thienyl)acetyl)oxy)-9,9-dimethyl-3-oxa-9-azoniatricyclo(3.3.1.0(2,4))nonane bromide, BA 679 BR, BA-679 BR.

Datenbankname	EMBASE
Suchoberfläche	Elsevier
Datum der Suche	28.08.2012
Zeitsegment	Bis 2012 KW 35
Suchfilter	Filter für randomisierte kontrollierte Studien ³⁹ Filter für EMBASE-Ergebnisse: [embase]/lim NOT [medline]/lim ⁴⁰

#	Suchbegriffe	Ergebnis
1	'aclidinium bromide'/exp OR 'aclidinium bromide'	55
2	Eklira	0
3	#1 OR #2	55
4	'tiotropium bromide'/exp OR 'tiotropium bromide'	1088
5	Spiriva	224
6	#4 OR #5	1091
7	#3 AND #6	27
8	'chronic obstructive lung disease'/exp	18961
9	#7 AND #8	23
10	random:ab:ti	36554
11	double-blind:ab:ti	32209
12	'randomized controlled trial'/exp OR 'randomized controlled trial'	67879
13	#10 OR #11 OR #12	123279
14	#9 AND #13	3

³⁹In Wong et al. 2006 wird zur Suche nach RCT's in EMBASE (Suchoberfläche Ovid) folgende Sucheinschränkung empfohlen: "*random:.tw. [textword] OR placebo:.mp. [multiple posting] OR double-blind:.tw. [textword]*". In der verwendeten Suchoberfläche Elsevier wird für *.mp. [multiple posting]* der Feldname *:ti:ab [Title/Abstract]* verwendet. Für die hier beschriebene Suche nach RCT's mit direktem Vergleich wird auf die Sucheinschränkung *,placebo:ti:ab'* verzichtet. Als zusätzlicher Filter wird das Ergebnis der 'Extensive search' für den Studientyp Randomized controlled trial '*randomized controlled trial'/exp OR 'randomized controlled trial'*' verwendet.

⁴⁰ Um die Zahl der Duplikate zu reduzieren, werden nur Ergebnisse angezeigt, die nicht in MEDLINE angezeigt werden.

Datenbankname Cochrane
Suchoberfläche Cochrane Library
Datum der Suche 28.08.2012
Zeitsegment keine Einschränkung
Suchfilter nur klinische Studien [Trials]

#	Suchbegriffe	Ergebnis
1	(aclidinium bromide) or (aclidinium) or (eklira)	45
2	(tiotropium bromide) or (tiotropium) or (spiriva)	492
3	#1 AND #2	7
4	MeSH descriptor Pulmonary Disease, Chronic Obstructive explode all trees	1835
	copd	5878
5	#3 OR #4	6247
6	#4 AND #5	6

Suchstrategie zur Identifizierung von RCTs für einen indirekten Vergleich von Acclidiniumbromid und Tiotropiumbromid (Abschnitt 4.3.2.1):

Datenbankname	EMBASE
Suchoberfläche	Ovid
Datum der Suche	05.Juli.2012
Zeitsegment	1988 bis 2012 (Woche 25)
Suchfilter	Sprache: Deutsch und Englisch, Ausschluss von Tierstudien

#	Suchbegriffe	Ergebnis
1	tiotropium bromide/	2304
2	aclidinium bromide/	100
3	(tiotropium or spiriva or aclidinium bromide or Eklira).ab. or (tiotropium or spiriva or aclidinium bromide or Eklira).ti.	1040
4	(COPD or chronic obstructive pulmonary disease or COAD or chronic obstructive airway disease or chronic obstructive lung disease or chronic bronchitis or emphysema).ab. or (COPD or chronic obstructive pulmonary disease or COAD or chronic obstructive airway disease or chronic obstructive lung disease or chronic bronchitis or emphysema).ti.	53880
5	exp Pulmonary Disease, Chronic Obstructive/ or exp Chronic obstructive lung disease/	52044
6	(randomised or randomized or randomly or placebo or trial).ab. or (randomised or randomized or randomly or placebo or trial).ti.	778731
7	exp RANDOMIZED CONTROLLED TRIAL/	307240
8	exp controlled clinical trial/	436813
9	1 or 2 or 3	2515
10	4 or 5	72768
11	6 or 7 or 8	943769
12	9 and 10 and 11	563
13	limit 12 to (english or german)	532
14	limit 13 to yr="1989 -Current"	532
15	(animals not humans).sh.	0
16	14 not 15	532

Datenbankname	MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R)
Suchoberfläche	Ovid
Datum der Suche	05.Juli.2012
Zeitsegment	1946 bis 2012 (Woche 25)
Suchfilter	Sprache: Deutsch und Englisch, Ausschluss von Tierstudien

#	Suchbegriffe	Ergebnis
1	tiotropium bromide/	0
2	aclidinium bromide/	0
3	(tiotropium or spiriva or aclidinium bromide or Eklira).ab. or (tiotropium or spiriva or aclidinium bromide or Eklira).ti.	707
4	(COPD or chronic obstructive pulmonary disease or COAD or chronic obstructive airway disease or chronic obstructive lung disease or chronic bronchitis or emphysema).ab. or (COPD or chronic obstructive pulmonary disease or COAD or chronic obstructive airway disease or chronic obstructive lung disease or chronic bronchitis or emphysema).ti.	52948
5	exp Pulmonary Disease, Chronic Obstructive/ or exp Chronic obstructive lung disease/	18361
6	(randomised or randomized or randomly or placebo or trial).ab. or (randomised or randomized or randomly or placebo or trial).ti.	673152
7	exp RANDOMIZED CONTROLLED TRIAL/	330719
8	exp controlled clinical trial/	84400
9	1 or 2 or 3	707
10	4 or 5	56103
11	6 or 7 or 8	801565
12	9 and 10 and 11	307
13	limit 12 to (english or german)	292
14	limit 13 to yr="1989 -Current"	292
15	(animals not humans).sh.	3644611
16	14 not 15	292

Datenbankname Cochrane
Suchoberfläche -
Datum der Suche 05.Juli.2012
Zeitsegment 1989 bis 2012
Suchfilter Sprache: Deutsch und Englisch, Ausschluss von Tierstudien

#	Suchbegriffe	Ergebnis
1	MeSH descriptor Pulmonary Disease, Chronic Obstructive explode all trees	1809
2	(COPD or chronic obstructive pulmonary disease or COAD or chronic obstructive airway disease or chronic obstructive lung disease or chronic bronchitis or emphysema), from 1989 to 2012	8462
3	tiotropium or spiriva or eklira or aclidinium, from 1989 to 2012	523
4	((#1 OR #2) AND #3), from 1989 to 2012 in trials	426

Anhang 4-B: Suchstrategien – Suche in Studienregistern

Geben Sie nachfolgend die Suchstrategien für die in Abschnitt 4.2.3.3 beschriebene Suche in Studienregistern an, und zwar einzeln für jedes Studienregister. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. *clinicaltrials.gov*), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. *http://www.clinicaltrials.gov*), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Geben Sie zusätzlich an, für welchen Abschnitt von Modul 4 die Suche durchgeführt wurde.

Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	08.12.2010
Suchstrategie	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer	23

Suche nach RCTs mit direktem Vergleich (Abschnitt 4.3.1)

Studienregister clinicaltrials.gov

Internetadresse www.clinicaltrials.gov

Datum der Suche 29.08.2012

Suchstrategie((aclidinium bromide OR Eklira) AND (tiotropium bromide OR spiriva))
(COPD OR chronic obstructive pulmonary disease OR COAD OR chronic obstructive airway disease OR chronic obstructive lung disease)

Treffer 3

Studienregister WHO ICTRP

Internetadresse <http://apps.who.int/trialsearch/>

Datum der Suche 29.08.2012

Suchstrategie acclidinium bromide OR Eklira [intervention] AND COPD OR chronic obstructive pulmonary disease OR COAD OR chronic obstructive airway disease OR chronic obstructive lung disease [condition] AND [recruitment status: ALL]

Treffer 38

Studienregister Clinicalstudyresults.org

Internetadresse <http://www.clinicalstudyresults.org/>

Datum der Suche 29.08.2012

Suchstrategie -

Treffer Kein Zugriff möglich

Suche nach RCTs für den indirekten Vergleich (Abschnitt 4.3.2.1)**Tiotropiumbromid**

Studienregister clinicaltrials.gov
Internetadresse <http://www.clinicaltrials.gov/>
Datum der Suche 06.07.2012
Suchstrategie (tiotropium OR spiriva) [ALL-FIELDS]
AND copd [CONDITION]
AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer 140

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 06.07.2012
Suchstrategie (tiotropium OR spiriva) [INTERVENTION] AND copd
[CONDITION]
Treffer 77

Studienregister Clinicalstudyresults.org
Internetadresse <http://www.clinicalstudyresults.org/>
Datum der Suche 06.07.2012
Suchstrategie -
Treffer Kein Zugriff möglich

Acridiniumbromid

Studienregister clinicaltrials.gov
Internetadresse <http://www.clinicaltrials.gov/>
Datum der Suche 06.07.2012
Suchstrategie (aclidinium OR eklira) [ALL-FIELDS]
AND copd [CONDITION]
AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
Treffer 22

Studienregister WHO ICTRP
Internetadresse <http://apps.who.int/trialsearch/AdvSearch.aspx>
Datum der Suche 06.07.2012
Suchstrategie (aclidinium OR eklira) [INTERVENTION] AND copd
[CONDITION]
Treffer 2

Studienregister Clinicalstudyresults.org
Internetadresse <http://www.clinicalstudyresults.org/>
Datum der Suche 06.07.2012
Suchstrategie -
Treffer Kein Zugriff möglich

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Studien mit Ausschlussgrund

Listen Sie die im Volltext gesichteten und ausgeschlossenen Studien unter Nennung des Ausschlussgrunds auf.

In der bibliografischen Literaturrecherche nach RCTs mit direktem Vergleich wurden keine im Volltext gesichteten Publikationen ausgeschlossen.

In der bibliografischen Literaturrecherche nach RCTs für den indirekten Vergleich wurden 47 nach Volltestscreening ausgeschlossen. Diese sind im Folgenden, sortiert nach Ausschlussgrund und fortlaufend nummeriert, aufgelistet.

Ausschlussgrund A2: Intervention nicht wie definiert (n=3)	
1.	Aaron SD, Vandemheen KL, Fergusson D, Maltais F, Bourbeau J, Goldstein R, et al. (2007): Tiotropium in combination with placebo, salmeterol, or fluticasone-salmeterol for treatment of chronic obstructive pulmonary disease: a randomized trial. <i>Annals of Internal Medicine</i> ; 146(8):545-55.
2.	Ambrosino N, Foglio K, Balzano G, Paggiaro PL, Lessi P, Kesten S, et al. (2008): Tiotropium and exercise training in COPD patients: effects on dyspnea and exercise tolerance. <i>International Journal of COPD</i> ; 3(4):771-80.
3.	Jones PW, Rennard SI, Agusti A, Chanez P, Magnussen H, Fabbri L, et al. (2011): Efficacy and safety of once-daily aclidinium in chronic obstructive pulmonary disease. <i>Respiratory Research</i> 26; 12-55.
Ausschlussgrund A3: Vergleichstherapie nicht wie definiert (n=12)	
4.	Tashkin DP, Pearle J, Iezzoni D, Varghese ST (2009): Formoterol and Tiotropium compared with Tiotropium alone for treatment of COPD. <i>Copd: Journal of Chronic Obstructive Pulmonary Disease</i> ; 6(1):17-25.
5.	Vogelmeier C, Hederer B, Glaab T, Schmidt H, Rutten-van Molken MP, Beeh KM, et al. (2011): Tiotropium versus salmeterol for the prevention of exacerbations of COPD. <i>New England Journal of Medicine</i> ; 364(12):1093-103.
6.	Wedzicha JA, Calverley PM, Seemungal TA, Hagan G, Ansari Z, Stockley RA, et al. (2008): The prevention of chronic obstructive pulmonary disease exacerbations by salmeterol/fluticasone propionate or Tiotropium bromide. <i>American Journal of Respiratory & Critical Care Medicine</i> ; 177(1):19-26.
7.	Perng DW, Tao CW, Su KC, Tsai CC, Liu LY, Lee YC (2009): Anti-inflammatory effects of salmeterol/fluticasone, Tiotropium/fluticasone or Tiotropium in COPD. <i>European Respiratory Journal</i> ; 33(4):778-84.
8.	Briggs DD, Jr., Covelli H, Lapidus R, Bhattacharya S, Kesten S, Cassino C (2005): Improved daytime spirometric efficacy of Tiotropium compared with salmeterol in patients with COPD. <i>Pulmonary Pharmacology & Therapeutics</i> ; 18(6):397-404.

9.	Cazzola M, Ando F, Santus P, Ruggeri P, Di MF, Sanduzzi A, et al. (2007): A pilot study to assess the effects of combining fluticasone propionate/salmeterol and Tiotropium on the airflow obstruction of patients with severe-to-very severe COPD. <i>Pulmonary Pharmacology & Therapeutics</i> ; 20(5):556-61.
10.	Hoshino M, Ohtawa J (2011): Effects of adding salmeterol/fluticasone propionate to Tiotropium on airway dimensions in patients with chronic obstructive pulmonary disease. <i>Respirology</i> ; 16(1):95-101.
11.	Buhl R, Dunn LJ, Disdier C, Lassen C, Amos C, Henley M, et al. (2011): Blinded 12-week comparison of once-daily indacaterol and Tiotropium in COPD. <i>European Respiratory Journal</i> ; 38(4):797-803.
12.	Fabbri LM (2011): Tiotropium reduced exacerbations more than salmeterol in moderate-to-very severe COPD. <i>Annals of Internal Medicine</i> ; 155(2):JC1-3.
13.	Barnes PJ, Pocock SJ, Magnussen H, Iqbal A, Kramer B, Higgins M, et al. (2010): Integrating indacaterol dose selection in a clinical study in COPD using an adaptive seamless design. <i>Pulmonary Pharmacology & Therapeutics</i> ; 23(3):165-71.
14.	Donohue JF, Menjoge S, Kesten S (2003): Tolerance to bronchodilating effects of salmeterol in COPD. <i>Respiratory Medicine</i> ; 97(9):1014-20.
15.	Kurashima K, Hara K, Yoneda K, Kanauchi T, Kagiya N, Tokunaga D, et al. (2009): Changes in lung function and health status in patients with COPD treated with tiotropium or salmeterol plus fluticasone. <i>Respirology</i> ; 14(2):239-44.
Ausschlussgrund A5: Keine patientenrelevanten Endpunkte oder keine patientenrelevanten Endpunkte zum vergleichbaren Zeitpunkt (n=3)	
16.	Calverley PM, Stockley RA, Seemungal TA, Hagan G, Willits LR, Riley JH, et al. (2011): Reported pneumonia in patients with COPD: findings from the INSPIRE study. <i>Chest</i> ; 139(3):505-12.
17.	Niewoehner DE, Lokhnygina Y, Rice K, Kuschner WG, Sharafkhaneh A, Sarosi GA, et al. (2007): Risk indexes for exacerbations and hospitalizations due to COPD. <i>Chest</i> ; 131(1):20-8.
18.	Rutten-van Molken MP, Oostenbrink JB, Tashkin DP, Burkhart D, Monz BU (2006): Does quality of life of COPD patients as measured by the generic EuroQol five-dimension questionnaire differentiate between COPD severity stages? <i>Chest</i> ; 130(4):1117-28.
Ausschlussgrund A8: Mehrfachpublikationen ohne (für die Bewertung des Zusatznutzens) relevante Zusatzinformation (n=1)	
19.	Stanbrook MB (2009): Tiotropium reduced exacerbations but not rate of FEV 1 decline in patients with COPD using other respiratory medications. <i>Evidence-Based Medicine</i> ; 14(2): 42-3.
Ausschlussgrund A9: Publikation nicht in englischer oder deutscher Sprache (n=4)	
20.	Yin KS, Zhang DP, Shi Y, Sun LH, Min R, Xiao YL, et al. (2010): [A

	randomized, double-blind, placebo-control study of once-daily inhaled tiotropium in chronic obstructive pulmonary disease]. ZhonghuaJiehehehuxizazhi= ZhonghuaJiehehehuxizazhi= Chinese journal of tuberculosis and respiratory diseases; 33:519-23.
21.	Fang LZ, Liang X, Zhang JQ, Liu L, Fu WP, Zhao ZH, et al. (2008): [Combination of inhaled salmeterol/fluticasone and tiotropium in the treatment of chronic obstructive pulmonary disease: a randomised controlled trial]. ZhonghuaJiehehehuxizazhi= ZhonghuaJiehehehuxizazhi= Chinese journal of tuberculosis and respiratory diseases; 31:811-4.
22.	Sun LH, Tan Y, Qiao Y, Fang SR, Xie H, et al. (2007): Evaluation of clinical effect and safety of tiotropium bromide in treating stable chronic obstructive pulmonary disease. ZhongguoXinyaoyuLinchuangZazhi; 26:328-31.
23.	Wang HY, Xiao Y, Shi LL, Gong Q, Wen ZG (2007): Study on efficacy and safety of tiotropium powder in patients with stable COPD. Chinese Journal of New Drugs; 16:1119-22.

Anhang 4-D: Liste der abgebrochenen Studien

Benennen Sie in der nachfolgenden Tabelle die abgebrochenen Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4.7-1 (Anhang): Studienpool, abgebrochene Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben. b: Studie, für die der Unternehmer Sponsor war oder an der der Unternehmer anderweitig finanziell beteiligt war. c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

Es existieren keine abgebrochenen Studien mit dem zu bewertenden Arzneimittel.

Anhang 4-E: Liste der laufenden Studien

Benennen Sie in der nachfolgenden Tabelle die laufenden Studien aus dem durch die verschiedenen Suchschritte in Abschnitt 4.3.1.1 identifizierten Studienpool.

Tabelle 4.7-2 (Anhang): Studienpool, laufende Studien, zu bewertendes Arzneimittel

Studie	Studienkategorie			verfügbare Datenquellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
M/34273/40	nein	ja	nein	nein	ja [NCT01471171]	nein
DW_EKL001	nein	nein	ja	nein	ja [NCT01636401]	nein
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Datenquelle(n) mit anzugeben.</p> <p>b: Studie, für die der Unternehmer (Almirall, S.A.) Sponsor ist oder an der der Unternehmer anderweitig finanziell beteiligt ist.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

Anhang 4-F: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4.7-3 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4.7-3 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4.7-3 (Anhang): Studiendesign und -methodik für Studie <Studienbezeichnung>

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	
4b	Studienorganisation und Ort der Studiendurchführung	
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	

Item a	Charakteristikum	Studieninformation
8a	Methode zur Generierung der zufälligen Zuteilung	
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante	

Item a	Charakteristikum	Studieninformation
	Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	
14b	Informationen, warum die Studie endete oder beendet wurde	
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.

<< Angaben des pharmazeutischen Unternehmers >>

Tabelle 4.7-4 (Anhang): Studiendesign und -methodik für Studie M/34273/23

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<ul style="list-style-type: none"> • Evaluation der Wirksamkeit von zweimal täglich (BID) inhaliertem Acclidiniumbromid 400 µg, bei mittleren bis schweren COPD-Patienten • Bewertung der Sicherheit und Verträglichkeit der Mehrfachgabe (zweimal täglich - BID) von inhaliertem Acclidiniumbromid 400 µg, in einer Patientenpopulation • Mit Hilfe einer zweiseitigen Hypothese wird die Vergleichbarkeit der Behandlungen (Acclidiniumbromid, Tiotropiumbromid, Placebo) geprüft.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>multizentrische Mehrfachdosis-Studie, doppelblind, double-dummy, randomisiert, 3-Behandlungsperioden-cross-over, placebo- und aktivkontrolliert</p> <p>Zuteilungsverhältnis: 1:1:1:1:1:1</p> <p>5 Patienten erhielten in der ersten 15-tägigen Behandlungsperiode Acclidiniumbromid (A), in der zweiten Placebo (P) und in der dritten Tiotropiumbromid (T).</p> <p>5 weitere Patientengruppen á 5 Personen erhielten folgende 5 Behandlungssequenzen: ATP, TAP, TPA, PAT, PTA.</p> <p>Jeder der 30 Patienten (aufgeteilt in die beschriebenen 6 Gruppen mit unterschiedlichen Behandlungssequenzen) wurde entsprechend des Behandlungsschemas 15 Tage lang jeweils mit Acclidiniumbromid, Placebo und Tiotropiumbromid behandelt.</p> <p>Zwischen den 3 Behandlungsperioden erfolgte jeweils eine Auswaschungsphase von 9-15 Tagen.</p>
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p><u>Protokollergänzung 1</u>, 12.02.2009</p> <p>Berücksichtigung von Vorgaben und Empfehlungen der Ethikkommission bezüglich:</p> <ul style="list-style-type: none"> • Ausschluss von Patienten aus der Studie • Beendigung der Studie • Veröffentlichungspolitik <p><u>Protokollergänzung 2</u>, 23.03.2009</p> <p>Verlängerung der Auswaschungs-Periode von 6-9 auf 9-15 Tage, um das Auswaschen von Tiotropiumbromid, welches eine längere Halbwertszeit als Acclidiniumbromid besitzt, zu gewährleisten.</p> <p>Beide Protokollergänzungen wurden von der Ethikkommission angenommen.</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<ul style="list-style-type: none"> • männliche und weibliche Erwachsenen (≥40 Jahre) mit einer stabilen mittleren bis schweren COPD (nach GOLD 2006 Kriterien). • Raucher oder Ex-Raucher, die mindestens 10 Jahre lang eine Schachtel Zigaretten pro Tag geraucht haben (10 Packyears). • FEV1 ≥30% und <80% vom erwarteten Normalwert vor der

Item ^a	Charakteristikum	Studieninformation
		<p>Anwendung des Bronchodilatators Salbutamol und FEV1/FVC <70% nach Anwendung des Bronchodilatators Salbutamol.</p> <ul style="list-style-type: none"> • Keine Krankheitsgeschichte oder aktuelle Diagnose Asthma. • Keine Anzeichen einer Exazerbation in den letzten 6 Wochen vor Studienbeginn. • Keine Atemwegs- oder Herz-Kreislaufkrankungen oder Laborauffälligkeiten. • Keine Kontraindikation bzgl. Anticholinergika, wie z.B. eine bekannte symptomatische pro-statische Hypertrophie, Blasenhalsostruktion, akute Harnretention oder Engwinkel Glaukom.
4b	Studienorganisation und Ort der Studiendurchführung	<p>Die Studie wurde an zwei Studienorten (Berlin und Großhansdorf bei Hamburg) in Deutschland durchgeführt.</p> <p><u>Sponsor der Studie:</u> Laboratorios Almirall, S.A Research Centre Laureà Miró 408-410 08980 Sant Feliu de Llobregat Barcelona, Spanien</p> <p><u>Studienorganisation:</u> Early Phase Clinical Unit - Berlin PAREXEL International GmbH Verortet im Klinikum Westend, Haus 18 Spandauer Damm 130 14050 Berlin, Deutschland</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Name: Acclidiniumbromid Anwendung: Inhalation mit dem Genuair[®]-Multidosenpulverinhalator Dosierung: jeweils 1 Sprühstoß 400 µg Acclidiniumbromid am Morgen (08:00-10:00) und am Abend (20:00-22:00)</p> <p>Name: Placebo für Acclidiniumbromid Anwendung: Inhalation mit dem Genuair[®]-Multidosenpulverinhalator Dosierung: jeweils 1 Sprühstoß Placebo am Morgen (08:00-10:00) und am Abend (20:00-22:00)</p> <p>Name: Tiotropiumbromid Anwendung: Inhalation mit dem HandiHaler[®] Dosierung: 1 Sprühstoß 18 µg Tiotropiumbromid am Morgen (08:00-10:00)</p> <p>Name: Placebo für Tiotropiumbromid Anwendung: Inhalation mit dem HandiHaler[®] Dosierung: 1 Sprühstoß Placebo am Morgen (08:00-10:00)</p>
6	Zielkriterien	

Item ^a	Charakteristikum	Studieninformation
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p><i>Primäre Endpunkte:</i></p> <ul style="list-style-type: none"> • Veränderung im Vergleich zur Baseline-Erhebung (Change from baseline, CFB) der Kurve unter den FEV1-Werten über 12 Stunden ab der morgendlichen Inhalation (FEV1 AUC₀₋₁₂) am 15. (letzten) Tag der Behandlung <p><i>Sekundäre Endpunkte:</i></p> <ul style="list-style-type: none"> • CFB FEV1 AUC₀₋₁₂ am 1. Tag der Behandlung • CFB FVC AUC₀₋₁₂ am 1. und 15. Tag der Behandlung • CFB FEV1 und FVC AUC₀₋₂₄ am 1. und 15. Tag der Behandlung • CFB FEV1 und FVC AUC₁₂₋₂₄ am 1. und 15. Tag der Behandlung • CFB Ausgangs- (Trough) FEV1-Wert am Morgen und am Abend am 1. und 15. Tag der Behandlung • CFB bester (peak) FEV1- und FVC Wert am Morgen und am Abend am 1. und 15. Tag der Behandlung • CFB FEV1- und FVC Wert zu jeder spezifischen Tageszeit • CFB der begleitenden Medikation (Anzahl der Sprühstöße, durch den Patienten dokumentiert) nach ein und zwei Wochen der Behandlung • CFB COPD-Symptome nach ein und zwei Wochen der Behandlung <p><i>Sicherheit:</i></p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse (AEs) • Schwere unerwünschte Ereignisse (SAEs) <p><i>Zusätzliche Endpunkte:</i></p> <ul style="list-style-type: none"> • Handhabbarkeit des Inhalators • Dauer der Behandlung je Patient • Arzneimittel, die während der Run-in Phase eingenommen wurden, sowie Begleitmedikation • Anzahl der Studienabbrecher (%) und Abbruchgründe <p>Ein zentrales Spirometrieunternehmen (Cardinal Health GmbH) hat die Spirometer oder Pneumotachographen und sämtliche andere benötigte Ausrüstung (Computer, Kalibrationspumpe, Drucker, Papier, Tinte, etc.) und ein detailliertes Studienhandbuch an alle an dieser klinischen Studie beteiligten Standorte geliefert und hat dort auch Fortbildungen (zur Qualifizierung von für die Durchführung von Spirometrien verantwortlichen Technikern) durchgeführt. Diese Spirometer durften nur für diese spezifische klinische Studie verwendet werden. Die Spirometer haben FVC (maximales Volumen der ausgeatmeten Luft bei maximaler forcierter Expirationsanstrengung ausgehend von einer maximalen Inspiration) und FEV1 (in Litern ausgedrücktes Luftvolumen, das während der ersten Sekunde der Durchführung von FVC ausgeatmet wird) gemessen und erfüllten die Empfehlungen der American Thoracic Society (ATS) und der European Respiratory Society (ERS) in Bezug auf Genauigkeit und Präzision. Empfehlungen zu Instrumenten mussten befolgt werden, um genaue und vergleichbare spirometrische Daten sicherzustellen.</p>

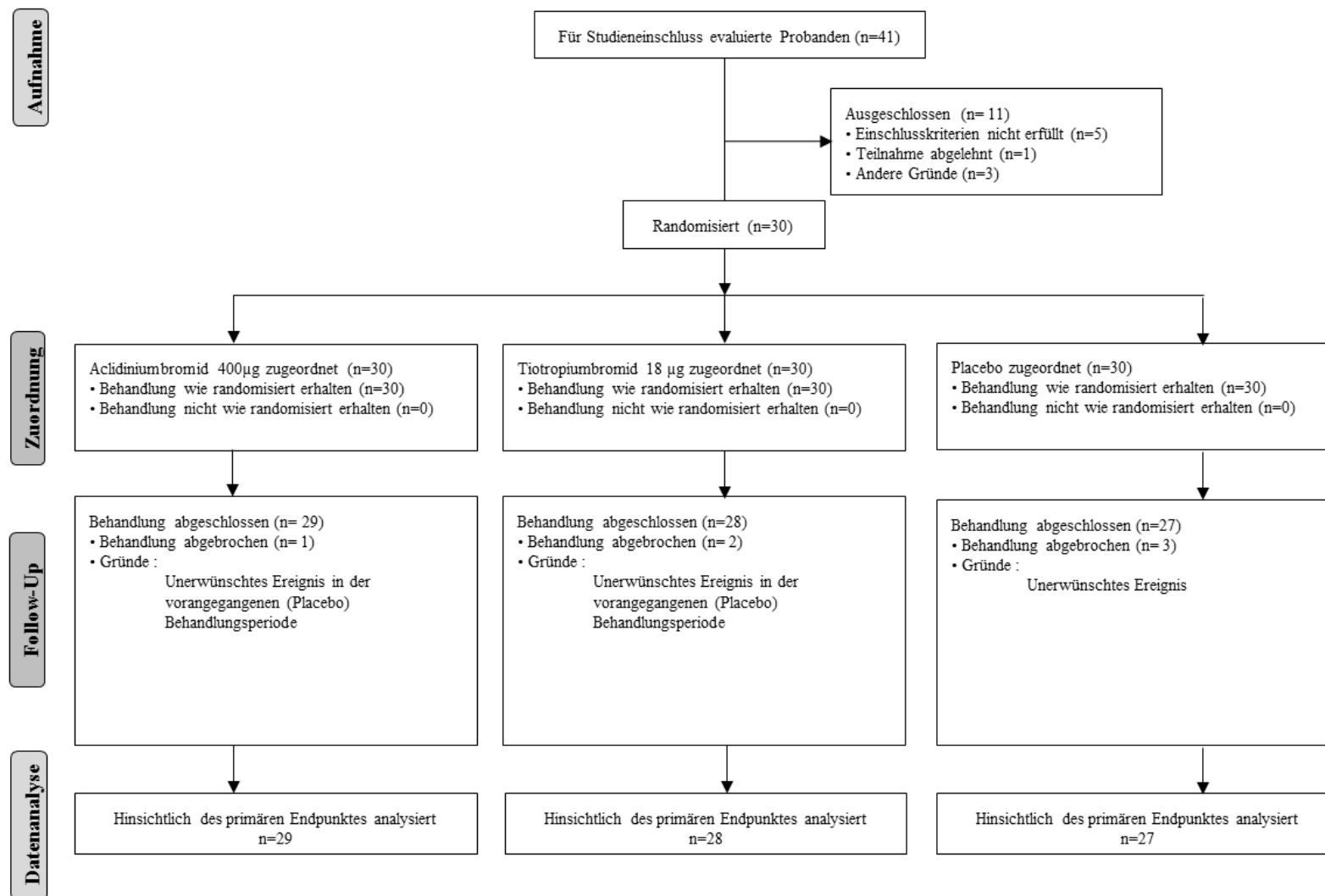
Item ^a	Charakteristikum	Studieninformation
		<p>An jedem Tag, an dem das System genutzt wurde und nach jeder Demontage, Reinigung und/oder Austausch des Sensors musste eine Kalibrierung durchgeführt werden.</p> <p>Spirometrien mussten bei Temperaturen zwischen 17 und 40°C durchgeführt werden. Es mussten Vorkehrungen getroffen werden, um alle Messungen bei in etwa gleicher Temperatur durchzuführen. Ein Notizbuch zur täglichen (zumindest an Tagen, an denen spirometrische Messungen durchgeführt wurden) Dokumentation der Umgebungstemperatur und des Drucks musste geführt werden.</p> <p>Im Falle von Temperatur- und/oder barometrischen Druckschwankungen innerhalb eines Tages musste die Kalibrierung vor der Durchführung jeglicher Spirometrie wiederholt werden.</p> <p>Lungenfunktionstests wurden von sehr erfahrenem Personal durchgeführt. Zu jeder spezifischen Tageszeit mussten in Übereinstimmung mit den Akzeptanz- und Wiederholbarkeitskriterien von ATS/ERS 3 technisch adäquate Messungen durchgeführt werden, die automatisch vom computergestützten Spirometer mittels vorprogrammierter Überprüfungen ausgewertet wurden. Diese Informationen wurden den Technikern am jeweiligen Studienort in Form von Mitteilungen angezeigt.</p> <p>Wenn die Akzeptanz- und Wiederholbarkeitskriterien eingehalten worden waren, konnte die Testsitzung abgeschlossen werden. Wenn eine oder beide Kriterien nicht erfüllt worden waren, wurden weitere Messungen durchgeführt bis entweder beide Kriterien durch eine Analyse von zusätzlichen, akzeptablen Spirogrammen erfüllt waren ODER bis insgesamt 8 Tests durchgeführt waren, außer der Patient war hierzu nicht in der Lage.</p> <p>Obwohl nur die höchsten Werte für FEV1 und FVC in der Analyse verwendet wurden, wurden alle Manöver gespeichert. Diese Daten wurden durch den Forscher elektronisch an das zentrale Datenzentrum des zentralen Spirometrieunternehmens übermittelt. Die Daten wurden überprüft und regelmäßig an die für die gesamten Studiendaten verantwortliche Datenverarbeitungsgruppe übermittelt.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Die folgenden Sekundärvariablen wurden zusätzlich zu den zuvor beschriebenen definiert:</p> <ul style="list-style-type: none"> • Absolute Werte von FEV1 und FVC zu jeder spezifischen Tageszeit an den Tagen 1 und 15 der Behandlung • CFB FEV1 AUC0-12 am Tag 15 der Behandlung • CFB FEV1 AUC12-24 am Tag 15 der Behandlung • CFB des morgendlichem und abendlichem Vordosis (trough)-FEV1 am Tag 15 der Behandlung
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Da es sich hier um eine explorative klinische Studie in der Phase IIa handelt, wurde die Anzahl von 24 Probanden (4 pro Behandlungssequenz) zum Erreichen der Studienziele als ausreichend erachtet.</p> <p>Unter Berücksichtigung einer Abbruchrate von 10% wurden insgesamt 30 Patienten (5 Patienten pro Behandlungssequenz)</p>

Item ^a	Charakteristikum	Studieninformation
		randomisiert.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	NA
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Computergenerierter Randomisierungsplan unter Verwendung der Patienten-Identifikationsnummer.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Während Visite 1 wurden die Patienten mittels eines ausgeglichenen Randomisierungsverhältnisses (1:1:1:1:1) einer der 6 Behandlungssequenzen zugeordnet
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Bei der Unterschrift der Einwilligungserklärung ordnete das Forschungspersonal jedem Patienten eine Patienten-Identifikationsnummer zu, was den Anfang der Untersuchungsphase markierte. Dem Patienten zugeteilte Nummer setzte sich aus zwei Teilen zusammen: <ul style="list-style-type: none"> • Die ersten 4 Ziffern (fest) repräsentierten die Forschernummer, wie sie jedem Forscher von Almirall zugeordnet worden war • Die nächsten 2 Ziffern (aufsteigend) wurden sequentiell innerhalb jedes Studienortes bei Bestätigung der Teilnahme der Patienten an der Studie durch die Unterschrift auf der Einwilligungserklärung, ausgehend von 01, vergeben. Die Patienten-Identifizierungsnummer wurde verwendet, um die Patienten im Laufe der Studie zu identifizieren und wurde auf jeder Seite des Prüfbogens (Case Report Form, CRF) festgehalten. Randomisierungsdaten wurden streng vertraulich behandelt und waren bis zur Offenlegung der zugeordneten Behandlung aller Versuchspatienten nach der Schließung der Datenbank beim Abschluss der Studie nur berechtigten Personen zugänglich. Erst nach dem Abschluss der Studie und nach der Überprüfung und dem Abschluss der Datensammlung wurden die Randomisierungs-codes zur Datenanalyse zugänglich gemacht.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Vor Beginn des Versuchs wurde ein computergenerierter Randomisierungsplan erstellt, um es dem Statistikteam bei Almirall, S.A. zu ermöglichen, gemäß dem zutreffenden Standardverfahren einer Randomisierungsnummer eine Behandlungssequenz zuzuordnen. Die Randomisierung wurde durchgeführt, um jegliche mögliche Verzerrung aufgrund der Reihenfolge der Interventionszuteilung zu vermeiden. Die Blockgröße wurde in Übereinstimmung mit dem Leiter der klinischen Studie und dem Statistiker festgelegt und wurde den Forschern nicht mitgeteilt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die	a) und b) trifft zu Da Randomisierungsdaten wurden streng vertraulich behandelt und waren bis zur Offenlegung der zugeordneten Behandlung aller Versuchspatienten nach der Schließung der Datenbank beim Abschluss der Studie nur berechtigten Personen zugänglich. Erst nach dem Abschluss der Studie und nach der Überprüfung und dem Abschluss der Datensammlung wurden die

Item ^a	Charakteristikum	Studieninformation
	Verblindung vorgenommen?	Randomisierungscode zur Datenanalyse zugänglich gemacht.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Um Verzerrungen zu minimieren, wurde ein Double-Dummy-Verfahren zur Sicherstellung der doppelblinden Durchführung des Versuchs angewandt. Zusätzlich wiesen die passenden Placebos zu Acclidiniumbromid und Tiotropium, abgesehen vom aktiven Wirkstoff, die gleiche äußere Erscheinungsform und die gleiche inhaltliche Zusammenstellung auf.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Alle statistischen Vergleiche waren zweiseitige Hypothesentests und das Signifikanzniveau wurde bei 0.05 festgelegt. Alle Konfidenzintervalle (CIs) lagen zweiseitig bei einem Konfidenzniveau von 95%.</p> <p>Die primären und sekundären Endpunkte wurden mit der Analyse des Kovarianzmodells (ANCOVA) für Crossover-Designs (vgl. CSR M/34273/23, Seite 61-62) ausgewertet.</p> <p>Die Variation zwischen Patienten wurde als zufällig interpretiert und die Reihenfolgeeffekte wurden mittels Mittelwerten zwischen den Patienten in Reihenfolge im obigen ANCOVA-Modell getestet. Zeiträume und Behandlungseffekte wurden unter Verwendung der Fehlervariablen getestet. Die Behandlungseffekte wurden mittels der kleinsten Quadratmittelwerte (LS, least square) mit einer Standardabweichung (SE, standard error) und 95% KI aus dem obigen ANCOVA-Modell näherungsweise ermittelt. Die Vergleiche der Behandlungen wurden auf der Basis von Unterschieden der Behandlungsfaktoren durchgeführt. Der Unterschied zwischen den Behandlungen wurde durch den Unterschied zwischen bereinigten Mittelwerten (LS-Mittelwerte), deren SE und 95% KI näherungsweise ermittelt. Da es sich hier um eine explorative klinische Studie handelt, wurde der Fehler erster Art nicht für mehrfache Behandlungsvergleiche angeglichen.</p> <p>Die entsprechenden Kovarianzen für jedes Modell wurden im statistischen Analyseplan (SAP, Statistical Analysis Plan) beschrieben.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt	<ul style="list-style-type: none"> • Randomisiert: Acclidiniumbromid 400 µg BID (30), Tiotropiumbromid 18 µg OD (30), Placebo BID (30) • Geplante Behandlung erhalten: Acclidiniumbromid 400 µg BID (26), Tiotropiumbromid 18 µg OD (25), Placebo BID (24) • Für Primärergebnis analysiert: Acclidiniumbromid 400 µg BID

Item ^a	Charakteristikum	Studieninformation
	wurden	(29), Tiotropiumbromid 18 µg OD (28), Placebo BID (27)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Von der Gesamtpopulation von 30 Patienten wurden 3 Patienten aufgrund von AEs von der Studie ausgeschlossen. Alle drei Patienten befanden sich zum Zeitpunkt des Auftretens des AEs in der Behandlungsperiode mit Placebo. Der erste Patient befand sich in der letzten (TAP), der zweite in der ersten (PAT) und der dritte in der zweiten (APT) Behandlungsperiode. Danach fehlen in der ITT-Population drei Placebo-Patienten, zwei Tiotropiumbromid-Patienten und ein Acclidiniumbromid-Patient.</p> <p>Acclidiniumbromid 400 µg BID: 1; unerwünschtes Ereignis (in der vorangegangenen Behandlungsperiode): 1</p> <p>Tiotropiumbromid 18 µg OD: 2; unerwünschtes Ereignis (in der vorangegangenen Behandlungsperiode): 2</p> <p>Placebo BID: 3; unerwünschtes Ereignis: 3</p>
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	<p>Startdatum der Studie (erste Untersuchung): 09. März 2009</p> <p>Abschlussdatum der Studie (letzter Besuch des letzten Patienten): 01. Juli 2009</p>
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar.



Hinweis: Bei dem vorliegenden Studiendesign handelt es sich um ein Cross-over-Design. Der Patientenfluss ist für die im CONSORT-Statement geforderten ausgewerteten Therapiearme, nicht die Randomisierung in die 6 Behandlungssequenzen, dargestellt.

Tabelle 4.7-5 (Anhang): Studiendesign und -methodik für Studie M/34273/39

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<p>1. Evaluierung der Wirksamkeit von inhaliertem Acclidiniumbromid 400 µg BID innerhalb 24h gegenüber Placebo bei Patienten mit moderater bis schwerer Ausprägung von COPD</p> <p>2. Evaluierung der Bronchodilatation von inhaliertem Acclidiniumbromid 400 µg BID in der Nacht gegenüber Tiotropiumbromid bei Patienten mit moderater bis schwerer Ausprägung von COPD</p> <p>3. Bewertung der Sicherheit und Verträglichkeit von inhaliertem Acclidiniumbromid 400 µg BID in der gleichen Zielpopulation.</p> <p>Mit Hilfe einer zweiseitigen Hypothese wird die Vergleichbarkeit der Behandlungen (Acclidiniumbromid, Tiotropiumbromid, Placebo) geprüft.</p>
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>multizentrische Mehrfachdosis-Studie, doppelblind, double-dummy, randomisiert, parallel, placebo- und aktivkontrolliert, multizentrische klinische Studie.</p> <p>Zuteilungsverhältnis: 2:2:1 (Acclidiniumbromid, Tiotropiumbromid, Placebo)</p>
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Nach Studienbeginn wurde keine grundsätzliche Änderung der im Protokoll beschriebenen Methodik vorgenommen. Im statistischen Analyseplan (SAP) wurde lediglich ergänzt, Sensitivitätsanalysen der primären und sekundären Variablen vorzunehmen. Dabei flossen Alter und Baselinewert des FEV₁ - Wertes als Kovariate und Behandlungsarm, Geschlecht, Woche und die Interaktion zwischen Behandlungsarm und Woche als Faktoren in das „mixed model for repeated measures“ (MMRM) ein. Außerdem wurden im SAP die Ableitung von Variablen und der Umgang mit fehlenden Werten spezifiziert</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<ul style="list-style-type: none"> • männliche und weibliche Erwachsenen (≥40 Jahre) mit einer stabilen mittleren bis schweren COPD (nach GOLD 2006 Kriterien). • Raucher oder Ex-Raucher, die mindestens 10 Jahre lang eine Schachtel Zigaretten pro Tag geraucht haben (10 Packyears). • FEV₁ ≥30% und <80% vom erwarteten Normalwert vor der Anwendung des Bronchodilatators Salbutamol und FEV₁/FVC <70% nach Anwendung des Bronchodilatators Salbutamol. • Keine Krankheitsgeschichte oder aktuelle Diagnose Asthma. • Keine Anzeichen einer Exazerbation in den letzten 6 Wochen vor Studienbeginn. • Keine Atemwegs- oder Herz-Kreislaufkrankungen oder Laborauffälligkeiten. • Keine Kontraindikation bzgl. Anticholinergika, wie z.B. eine bekannte symptomatische pro-statische Hypertrophie, Blasenhalsostruktion, akute Harnretention oder Engwinkel Glaukom.
4b	Studienorganisation und Ort der	Die Studie wurde an 49 Studienorten in Deutschland (20), Polen

Item ^a	Charakteristikum	Studieninformation
	Studiendurchführung	<p>(13), Tschechische Republik (3) und Ungarn (5) durchgeführt.</p> <p><u>Sponsor der Studie:</u> Laboratorios Almirall, S.A Research Centre Laureà Miró 408-410 08980 Sant Feliu de Llobregat Barcelona, Spanien</p> <p><u>Studienorganisation:</u> Institut für Atemwegsforschung GmbH (INSAF) Biebricher Allee 34 66187 Wiesbaden, Deutschland</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Name: Acridiniumbromid Anwendung: Inhalation mit dem Genuair[®]-Multidosenpulverinhalator Dosierung: jeweils 1 Sprühstoß 400 µg Acridiniumbromid am Morgen (08:00-10:00) und am Abend (20:00-22:00)</p> <p>Name: Placebo für Acridiniumbromid Anwendung: Inhalation mit dem Genuair[®]-Multidosenpulverinhalator Dosierung: jeweils 1 Sprühstoß Placebo am Morgen (08:00-10:00) und am Abend (20:00-22:00)</p> <p>Name: Tiotropiumbromid Anwendung: Inhalation mit dem HandiHaler[®] Dosierung: 1 Sprühstoß 18 µg Tiotropiumbromid am Morgen (08:00-10:00)</p> <p>Name: Placebo für Tiotropiumbromid Anwendung: Inhalation mit dem HandiHaler[®] Dosierung: 1 Sprühstoß Placebo am Morgen (08:00-10:00)</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p><i>Primärer Endpunkt:</i> • Veränderung im Vergleich zur Baseline-Erhebung (Change from baseline, CFB) der Kurve unter den FEV1-Werten über 24 Stunden ab der morgendlichen Inhalation (FEV1 AUC₀₋₂₄) am letzten Tag der Behandlung (nach 6 Wochen)</p> <p><i>Sekundärer Endpunkt:</i> • CFB FEV1 AUC₁₂₋₂₄ am 1. Tag der Behandlung</p> <p><i>Zusätzliche Endpunkte:</i> • CFB FVC und FEV1 AUC₀₋₁₂, AUC₀₋₂₄, AUC₁₂₋₂₄, am 1. Tag der Behandlung • CFB FVC und FEV1 AUC₀₋₁₂, FVC AUC₀₋₂₄ und FVC AUC₀₋₂₄ am letzten Tag der Behandlung • CFB Ausgangs- (Trough) FEV1- und FVC- Wert am Morgen und am Abend am 1. und letzten Tag der Behandlung</p>

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • CFB bester (peak) FEV1- und FVC Wert am Morgen und am Abend am 1. und letzten Tag der Behandlung • Zeit bis zum morgendlichen FEV1 am 1. und letzten Tag • CFB FEV1- und FVC Wert zu jeder spezifischen Tageszeit am 1. und letzten Tag • CFB der begleitenden Medikation nach sechs Wochen der Behandlung • CFB des Anteils der Tage ohne begleitende Medikation an der gesamten Studiendauer • CFB des Anteils der Tage ohne COPD-Symptome an der gesamten Studiendauer • Anteil der Patienten, die einen der beiden Inhalatoren bevorzugen, und Anteil der Patienten, die den Inhalator weiterhin verwenden wollen <p><i>Sicherheit:</i></p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse (AEs) • Schwere unerwünschte Ereignisse (SAEs) • Blutdruck, Herzfrequenz <p>Ein zentrales Spirometrieunternehmen hat die Spirometer (inklusive Pneumotachograph) und sämtliche benötigte Ausrüstung (Computer, Kalibrationspumpe, Drucker, Papier, Tinte, etc.) und ein detailliertes Studienhandbuch an alle an dieser klinischen Studie beteiligten Standorte geliefert und hat dort auch Fortbildungen (zur Qualifizierung von für die Durchführung von Spirometrie verantwortlichen Techniken) durchgeführt. Diese Spirometer dürfen nur für diese spezifische klinische Studie verwendet werden. Die Spirometer haben FVC (maximales Volumen der ausgeatmeten Luft bei maximaler forcierter Expirationsanstrengung aus einer Position der maximalen Inspiration) und FEV1 (in Litern ausgedrücktes Luftvolumen, das während der ersten Sekunde der Durchführung von FVC ausgeatmet wird) gemessen und erfüllten die Empfehlungen der American Thoracic Society (ATS) und der European Respiratory Society (ERS) in Bezug auf Genauigkeit und Präzision.</p> <p>Empfehlungen zu Instrumenten mussten befolgt werden, um genaue und vergleichbare spirometrische Daten sicherzustellen. An jedem Tag, an dem das System genutzt wird und nach jeder Demontage, Reinigung und/oder Austausch des Pneumotachographen muss eine Kalibrierung durchgeführt werden.</p> <p>Spirometrie musste bei Temperaturen zwischen 17 und 40°C durchgeführt werden. Es mussten Vorkehrungen getroffen werden, um alle Manöver bei in etwa gleicher Temperatur durchzuführen. Ein Notizbuch zur täglichen (zumindest an Tagen, an denen spirometrische Vorgänge durchgeführt wurden) Dokumentation der Umgebungstemperatur und des Drucks musste geführt werden.</p> <p>Im Falle von Temperatur- und/oder barometrischen Druckschwankungen innerhalb eines Tages musste die</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Kalibrierung vor der Durchführung jeglicher Spirometrie wiederholt werden.</p> <p>Lungenfunktionstests wurden von sehr erfahrenem Personal durchgeführt. Zu jeder spezifischen Tageszeit mussten in Übereinstimmung mit den Akzeptanz- und Wiederholbarkeitskriterien von ATS/ERS 3 technisch adäquate Messungen durchgeführt werden, die automatisch vom computerisierten Spirometer mittels vorprogrammierten Überprüfungen ausgewertet wurden. Diese Information wurde den Technikern am jeweiligen Studienort in Form von Mitteilungen angezeigt, sodass sie weiterhin die notwendigen Versuche durchführten.</p> <p>Wenn die Akzeptanz- und Wiederholbarkeitskriterien eingehalten worden waren, konnte die Testsitzung abgeschlossen werden. Wenn eine oder beide Kriterien nicht erfüllt worden waren, wurden von der Ausrüstung zusätzliche Manöver angefordert bis entweder beide Kriterien durch eine Analyse von zusätzlichen, akzeptablen Spirogrammen erfüllt waren ODER bis insgesamt 8 Tests durchgeführt waren, außer der Patient war hierzu nicht in der Lage.</p> <p>Der Schweregrad von COPD-Symptomen wurde mit E-RS (EXACT-Respiratorische Symptome) evaluiert. Die Patienten beantwortete die Fragen des E-RS jeden Abend vor dem Zubettgehen in einem elektronischen Tagebuch als Teil von EXACT. E-RS wurde in allen lokalen Sprachen unter Verwendung von zertifizierten und lizenzierten Übersetzungen zur Verfügung gestellt.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Zusätzlich zu denen im Studienprotokoll definierten Endpunkt „Anteil der Patienten, den Inhalator weiterhin verwenden wollen“ wurde im statistischen Analyseplan (SAP) der Endpunkt „Anteil der Patienten, die einen der beiden Inhalatoren bevorzugen“ definiert.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Ein Stichprobengröße von 385 COPD-Patienten (77 und 154 Patienten in der Placebogruppe und in jeder der aktiven Behandlungsgruppen) liefert mindestens eine 99-prozentige Wahrscheinlichkeit, einen Unterschied von 130 ml gegenüber dem Ausgangswert in der normalisierten FEV1-Fläche unter der Kurve in einem 24h-Zeitfenster gleich nach der morgendlichen Arzneimittelgabe (AUC0-24) nach 6 Wochen Behandlung zwischen der Gruppe mit Acridiniumbromid 400 µg und der Placebogruppe mit einem zweiseitigen Signifikanzniveau unter der Annahme einer allgemeinen Standardabweichung von 210 ml zu ermitteln. Auch besteht bei diesem Probandenumfang eine 99-prozentige Wahrscheinlichkeit, einen Unterschied zwischen Acridiniumbromid 400 µg BID und Placebo gegenüber dem Ausgangswert des normalisierten FEV1 AUC12-24 nach 6 Wochen Behandlung zu finden, der 130 mL entspricht (unter der Annahme, dass SD = 210 ml).</p> <p>FEV1 AUC12-24 nach 6 Wochen Behandlung entspricht 130 ml (unter der Annahme, dass SD = 210 ml).</p> <p>Mit insgesamt 154 Patienten in jeder der aktiven</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Behandlungsgruppen besteht eine 83-prozentige Wahrscheinlichkeit, einen Unterschied zwischen Acclidiniumbromid 400 µg BID und Tiotropiumbromid gegenüber dem Ausgangswert des normalisierten FEV1 AUC12-24 nach 6 Wochen Behandlung zu ermitteln, der 70 mL entspricht (unter Annahme einer allgemeinen Standardabweichung von 210 ml). Annahmen bezüglich SD von AUC0-24 und AUC12-24 wurden von vorherigen Studien mit Acclidinium übernommen.</p> <p>Unter Berücksichtigung einer Abbruchrate von 5% wurde geplant insgesamt 405 Patienten (162 mit Acclidiniumbromid 400 µg, 162 mit Tiotropiumbromid 18 µg und 81 mit Placebo) zu randomisieren.</p> <p>Zusätzlich wurde angenommen, dass 35% der für den Studieneinschluss evaluierten Patienten nicht randomisiert werden, was zu 625 evaluierten Patienten führt.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Interaktives Voice Response System (IVRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Bei Visite 1 wurden Patienten, die die Ein-/Ausschlusskriterien erfüllten, von IVRS einer der 3 Behandlungssequenzen (Acclidiniumbromid, Tiotropiumbromid, Placebo) unter Verwendung eines Randomisierungsverhältnisses von (2:2:1) zugewiesen
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	<p>Bei der Untersuchungsvisite (Visite -1) rief das Forschungspersonal IVRS auf, um jedem Patienten eine Patienten-Identifikationsnummer zuzuweisen. Somit begann der Untersuchungszeitraum. Diese Nummer setzte sich aus zwei Teilen zusammen:</p> <ul style="list-style-type: none"> • Die ersten 4 Ziffern (fest) repräsentierten die Forschernummer, wie sie jedem Forscher von Almirall zugeordnet worden war. • Die nächsten 2 Ziffern (aufsteigend) werden, beginnend mit 01, sequentiell und automatisch von IVRS in jedem Zentrum bei der Aufnahme der Patienten während der Untersuchungsvisite vergeben. <p>Die Patienten-Identifizierungsnummer wurde verwendet, um die Patienten im Laufe der Studie zu identifizieren.</p> <p>Die Forscher wurden mit den nötigen Mitteln ausgestattet, um im Notfall die Offenlegung durch das IVRS durchführen zu können. Die Pharmakovigilanzereinheit des Auftragsforschungsinstituts und der Sponsor wurden umgehend über die Offenlegung der Behandlung informiert. Das Brechen des Verhaltenskodex am Forschungsstandort hätte den Patienten umgehend von der weiteren Teilnahme an der Studie disqualifiziert.</p>
10	Randomisierung, Durchführung Wer führte die Zuteilung durch,	Vor Beginn der Studie wurde ein computergenerierter Randomisierungsplan erstellt, um es dem Statistik-/Programmierungsteam bei Almirall zu ermöglichen, gemäß dem

Item ^a	Charakteristikum	Studieninformation
	wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	zutreffenden Standardverfahren einer Randomisierungsnummer eine Behandlungssequenz zuzuordnen. Die Randomisierung wurde durchgeführt, um jegliche mögliche Verzerrung aufgrund der Reihenfolge der Interventionszuteilung zu vermeiden. Die Blockgröße wurde in Übereinstimmung mit dem Leiter der klinischen Studie und dem Statistiker festgelegt und wurde den Forschern nicht mitgeteilt.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a.) und b.) treffen zu Almirall, S.A., Forest, Studienmitarbeiter und Studienteilnehmer wurden verblindet, so dass nicht bekannt wurde, ob der Patient Acridiniumbromid oder Placebo erhielt.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Um Verzerrungen zu minimieren, wurde ein Double-Dummy-Verfahren zur Sicherstellung der doppelblinden Durchführung des Versuchs angewandt. Zusätzlich wiesen die passenden Placebos zu Acridiniumbromid und Tiotropiumbromid, abgesehen vom aktiven Wirkstoff, die gleiche äußere Erscheinungsform und die gleiche inhaltliche Zusammenstellung auf.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Alle statistischen Vergleiche waren zweiseitige Hypothesentests und das Signifikanzniveau wurde auf 0.05 festgelegt . Alle Konfidenzintervalle (KIs) lagen zweiseitig bei einem Konfidenzniveau von 95%. Alle Wirksamkeitsvariablen mit Ausnahme der Zeit zum FEV1-Maximum (peak), dem Prozentsatz von Patienten, die eins der beiden Geräte bevorzugen und dem Prozentsatz von Patienten, die mit dem Gerät weitermachen wollten, wurden über die Kovarianz (ANCOVA) analysiert, wobei Behandlung und Geschlecht Faktoren und der entsprechende Ausgangswert sowie das Alter die Kovarianzen waren. Alle Vergleiche zwischen den Gruppen wurden mit dem geeigneten Kontrast im ANCOVA-Modell überprüft. Für alle paarweisen Vergleiche zwischen den Gruppen wurden LS-Mittelwerte und 95% Konfidenzintervalle vorgegeben. Im Falle extremer Abweichungen von angenommenen Normalwerten oder Vorhandensein von Ausreißern in den Daten, wurden die Variablen "Änderung vom Ausgangswert zu Tagen ohne Medikation in Prozent", "Änderung vom Ausgangswert zu Tagen ohne COPD-Symptome in Prozent" und "Änderung vom Ausgangswert zur durchschnittlichen Anwendung von Medikation" mit einem normalverteilten ANCOVA-Modell analysiert, wobei Behandlung und Geschlecht Faktoren und normalverteilte Ausgangswerte sowie das Alter die Kovarianzen waren. Für die paarweisen Vergleiche wurden zweiseitige Konfidenzintervalle (bei 95% Konfidenzniveau) vorgegeben. Der Zeitraum zum FEV1-Maximum (peak) wurde beschreibend

Item ^a	Charakteristikum	Studieninformation
		<p>analysiert.</p> <p>Der Anteil von Patienten, die eines der beiden Inhalationsgeräte bevorzugen und der Anteil von Patienten, die mit dem Inhalationsgerät weitermachen wollten, wurden näherungsweise ermittelt. Das zugehörige 95% KI wird für die Behandlungsgruppen und für alle Patienten gesamt dargestellt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Der primäre Vergleich der Primärvariablen fand zwischen den Gruppen Acclidiniumbromid 400 µg BID und Placebo statt.</p> <p>Wenn der obige Vergleich auf dem Niveau 0.05 signifikant war, wurde ein Test der sekundären Wirksamkeitsvariablen (Änderung gegenüber dem Ausgangswert des normalisierten FEV1 AUC12-24 nach 6 Wochen Behandlung) durchgeführt.</p> <p>Der erste Vergleich dieser Variablen fand zwischen den Gruppen mit Acclidiniumbromid 400 µg BID und Placebo statt. Wenn dieser Vergleich auf dem Niveau 0.05 signifikant war, wurde ein Vergleich zwischen Acclidiniumbromid 400 µg BID und Tiotropiumbromid durchgeführt.</p> <p>Alle Vergleichsgruppen (Acclidiniumbromid 400 µg BID vs. Placebo, Acclidiniumbromid 400 µg BID vs. Tiotropiumbromid, Tiotropiumbromid vs. Placebo) für alle anderen Endpunkte wurden durchgeführt, jedoch als sekundär angesehen, weshalb keine Anpassung für multiples Testen vorgenommen wurde.</p>
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Randomisiert: Acclidiniumbromid 400 µg BID (171), Tiotropiumbromid 18 µg OD (158), Placebo BID (85) • Beabsichtigte Behandlung erhalten: Acclidiniumbromid 400 µg BID (162), Tiotropiumbromid 18 µg OD (149), Placebo BID (80) • Für Primärergebnis analysiert: Acclidiniumbromid 400 µg BID (171), Tiotropiumbromid 18 µg OD (158), Placebo BID (85) <p>Flussdiagramm (siehe unten).</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • Acclidiniumbromid 400 µg BID: 5; unerwünschter Zwischenfall:3 ; Patientenwunsch:2 • Tiotropiumbromid 18 µg BID:4; unerwünschter Zwischenfall:3; Patientenwunsch:1 • Placebo BID:5; unerwünschter Zwischenfall:4 ; unzureichende Wirksamkeit:1
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	<p>Startdatum der Studie (erste Untersuchung): 17. Oktober 2011</p> <p>Abschlussdatum der Studie (letzter Besuch des letzten Patienten):</p>

Item ^a	Charakteristikum	Studieninformation
		14.März 2012
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		

Stellen Sie den Patientenfluss in einem Flow-Chart je Studie dar

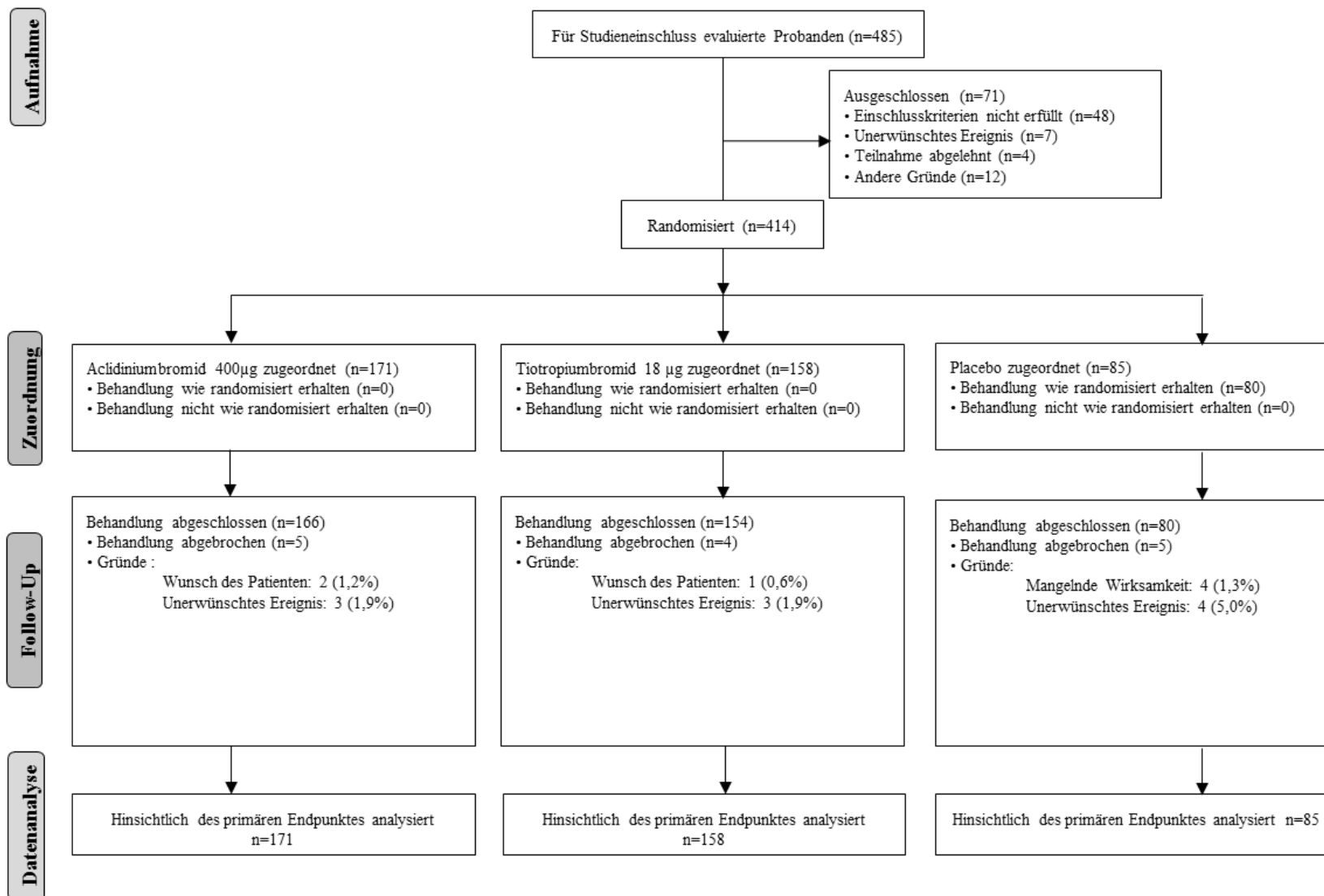


Tabelle 4.7-6 (Anhang): Studiendesign und -methodik für Studie ATTAIN

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genaue Ziele, Fragestellung und Hypothesen	<ul style="list-style-type: none"> • Beurteilung der langfristigen bronchodilatatorischen Wirksamkeit von zweimal täglich (BID) 200 µg und 400 µg inhalativem Aclidiniumbromid bei COPD-Patienten im Vergleich mit Placebo. • Beurteilung der Vorteile von 200 µg und 400 µg Aclidiniumbromid BID im Vergleich mit Placebo, bei krankheitsbezogenem Gesundheitsstatus, COPD-Symptomatik und COPD-Exazerbationen. • Beurteilung der langfristigen Sicherheit und Verträglichkeit von 200 µg und 400 µg inhalativem Aclidiniumbromid BID im Vergleich mit Placebo in einer identischen Zielpopulation.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Parallelgruppe, 200 µg Aclidiniumbromid BID, 400 µg Aclidiniumbromid BID oder Placebo BID in einem Randomisierungsverhältnis von 1:1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Speziell für Frankreich wurde am 28. Juli 2010 eine Änderung vorgenommen. Es wurden die folgenden Einschlusskriterien ergänzt, die für französische Patienten gelten: <ul style="list-style-type: none"> • Patienten, die nicht in der Lage sind, ihre Einwilligung abzugeben, Patienten im zustimmungspflichtigen Alter unter Vormundschaft oder Patienten mit schwacher Konstitution.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<ul style="list-style-type: none"> • Erwachsene männliche oder weibliche Patienten ab 40 Jahren mit stabiler mittelschwerer bis schwerer COPD (Definition gemäß GOLD-Richtlinien [Global Initiative for Chronic Obstructive Lung Disease]) und stabiler Atemwegsobstruktion. • Aktuelle oder Ex-Raucher mit ≥ 10 Packungsjahren. • Forciertes Expirationsvolumen in einer Sekunde (Einsekundenkapazität, FEV₁) 10 bis 15 Minuten nach Verabreichung von Salbutamol ≥ 30 % und < 80 % des prognostizierten Normalwertes und FEV₁/forcierte Vitalkapazität (FVC) < 70 %. • Keine frühere oder aktuelle Asthma-Diagnose. • Keine Anzeichen für COPD-Exazerbationen innerhalb von sechs Wochen vor dem Screening. • Keine Hospitalisierung wegen COPD-Exazerbationen innerhalb von drei Monaten vor dem Screening. • Keine klinisch signifikanten respiratorischen und/oder kardiovaskulären Erkrankungen oder Laborwertabweichungen. • Keine Überempfindlichkeit oder Kontraindikation ggü. anticholinergen Arzneimitteln, wie bekannte symptomatische

Item ^a	Charakteristikum	Studieninformation
		Prostatahypertrophie, Blasenhalsobstruktion, akute Harnretention oder Engwinkelglaukom. • Keine Teilnahme an früheren Acridiniumbromid-Studien.
4b	Studienorganisation und Ort der Studiendurchführung	Diese Studie wurde in 103 Prüfzentren durchgeführt (von denen 100 ihre Patienten randomisierten): zehn Zentren in der Tschechischen Republik, fünf Zentren in Frankreich, 17 Zentren in Deutschland, 13 Zentren in Ungarn, drei Zentren in Italien, ein Zentrum in Peru, 21 Zentren in Polen, zehn Zentren in der Russischen Föderation, 5 Zentren in Spanien, 13 Zentren in Südafrika und 5 Zentren in der Ukraine. Diese Studie wurde finanziert von Almirall S.A., Barcelona, Spanien und Forest Laboratories, Inc, NY, USA. Leiter der klinischen Studie: Rosa Segarra Ctra. Laureà Miró, 408-410 08980, Sant Feliu de Llobregat Barcelona, Spanien
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Name: Acridiniumbromid; Verabreichungsweg: Orale Inhalation mithilfe des Genuair® Mehrfachdosis-Trockenpulver-Inhalators; Dosierungsform: Inhalationspulver-Dosis und Regime: 1 Sprühstoß mit 200 µg oder 400 µg am Morgen (08:00 – 10:00) und am Abend (20:00 – 22:00). Name: Placebo im Vergleich mit Acridiniumbromid; Verabreichungsweg: Orale Inhalation mithilfe des Genuair® Mehrfachdosis-Trockenpulver-Inhalators; Dosierungsform: Inhalationspulver-Dosis und Regime: 1 Sprühstoß Placebo am Morgen (08:00 – 10:00) und am Abend (20:00 – 22:00).
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primärer Wirksamkeitsendpunkt: • Veränderung gegenüber des Baseline-Trough-FEV1-Wertes am Morgen vor Verabreichung der Dosis in Woche 24 für den Zulassungsantrag in der Europäischen Union (EU) und in Woche 12 für den Antrag in den USA. Sekundäre Wirksamkeitsendpunkte: • Veränderung gegenüber der Baseline beim Peak-FEV1 in Woche 24 für den Zulassungsantrag in der EU und in Woche 12 für den Antrag in den USA. • Anteil (%) der Patienten mit klinisch relevanter Verbesserung (≥ 1 Einheit) beim medianen Wert des Transitional Dyspnoea Index (TDI) in Woche 24. • Anteil (%) der Patienten mit klinisch relevanter Verbesserung (≥ 4 Einheiten) im Vergleich zur Baseline bei der Gesamtpunktzahl des Saint George's Respiratory Questionnaire (SGRQ) in Woche 24. • COPD-Symptome und Exazerbationen (Aufzeichnung in einem elektronischen Tagebuch) • Gebrauch der Bedarfsmedikation. • Zusätzliche Wirksamkeit (Nennung wesentlicher zusätzlicher

Item ^a	Charakteristikum	Studieninformation
		<p>Variablen): Lungenfunktionstests (FEV1, FVC und inspiratorische Kapazität [IC]) zur Ermittlung der Trough- und Peak-Werte sowie des Wertes zum Zeitpunkt der Verabreichung, einschließlich Messungen aus der 12-stündigen seriellen Spirometrie-Substudie, normalisierter Bereich unter der Kurve (AUC) von 0 bis 3 Stunden (AUC0-3Std); medianer TDI-Wert, SGRQ-Gesamtpunktzahl, EXACT (Exacerbations of Chronic Pulmonary Disease Tool) respiratorische Symptome (E-RS), erfasst im EXACT zu den vom Patienten angegebenen Ergebnissen (EXACT-PRO), nächtliche und morgendliche COPD-Symptome, tägliche Verwendung von schmerzstillenden Medikamenten, EQ-5D-Fragebogen, COPD-Exazerbationen auf Grundlage der Nutzung von Ressourcen des Gesundheitswesens (elektronischer Patientenerhebungsbogen [eCRF]) und COPD-Exazerbationen anhand EXACT-PRO.</p> <p>Sicherheit:</p> <ul style="list-style-type: none"> • Unerwünschter Ereignisse (UE) und schwerwiegende UE (SUE) • Laboruntersuchungen (Blutbild, Blutbiochemie, Urin und der Theophyllin-Konzentration) • Körperliche Untersuchungen einschließlich Blutdruckmessung und 12-Kanal-EKGs. Bei gebärfähigen Frauen wurde ein Schwangerschaftstest durchgeführt. <p>Alle Prüfzentren verfügten über eine identische Spirometrie-Ausrüstung und ein detailliertes Studienhandbuch, und erhielten die gleichen Schulungen. Die Spirometrie-Daten wurden elektronisch an ein Datenmanagementzentrum übermittelt und dort auf ihre Qualität überprüft. Nur technisch einwandfreie Messungen wurden akzeptiert. Die Ermittlung des Baseline Dyspnoea Index (BDI) und des Transitional Dyspnoea Index (TDI) erfolgte durch einen unabhängigen Interviewer, der keine Kenntnis von weiteren, bei dem jeweiligen Patienten erhobenen Parametern hatte, um Verzerrungen zu vermeiden.</p> <p>Die Dokumentation der Patientendaten erfolgte mit einem EDC-System in elektronischer Form.</p> <p>Die Verifizierung der Quelldaten führten klinisch-wissenschaftlichen Mitarbeiter des Auftragsforschungsinstituts Parexel durch.</p> <p>Das EDC-System wurde so konfiguriert, dass nur der Prüfarzt und autorisierte klinische Mitarbeiter auf die Daten zugreifen und den elektronischen Patientenerhebungsbogen (eCRF) korrigieren konnten. Alle Daten-Korrekturen oder Änderungen sowie der entsprechende Grund dafür wurden im Auditpfad chronologisch erfasst.</p> <p>Die Patienten erhielten von CRF Inc. ein in einen Terminplaner integriertes elektronisches Tagebuch, in dem sie jeden Tag ihre Daten aufzeichnen konnten. Die wissenschaftlichen Mitarbeiter in</p>

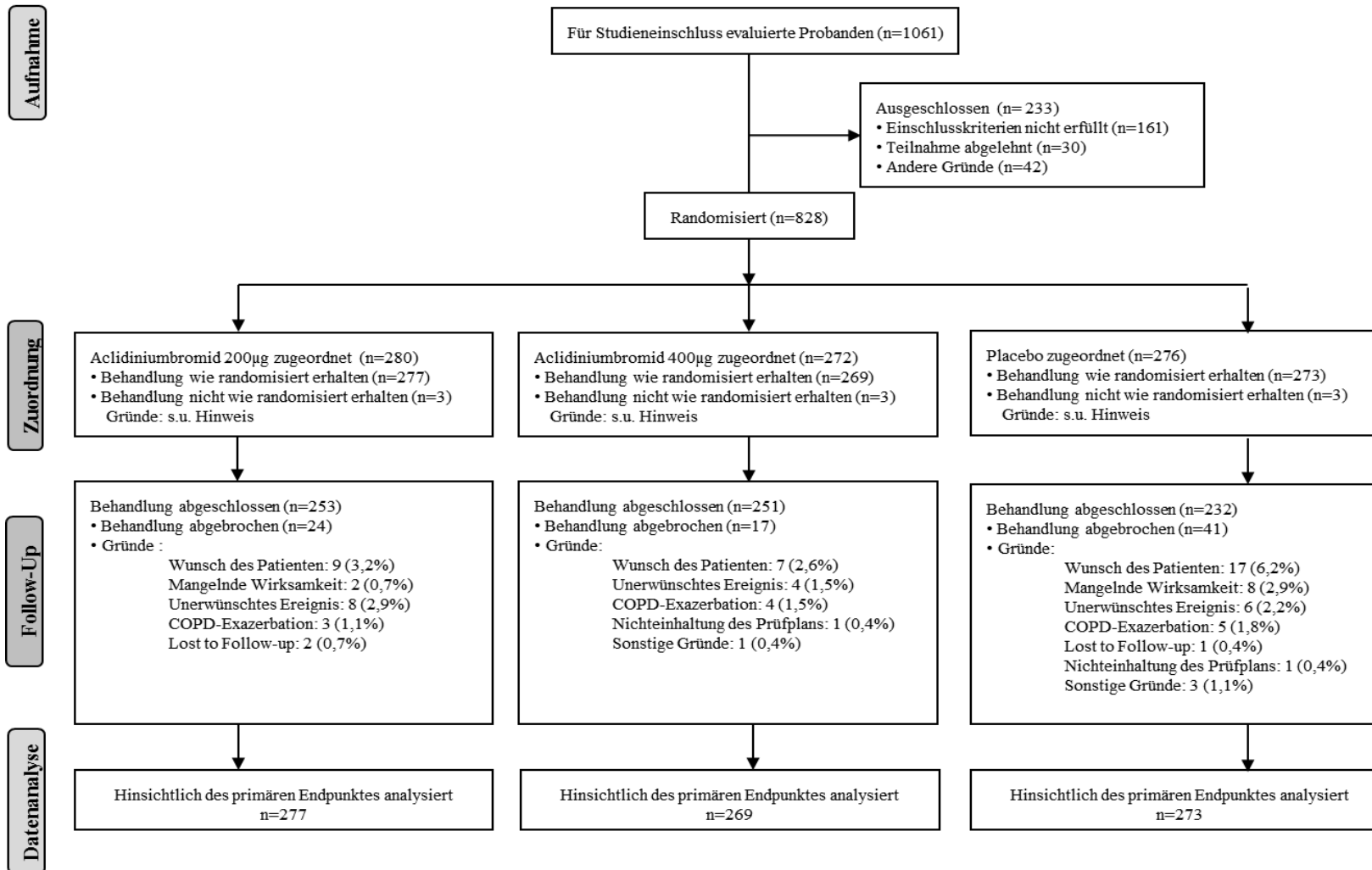
Item ^a	Charakteristikum	Studieninformation
		<p>den Studienzentren wurden in die Verwendung des Tagebuchs eingewiesen, um den Patienten so beim Screening die Funktionsweise zu erklären und die Erläuterungen ggf. beliebig oft zu wiederholen. Darüber hinaus wurden den Patienten schriftliche Informationen zur Verwendung des Terminplaners ausgehändigt.</p> <p>Das Datenmanagement der Studie wurde von PAREXEL gemäß Standardarbeitsanweisung durchgeführt und vom Datenmanagement der Almirall, S.A überwacht. Ein Auditpfad wurde eingerichtet, um die Authentizität und Integrität der klinischen Daten zu gewährleisten.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Wie im abschließenden statistischen Analyseplan (SAP) (mit Datum 12. Dezember 2010) aufgeführt, wurden die Analysestrategie und die spezifischen SAP-Methoden gegenüber den im Protokoll aufgeführten nicht mehr in wichtigen Punkten verändert. Zusätzliche Details wurden ergänzt, wie Variablen abgeleitet und fehlende Werte geschätzt werden sollen. Einige zusätzliche Analysen wurden definiert und einige herausgenommen. Gegenüber der im abschließenden SAP spezifizierten Analysen wurden folgende Änderungen vorgenommen:</p> <ul style="list-style-type: none"> • Der E-RS-Algorithmus wurde zur Schätzung eines täglichen E-RS-Gesamtpunktwerts verwendet sowie von Punktzahlen zu Atemnot, Husten und Sputum, Engegefühl im Brustbereich; die E-RS-Punktwerte basierten nicht auf dem EXACT-Punktwertsystem. • Der prozentuale Anteil der Morgende/Nächte mit spezifischen oder anderen Symptomen (dem elektronischen COPD-Tagebuch entnommen) in Woche 1, 4, 12, 18 und 24 sowie während der ersten 12 und 24 Behandlungswochen wurden mithilfe des parametrischen ANCOVA-Modells analysiert; Faktoren waren das Geschlecht und die Therapie sowie das Alter und eine entsprechende Baseline als Kovariaten zusätzlich zu den vorab spezifizierten normalen ANCOVA-Punktwerten.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>In jeder Behandlungsgruppe mussten insgesamt 244 Patienten die Studie abschließen. Etwa 270 Patienten wurden jeweils zu 200 µg Aclidiniumbromid BID, 400 µg Aclidiniumbromid BID und Placebo BID randomisiert, um eine Abbrecherquote von geschätzten 10% auffangen zu können. Es wurde ein Randomisierungsverhältnis von 1:1:1 (200 µg Aclidiniumbromid BID: 400 µg Aclidiniumbromid BID: Placebo) gewählt; insgesamt wurden somit 810 Patienten randomisiert. Vor der Randomisierung wurden die Patienten zudem einem Screening unterzogen, um mögliche beeinträchtigende Faktoren auszuschließen. Aus früheren Studien war bekannt, dass die Durchfallquote beim Screening bei etwa 35 % liegt. Daher mussten ca. 1.250 Patienten gescreent werden, um die anvisierten</p>

Item ^a	Charakteristikum	Studieninformation
		<p>810 Patienten für die Randomisierung zu erhalten.</p> <p>Als Annahme wurde zugrunde gelegt, dass die Gesamtpopulation von 810 randomisierten Patienten mit mindestens 90 %-iger Sicherheit einen signifikanten Unterschied zwischen Acclidiniumbromid und Placebo nachweisen könne, nämlich 90 ml beim morgendlichen Ausgangs-FEV₁ zur Baseline nach 24 Behandlungswochen für die Antragsstellung in der EU bzw. nach 12 Behandlungswochen für die Antragsstellung in den USA; angenommen wurde dabei eine Standardabweichung von 240 ml bei Verwendung zweiseitiger Tests und Anpassung für einen Vergleich mit verschiedenen Therapien bei einem Gesamt-Signifikanzniveau von 0,05.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Für diese Studie wurde keine Zwischenanalyse geplant.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Interaktives Voice-Response-System (IVRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	200 µg Acclidiniumbromid BID, 400 µg Acclidiniumbromid BID oder Placebo BID in einem Randomisierungsverhältnis von 1:1:1
9	<p>Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment)</p> <p>Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war</p>	Die Patienten wurden beim Screening im IVRS registriert (zwecks Zuweisung einer Patientenidentifikationsnummer). Nach der 14-tägigen Run-in-Phase und nach Durchführung der Qualifikationsbeurteilung vor der Randomisierung wurden die Patienten innerhalb der einzelnen Prüfzentren zu einer der drei Behandlungsgruppen randomisiert. Dies erfolgte mithilfe des IVRS.
10	<p>Randomisierung, Durchführung</p> <p>Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?</p>	Vor Beginn der Studie wurde ein computergenerierter Randomisierungsplan erstellt. In diesem Zusammenhang wurde den Behandlungsgruppen von der Almirall Statistik-/Programmiergruppe eine Randomisierungsnummer zugewiesen. In dem Moment, in dem der Patient die Einwilligungserklärung (ICF) unterzeichnete, wiesen die Studienmitarbeiter jedem Patient eine Patientenidentifikationsnummer zu, indem sie das IVRS anriefen und leiteten dann die Run-in-Phase ein. Die Patienten wurden mithilfe des IVRS randomisiert.

Item ^a	Charakteristikum	Studieninformation
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	<p>a.) und b.) treffen zu</p> <p>Almirall, S.A., Forest, Studienmitarbeiter und Studienteilnehmer wurden verblindet, so dass nicht bekannt wurde, ob der Patient Aclidiniumbromid oder Placebo erhielt. Die Randomisierungsdaten wurden streng vertraulich behandelt und waren nur für autorisierte Personen zugänglich, bis die zugewiesene Behandlungsart für alle Probanden entblindet und die Datenbank zum Studienende gesperrt wurde. Erst nach Beendigung der Studie sowie der Verifizierung und Sperrung der Daten wurden die Randomisierungs-codes zur Datenanalyse freigegeben.</p> <p>Die Prüfärzte erhielten keine versiegelten Notfall-Umschläge, in denen die Behandlungsgruppen identifiziert werden konnten. Im Notfall waren die Prüfärzte angehalten, das IVRS anzurufen, da die medizinische Handlungsentscheidung davon abhing, welches Prüfpräparat verabreicht wurde. Das Auftragsforschungsinstitut, das die Studie durchgeführt hatte, sowie Almirall wurden bei Entblindung der Behandlung sofort informiert. Nachdem der Code am Prüfzentrum offengelegt wurde, wurde der Patient von der Studie ausgeschlossen.</p>
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Um Verzerrungen zu minimieren, wurden Placebo und Aclidiniumbromid mit demselben Inhalator (die gleiche äußerer Form und Zusammensetzung) appliziert. Nur der Wirkstoff im Inhalator war unterschiedlich.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Alle Wirksamkeitsanalysen wurden in der ITT-Population mittels der LOCF-Methode (last-observation-carried-forward) durchgeführt, wenn der fehlende Wert vorhanden war. Die Primärvariable für den Vergleich zwischen Aclidiniumbromid und Placebo wurde mittels der Hochberg-Prozedur für Multiplizität adjustiert. Die sekundären Wirksamkeitsvariablen wurden untersucht, wenn die Primärvariable in beiden Therapievergleichen (200 µg Aclidiniumbromid BID ggü. Placebo und 400 µg Aclidiniumbromid BID ggü. Placebo) signifikant war. Die sequentielle Testung erfolgte in folgender Reihenfolge:</p> <ul style="list-style-type: none"> o Veränderung gegenüber der Baseline beim Peak-FEV₁ in Woche 24 für den Zulassungsantrag in der EU und in Woche 12 für den Antrag in den USA; o Prozentualer Anteil der Patienten mit klinisch relevanter Verbesserung (≥1 Einheit) beim focalen TDI in Woche 24; und o Prozentualer Anteil der Patienten mit klinisch relevanter Verbesserung (≥4 Einheiten) verglichen mit der Baseline beim SGRQ-Gesamtpunktwert in Woche 24. <p>Die Veränderung gegenüber der Baseline beim Peak-FEV₁ wurde untersucht, indem zunächst die Hochberg-Prozedur auf das 5 %-Signifikanzniveau angewandt wurde. Prozess für den Übergang</p>

Item ^a	Charakteristikum	Studieninformation
		<p>zum sequenziellen Verfahren: Die Untersuchungen wurden mit der nächsten Variable fortgeführt, wenn mindestens eine Null-Hypothese bei den beiden Behandlungsvergleichen zurückgewiesen wurde; andernfalls wurden die sequenziellen Verfahren wegen inferentieller Zwecke gestoppt. Waren beide Dosierungen (Therapievergleiche) signifikant, dann wurde die Hochberg-Prozedur angewandt, um Korrekturen im Hinblick auf Mehrfach-Therapievergleiche vorzunehmen; andernfalls wurden keine Korrekturen vorgenommen und die abgegebene Dosis stand in der Sequenz für keine der verbleibenden Sekundärvariablen für inferentielle Tests zur Verfügung. Die primären Wirksamkeitsvariablen (Veränderung ggü. Baseline beim Ausgangs-FEV₁) und sekundäre Variablen (Veränderung ggü. Baseline beim Peak-FEV₁) wurden mithilfe eines Kovarianzanalyse-Modells (ANCOVA) analysiert; dieses umfasste als Faktoren das Geschlecht und die Behandlungsgruppen und als Kovariaten das Baseline-FEV₁ und das Alter. Die zweigliedrigen Sekundärvariablen wurden mithilfe logistischer Regressionsmodelle analysiert, bei denen Baseline, Alter, Geschlecht und Behandlung als exploratorische Variable herangezogen wurden. Die Sicherheitsergebnisse wurden in der Sicherheitspopulation analysiert. Die Sicherheitsergebnisse (UE, Laborparameter, Vitalzeichen und 12-Kanal-EKG) wurden mithilfe deskriptiver Statistiken über die gesamte Zeit für die jeweilige Behandlungsgruppe zusammengefasst. Für die letzten drei Sicherheitsergebnisse wurden potenzielle klinisch signifikante Werte nach Studienbeginn für die einzelnen Behandlungsgruppen tabellarisiert.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Für die Primärvariable wurde ein Modell mit gemischten Effekten (Mixed-Effects Model) für Wiederholungswerte auf Basis beobachteter Fälle eingesetzt und eine Per-Protocol-Analyse anhand der LOCF-Methode als Empfindlichkeitsanalyse der ITT-Population bzw. als Ansatz für den Umgang mit fehlenden Daten durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante	<ul style="list-style-type: none"> • Randomisierung: 200 µg Aclidiniumbromid BID (280), 400 µg Aclidiniumbromid BID (272), Placebo BID (276) • geplante Behandlung erhalten: 200 µg Aclidiniumbromid BID (277), 400 µg Aclidiniumbromid BID (269), Placebo BID (273) • Analyse primärer Ergebnisse: 200 µg Aclidiniumbromid BID (277), 400 µg Aclidiniumbromid BID (269), Placebo BID (273)

Item ^a	Charakteristikum	Studieninformation
	Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Ablaufdiagramm (s. u.).
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • 200 µg Aclidiniumbromid BID: 24; Unerwünschtes Ereignis: 8; Nichteinhaltung des Prüfplans: 0; Lost-to-Follow-Up: 2; Persönlicher Wunsch des Patienten: 9; COPD-Exazerbation: 3; Mangelnde Wirksamkeit: 2; Andere: 0 • 400 µg Aclidiniumbromid BID: 17; Unerwünschtes Ereignis: 4; Nichteinhaltung des Prüfplans: 1; Lost-to-Follow-Up: 0; Persönlicher Wunsch des Patienten: 7; COPD-Exazerbation: 4; Mangelnde Wirksamkeit: 0; Andere: 1 • Placebo BID: 41; Unerwünschtes Ereignis: 6; Nichteinhaltung des Prüfplans: Aclidiniumbromid 1; Lost-to-Follow-Up: 1; Persönlicher Wunsch des Patienten: 17; COPD Exazerbation: 5; Mangelnde Wirksamkeit: 8; Andere: 3
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Datum des Studienbeginns (erstes Screening): 07. Oktober 2009 Datum des Abschlusses der Studie: (letzter Patientenkontakt): 08. November 2010
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		



Hinweis: Alle Patienten (insgesamt 9), die mit dem Prüfarzt mit der Nummer 1221 in Verbindung stehen, wurden nur als randomisiert gezählt. Die Patienten, die die Behandlung abgeschlossen haben + die Patienten, die abgebrochen haben, werden daher nicht zur Gesamtzahl der randomisierten Patienten (N) gezählt.

Tabelle 4.7-7 (Anhang): Studiendesign und -methodik für Studie ACCORD I

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<ul style="list-style-type: none"> • Beurteilung der langfristigen bronchodilatatorischen Wirksamkeit von inhalativem Acclidiniumbromid in einer Dosierung von entweder 200 µg oder 400 µg, verabreicht zweimal täglich (morgens und abends), mithilfe des Genuair Mehrfachdosis-Trockenpulverinhalators im Vergleich zu Placebo bei Patienten mit mittelschwerer bis schwerer COPD • Beurteilung der Sicherheit und Verträglichkeit von Acclidiniumbromid in einer Dosierung von entweder 200 µg oder 400 µg, verabreicht zweimal täglich (morgens und abends), im Vergleich mit Placebo • Beurteilung des Nutzens von Acclidiniumbromid in einer Dosierung von entweder 200 µg oder 400 µg, verabreicht zweimal täglich (morgens und abends), im Hinblick auf COPD Exazerbationen, den krankheitsbezogenen Gesundheitsstatus gemessen anhand des St. George's Respiratory Questionnaire (SGRQ), COPD-Symptomen und weiteren Ergebnissen bei Patienten mit mittelschwerer bis schwerer COPD
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Parallelgruppe, 200 µg Acclidiniumbromid BID, 400 µg Acclidiniumbromid BID oder Placebo BID in einem Randomisierungsverhältnis von 1:1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Amendment 2 mit Datum vom 14. Juli 2009 wurde nach Studienbeginn am Originalprüfplan vorgenommen.</p> <ul style="list-style-type: none"> • Ergänzung des EQ-5D-Fragebogens zur Lebensqualität für noch nicht randomisierte Patienten • Erläuterung, dass BDI und TDI von einem unabhängigen Interviewer ermittelt wurden • Verdeutlichung der Einschlusskriterien im Hinblick auf die Messungen von FEV₁ und FVC und den Zeitpunkt (vor und nach der Dosisverabreichung), die zur Beurteilung verwendet wurden, ob die Ergebnisse der Lungenfunktionsprüfung die Reproduzierbarkeitskriterien bei Besuch 1 stützen • Herausnahme der Einschlusskriterien hinsichtlich der akzeptablen Blutdruckwerte – um diese allein der Beurteilung des Prüfarztes zu überlassen • Klärung des Zeitpunkts für die Spirometrie bei Besuch 1 • Verkürzung des Rauchverbots von acht Stunden auf eine Stunde vor jedem Besuch und Anpassung anderer Einschränkungen gemäß ATS/ERS-Kriterien • Ergänzung weiterer Wirksamkeitsanalysen und Neuordnung der Variablen für eine größere Eindeutigkeit
4	Probanden/Patienten	

Item ^a	Charakteristikum	Studieninformation
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<ul style="list-style-type: none"> • Männliche und weibliche ambulant behandelte Patienten, mindestens 40 Jahre alt, ehemalige und aktuelle Raucher mit mindestens 10 Packungsjahren. • Stabile, mittelschwere bis schwere COPD, definiert anhand der GOLD-Kriterien (Global Initiative for Chronic Obstructive Lung Disease) und stabile Atemwegsobstruktion. • Forciertes Expirationsvolumen in einer Sekunde (Einsekundenkapazität, FEV₁) nach 10 bis 15 Minuten Verabreichung von Albuterol/Salbutamol $\geq 30\%$ bis $< 80\%$ des prognostizierten Normalwertes. • Verhältnis FEV₁/forcierte Vitalkapazität (FVC) $< 70\%$. • Keine frühere oder aktuelle Asthma-Erkrankung. • Keine Infektion der Atemwege oder COPD-Exazerbation in den 6 Wochen (3 Monate, wenn die Exazerbation eine stationäre Aufnahme erforderte) vor Besuch 1. • Keine klinisch relevanten kardiovaskulären Anomalien, abweichenden Laboruntersuchungen oder Unregelmäßigkeiten im Elektrokardiogramm (EKG).
4b	Studienorganisation und Ort der Studiendurchführung	<p>106 Zentren in Nordamerika (100 Zentren in den USA und 6 Zentren in Kanada) führten für die Studie Patientenscreenings durch. 99 dieser Zentren randomisierten Patienten im Rahmen der Studie. Alle Zentren, die keine Patienten randomisierten, befanden sich in den USA</p> <p>Die Studie wurde von Forest Research Institute, Inc. (FRI) gesponsort, einer hundertprozentigen Tochtergesellschaft von Forest Laboratories, Inc. und Almirall, S.A.</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Name: Acclidiniumbromid; Verabreichungsweg: Orale Inhalation mithilfe des Genuair® Mehrfachdosis-Trockenpulver-Inhalators; Dosierungsform: Inhalationspulver-Dosis und Regime: 1 Sprühstoß mit 200 µg oder 400 µg am Morgen (08:00 – 10:00) und am Abend (20:00 – 22:00).</p> <p>Name: Placebo im Vergleich mit Acclidiniumbromid; Verabreichungsweg: Orale Inhalation mithilfe des Genuair® Mehrfachdosis-Trockenpulver-Inhalators; Dosierungsform: Inhalationspulver-Dosis und Regime: 1 Sprühstoß Placebo am Morgen (08:00 – 10:00) und am Abend (20:00 – 22:00).</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <ul style="list-style-type: none"> • Veränderung gegenüber der Baseline des morgendlichen Ausgangs- (Trough) FEV₁ in Woche 12 <p>Sekundärer Wirksamkeitsendpunkt:</p> <ul style="list-style-type: none"> • Veränderung gegenüber der Baseline des Peak-FEV₁ in Woche 12 <p>Zusätzliche Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • Lungenfunktionstest (FEV₁, FVC, inspiratorische Kapazität [IC]) zur Ermittlung der Trough- und Peak-Werte sowie des Wertes zum Zeitpunkt der Verabreichung; SGRQ; Transitional Dyspnoea Index (TDI); COPD-Exazerbationen (definiert als Anstieg der COPD-Symptome an mindestens 2

Item ^a	Charakteristikum	Studieninformation
		<p>aufeinanderfolgenden Tagen, die eine medizinische Intervention erforderten); Behandlung mit Notfallmedikation. Weitere vom Patienten selber vorgenommene Beurteilungen der Gesundheitsergebnisse umfassten den Fragebogen zu nächtlichen COPD-Symptomen und das täglich auszufüllende Schlaftagebuch.</p> <p>Sicherheit: Unerwünschte Ereignisse (UE), Laborparameter, Vitalzeichen, EKG und Holter-Monitoring (nur Patienten der Substudie)</p> <p>Vor Beginn der Studie wurde ein Treffen der Prüfarzte abgehalten. Ziel dieses Meetings war es, den Prüfplan zu überprüfen und Anweisungen zu allen Verfahren zur Verfügung zu stellen, die für die Durchführung (z. B. Erfassung klinischer Labor- und EKG-Werte, Dateneingabe) und den Überblick (z. B. Schulung zur Überwachung) über die Studie erforderlich waren. Vor Beginn der Studie nahmen alle Mitarbeiter der Prüzentren, die Spirometrie-Messungen, BDI/TDI- oder SGRQ-Befragungen durchführen sollten, an einer Schulung teil und wurden von CareFusion Research Services (Unterstützung klinischer Studien der Medikamentenerprobung) zertifiziert; Mitarbeiter, die EKG oder Holter-Monitoring durchführen sollten, wurden von ERT geschult.</p> <p>Um die Sicherheitsanalyse zu standardisieren, wurden Zentrallabors für die Verarbeitung und Analyse aller klinischen Laborwerte (Covance) sowie EKG und Holter-Monitoring (ERT) eingesetzt.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Auf Basis des Original-SAP vom 21. Juli 2009 wurde am 22. Dezember 2009 Amendment 1 übernommen. Vor der Entblindung der Studie wurden die folgenden wichtigen Änderungen vorgenommen:</p> <ul style="list-style-type: none"> • Die Analyse des EQ-5D-Fragebogens wurde gestrichen; Grund dafür war, dass zum Zeitpunkt der Genehmigung der zweiten Prüfplan-Änderung durch die Ethikkommission (Ergänzung des EQ-5D-Fragebogens) alle Patienten bereits randomisiert worden waren und daher kein Baseline-EQ-5D-Fragebogen zur Verfügung stand • Für die gesammelten Daten wurde folgender Wirksamkeitsendpunkt ergänzt: Anzahl der Tage des stationären Aufenthalts wegen COPD Exazerbation pro Patient/Jahr <p>Wichtige Änderungen und/oder zusätzliche statistische Analysen (nach Schließung der Datenbank:</p> <ul style="list-style-type: none"> • Die Analyse der Anzahl der Exazerbationstage pro Patient/Jahr wurde auf Basis eines Poisson-Regressionsmodells für alle Patienten aktualisiert, anstelle des ursprünglich geplanten Varianzanalyse-Modells • Die Analyse der Dauer bis zur ersten leichten, mittelschweren oder schweren COPD-Exazerbation wurde ergänzt • Für UE, die vor der Baseline erstmals auftraten und sich während der Behandlungsperiode mit gleicher oder verringerter Schwere wiederholten (Vergleichsarm), wurde eine neue Auflistung erstellt

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> Die zusätzlichen Analysen wurden für alle symptombezogenen Variablen aus dem Fragebogen zu nächtlichen COPD-Symptomen durchgeführt; Basis war der Durchschnitt der wöchentlichen Werte aus den 12 Studienwochen
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Für die Bestimmung der Fallzahl und der statistischen Aussagekraft wurde angenommen, dass in der ITT-Population für die Analyse der primären und sekundären Wirksamkeitsparameter jeder Behandlungsarm 165 Patienten umfassen muss. Um unter Annahme einer Standardabweichung von 0,240 l einen Therapieunterschied von 0,100 l für den primären Wirksamkeitsparameter (Veränderung gegenüber der Baseline beim morgendlichen Ausgangs- (Trough) FEV₁) in Woche 12 nachzuweisen, lag für beide Vergleiche in Bezug auf die zweimal täglich verabreichte Dosierung von Acridiniumbromid im Vergleich mit Placebo eine statistische Aussagekraft von mindestens 93 % vor (angepasst um die Einflüsse des Mehrfachvergleichsverfahrens), bei einem Gesamt-Signifikanzniveau von 0,05.</p> <p>Um wiederum unter Annahme einer Standardabweichung von 0,300 l einen Therapieunterschied von 0,150 l für den sekundären Wirksamkeitsparameter (Veränderung gegenüber der Baseline des Peak-FEV₁) in Woche 12 nachzuweisen, lag für beide Vergleiche in Bezug auf die zweimal täglich verabreichte Dosierung von Acridiniumbromid im Vergleich mit Placebo eine statistische Aussagekraft von mindestens 92 % vor (angepasst um die Einflüsse des Mehrfachvergleichsverfahrens), bei einem Gesamt-Signifikanzniveau von 0,05.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Für diese Studie wurde keine Zwischenanalyse geplant.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Interaktives Voice-Response-System (IVRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	200 µg Acridiniumbromid BID, 400 µg Acridiniumbromid BID oder Placebo BID in einem Randomisierungsverhältnis von 1:1:1
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale	Den Patienten wurde über das IVRS eine Identifikationsnummer zugewiesen, sobald sie die Einwilligungserklärung beim Screening unterzeichnet hatten. Der zweite Besuch erfolgte etwa zwei Wochen (± 3 Tage) nach dem Screening, um die Eignung der Patienten für eine doppelblinde Therapie im Rahmen der Studie zu prüfen. Wurde der Patient als für die Studie geeignet

Item ^a	Charakteristikum	Studieninformation
	Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	eingestuft, so wurde ihm eine Packung der Studienmedikation mit der entsprechenden Randomisierungsnummer zugewiesen.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Patienten erhielten von den Studienmitarbeitern über das IVRS eine Identifikationsnummer. Vom Sponsor wurde eine Liste der Codes für die Patientenrandomisierung erstellt. In dieser wurde jeder Patient anhand seiner Randomisierungsnummer identifiziert und die entsprechende zugewiesene Therapie aufgeführt (eine elektronische Version wurde an einem sicheren gesperrten Ort aufbewahrt). Die Randomisierungs-codes und das geplante Set wurden dem IVRS-Anbieter auf gesichertem Weg zur Verfügung gestellt. Die Patienten wurden mithilfe des IVRS randomisiert.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a.) und b.) treffen zu Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Für die Sicherheit und das Wohlbefinden des Patienten war der Prüfarzt in Notfallsituationen berechtigt, die Verblindung über das IVRS-System aufzuheben. Im Falle einer Entblindung wurde das Forest Research Institute umgehend informiert und es wurde eine umfassende schriftliche Erklärung vorgelegt. Nachdem der Code am Prüfzentrum offengelegt wurde, wurde der Patient sofort von der Studie ausgeschlossen.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Alle Inhalatoren waren äußerlich identisch und enthielten die gleiche Zusammensetzung; die Placebo-Inhalatoren enthielten allerdings keinen Wirkstoff. Der Wirkstoff war im Hinblick auf Geschmack, Erscheinung, Geruch oder Farbe nicht zu identifizieren, so dass die Verblindung nicht durchbrochen werden konnte.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Wirksamkeit: Die Wirksamkeitsanalysen basierten auf den primären LOCF-Werten der ITT-Population. Die primären Wirksamkeitsparameter (Veränderung ggü. Baseline beim morgendlichen Ausgangs-FEV ₁ (trough) in Woche 12) wurden mithilfe eines Kovarianzanalyse-Modells (ANCOVA) analysiert; dieses umfasste als Faktoren das Geschlecht und die Behandlungsgruppen und als Kovariaten das Baseline-FEV ₁ und das Alter. Sicherheit: Alle Sicherheitsparameter wurden deskriptiv ausgewertet. Die Sicherheitsanalysen basierten auf der Sicherheitspopulation.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und	Für die Primärvariable wurde ein Modell mit gemischten Effekten (Mixed-Effects Model) für wiederholte Messungen auf Basis

Item ^a	Charakteristikum	Studieninformation
	adjustierte Analysen	beobachteter Fälle (OC) eingesetzt und eine Per-Protocol-Analyse (PP) anhand der LOCF-Methode als Empfindlichkeitsanalyse der ITT-Population durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Randomisierung: 200 µg Acridiniumbromid BID (185), 400 µg Acridiniumbromid BID (190), Placebo BID (186) • Geplante Behandlung erhalten: 200 µg Acridiniumbromid BID (184), 400 µg Acridiniumbromid BID (190), Placebo BID (186) • Analyse primärer Ergebnisse: 200 µg Acridiniumbromid BID (184), 400 µg Acridiniumbromid BID (190), Placebo BID (185)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • 200 µg Acridiniumbromid BID: 33; Widerruf der Einwilligung: 6; Unerwünschtes Ereignis: 8; unzureichendes therapeutisches Ansprechen: 5; COPD-Exazerbation: 4; andere: 5; Verstoß gegen den Prüfplan: 1; Einschluss-/Ausschlusskriterien: 2; Lost-To-Follow-Up: 2 • 400 µg Acridiniumbromid BID: 24; Widerruf der Einwilligung: 7; Unerwünschtes Ereignis: 7; unzureichendes therapeutisches Ansprechen: 1; COPD Exazerbation: 1; andere: 3; Verstoß gegen den Prüfplan: 3; Einschluss-/Ausschlusskriterien: 2; Lost-To-Follow-Up: 0 • Placebo BID: 37; Widerruf der Einwilligung: 9; Unerwünschtes Ereignis: 7; unzureichendes therapeutisches Ansprechen: 10; COPD Exazerbation: 7; andere: 2; Verstoß gegen den Prüfplan: 2; Einschluss-/Ausschlusskriterien: 0; Lost-To-Follow-Up: 0
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Erster Kontakt zu erstem Patienten: 28. April 2009 Letzter Kontakt zu letztem Patienten: 03. November 2009
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		

Item ^a	Charakteristikum	Studieninformation

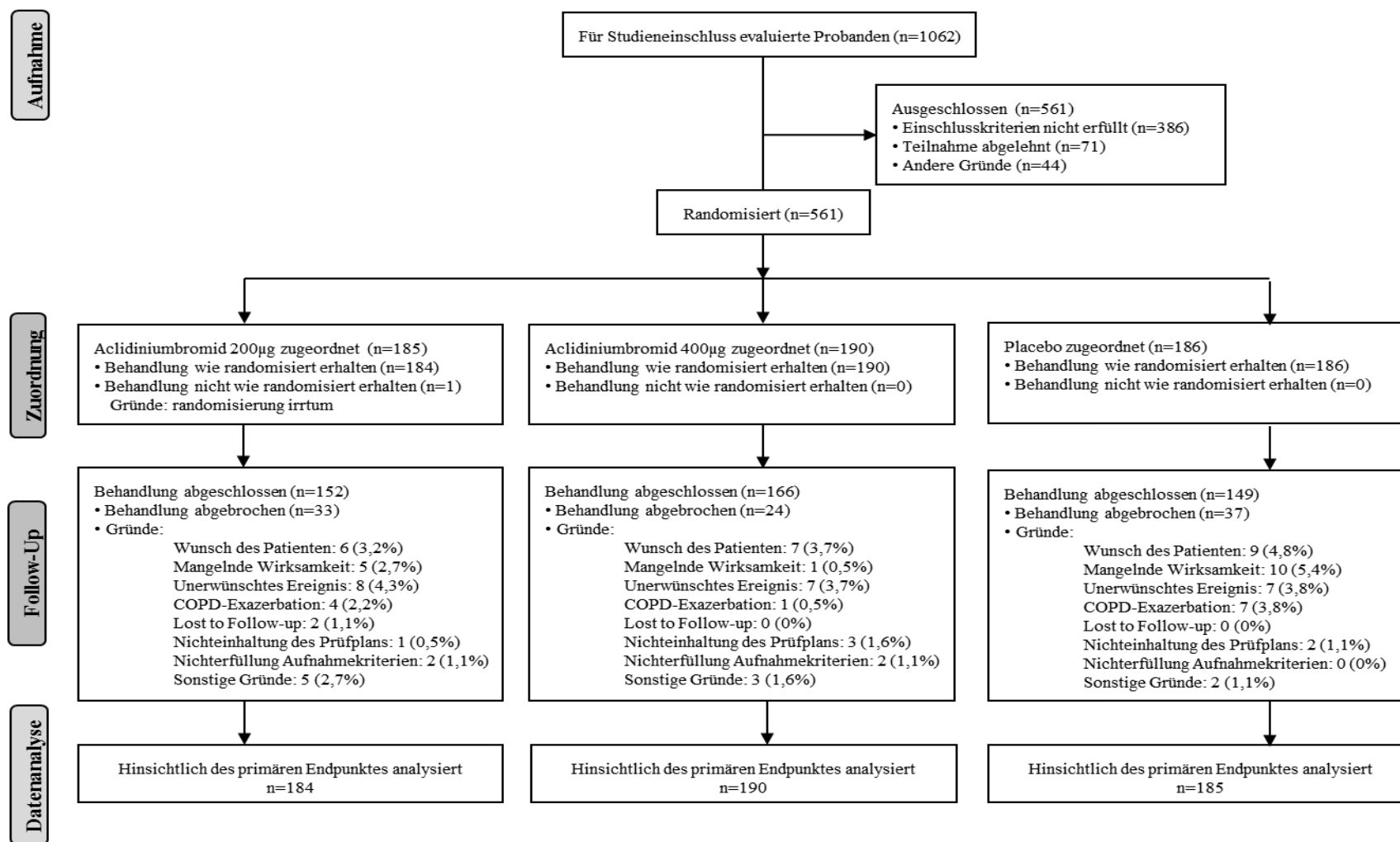


Tabelle 4.7-8 (Anhang): Studiendesign und -methodik für Studie ACCORD II

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<ul style="list-style-type: none"> • Beurteilung der bronchodilatatorischen Wirksamkeit von inhalativem Aclidiniumbromid in einer Dosierung von entweder 200 µg oder 400 µg, verabreicht zweimal täglich (BID) über 12 Wochen, verglichen mit Placebo bei Patienten mit mittelschwerer bis schwerer stabiler chronisch-obstruktiver Lungenerkrankung (COPD) • Beurteilung der langfristigen bronchodilatatorischen Sicherheit und Verträglichkeit von inhalativem Aclidiniumbromid in einer Dosierung von entweder 200 µg oder 400 µg, verabreicht zweimal täglich (BID) über 12 Wochen verglichen mit Placebo bei Patienten mit mittelschwerer bis schwerer stabiler COPD <p>Zusätzliches Ziel bei Part B (für Part A nicht weiter relevant, Gegenstand dieses CONSORT-Statements ist Part A):</p> <ul style="list-style-type: none"> • Beurteilung der langfristigen bronchodilatatorischen Sicherheit, Verträglichkeit und Wirksamkeit von inhalativem Aclidiniumbromid in einer Dosierung von 400 µg, verabreicht zweimal täglich (BID) über weitere 40 Wochen und Beurteilung des Nutzens für den Gesundheitsstatus (gemessen anhand Baseline/Transitional Dyspnoea Index (TDI) [BDI/TDI], St. George's Respiratory Questionnaire [SGRQ], EuroQol-Fragebogen zur Lebensqualität [EQ-5D] und dem COPD-Fragebogen zur Ressourcennutzung) und weiterer Ergebnisse bei Patienten mit mittelschwerer bis schwerer stabiler COPD
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Parallelgruppe, 200 µg Aclidiniumbromid BID, 400 µg Aclidiniumbromid BID oder Placebo BID in einem Randomisierungsverhältnis von 1:1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Ziel von Amendment 1 vom 02. März 2010:</p> <ul style="list-style-type: none"> • Änderung des Studiendesigns in Teil B von doppelblind zu offen • Herausnahme der 200-µg-Dosis aus Teil B • Überarbeitung der statistischen Analyse infolge dieser Änderungen • Ergänzung einer körperlichen Untersuchung am Ende von Teil A (Besuch 6), um sicherzustellen, dass alle Änderungen, die nach Beginn der Studie auftreten, am Ende der doppelblinden Behandlungsperiode ausgewertet und verzeichnet werden • Klärung der Prozesse für die Spirometrie-Schulung, den Umgang mit Daten und den besten Testreviews.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der	<ul style="list-style-type: none"> • Männliche und weibliche ambulant behandelte Patienten, mindestens 40 Jahre alt,

Item ^a	Charakteristikum	Studieninformation
	Probanden/Patienten	<ul style="list-style-type: none"> • aktuelle und ehemalige Raucher von mindestens 10 Packungsjahren pro Jahr • Stabile, mittelschwere bis schwere COPD, definiert anhand der GOLD-Kriterien (Global Initiative for Chronic Obstructive Lung Disease), stabile Atemwegsobstruktion • Forciertes Expirationsvolumen in 1 Sekunde (Einsekundenkapazität, FEV₁) 10-15 Minuten nach Verabreichung von Albuterol/Salbutamol ≥ 30 % bis < 80 % des prognostizierten Normalwertes • FEV₁/forcierte Vitalkapazität (FVC) < 70 % • Keine anamnetische oder aktuelle Asthma-Erkrankung • Keine Infektion der Atemwege oder COPD-Exazerbation in den 6 Wochen (3 Monate, wenn die Exazerbation eine stationäre Aufnahme erforderte) vor Besuch 1 • Keine klinisch relevanten kardiovaskulären Anomalien oder Unregelmäßigkeiten im Elektrokardiogramm (EKG)
4b	Studienorganisation und Ort der Studiendurchführung	<p>Diese Studie wurde in 112 Prüfzentren (Nordamerika) durchgeführt: 110 Zentren in den USA und 2 Zentren in Kanada. 103 Prüfzentren (101 Zentren in den USA und 2 Zentren in Kanada) randomisierten Patienten.</p> <p>Diese Arbeit wurde von Forest Research Institute, Inc. (FRI) finanziert, einer hundertprozentigen Tochtergesellschaft von Forest Laboratories, Inc. und Almirall, S.A.</p> <p>Medizinischer Leiter: Marco Taglietti Forest Research Institute Harborside Financial Center - Plaza V Jersey City, New Jersey 07311, USA</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Name: Acridiniumbromid; Verabreichungsweg: Orale Inhalation mithilfe des Genuair® Mehrfachdosis-Trockenpulver-Inhalators; Dosierungsform: Inhalationspulver-Dosis und Regime: 1 Sprühstoß mit 200 µg oder 400 µg am Morgen (08:00 – 10:00) und am Abend (20:00 – 22:00).</p> <p>Name: Placebo im Vergleich mit Acridiniumbromid; Verabreichungsweg: Orale Inhalation mithilfe des Genuair® Mehrfachdosis-Trockenpulver-Inhalators; Dosierungsform: Inhalationspulver-Dosis und Regime: 1 Sprühstoß Placebo am Morgen (08:00 – 10:00) und am Abend (20:00 – 22:00).</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von	<p>Primäre Wirksamkeitsendpunkte</p> <ul style="list-style-type: none"> • Veränderung gegenüber der Baseline beim morgendlichen Ausgangs- (Trough) FEV₁ in Woche 12 <p>Sekundäre Wirksamkeit:</p> <ul style="list-style-type: none"> • Veränderung gegenüber der Baseline beim Peak-FEV₁ in Woche 12 <p>Zusätzliche Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • Lungenfunktionstest (FEV₁, FVC, inspiratorische Kapazität [IC]) zur Ermittlung der Trough- und Peak-Werte sowie des Wertes zum Zeitpunkt der Verabreichung; St. George's Respiratory

Item ^a	Charakteristikum	Studieninformation
	Erhebungsinstrumenten	<p>Questionnaire (SGRQ); Transitional Dyspnoea Index, (TDI); COPD-Exazerbationen (definiert als ein Anstieg der COPD-Symptome an mindestens 2 aufeinanderfolgenden Tagen, die eine medizinische Intervention erforderten); Behandlung mit Notfallmedikation.</p> <p>Sicherheitsendpunkte:</p> <ul style="list-style-type: none"> • Aufzeichnung unerwünschter Ereignisse (UE), Laborparameter, Vitalzeichen, körperliche Untersuchung, EKG und Holter-Monitoring (nur Patienten der Substudie) <p>Vor Beginn der Studie wurde ein Treffen der Prüfärzte abgehalten. Ziel dieses Meetings war es, den Prüfplan zu überprüfen und Anweisungen zu allen Verfahren zur Verfügung zu stellen, die für die Durchführung (z. B. Erfassung klinischer Labor- und EKG-Werte, Dateneingabe) und den Überblick (z. B. Schulung zur Überwachung) über die Studie erforderlich waren. Vor Beginn der Studie mussten alle Mitarbeiter der Prüfzentren, die Spirometrie-Messungen, BDI/TDI- oder SGRQ-Befragungen durchführen sollten, an einer Schulung teilnehmen und von CareFusion Research Services (Unterstützung klinischer Studien der Medikamentenerprobung) zertifiziert werden; Mitarbeiter, die EKG oder Holter-Monitoring durchführen sollten, wurden von ERT geschult.</p> <p>Um die Analyse von Sicherheitsbewertungen zu standardisieren, wurden Zentrallabors für die Verarbeitung und Analyse aller klinischen Laboruntersuchungen (Covance) sowie EKG und Holter-Monitoring (ERT) eingesetzt.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Es wurden keine Änderungen an den geplanten Analysen vorgenommen. Zusätzliche Post-hoc-Analysen:</p> <ul style="list-style-type: none"> • Vergleiche der COPD-Schweregrade und des FEV₁ (in l) zur Baseline, um zu untersuchen, ob die Werte in den einzelnen Behandlungsgruppen ausgeglichen waren. • Subgruppenanalyse des primären Wirksamkeitsparameters in zwei nach COPD-Schweregrad gebildeten Subgruppen (leicht + mittelschwer und schwer + sehr schwer). Diese Analyse wurde aufgrund eines Ungleichgewichts beim COPD-Schweregrad zur Baseline zwischen den Behandlungsgruppen durchgeführt.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Für die Beurteilung der Probengröße und der statistischen Aussagekraft für Teil A der Studie wurde angenommen, dass in der ITT-Population für die Analyse der primären und sekundären Wirksamkeitsparameter jede Behandlungsgruppe 165 Patienten umfasste. Um unter Annahme einer Standardabweichung von 0,241 einen Therapieunterschied von 0,101 für den primären Wirksamkeitsparameter (Veränderung gegenüber der Baseline beim morgendlichen Ausgangs- (Trough) FEV₁) in Woche 12 nachzuweisen, lag für beide Vergleiche in Bezug auf die BID-Dosierung von Aclidiniumbromid im Vergleich mit Placebo eine statistische Aussagekraft von mindestens 93 % vor (angepasst um die Einflüsse des Mehrfachvergleichsverfahrens), bei einem Gesamt-Signifikanzniveau von 0,05. Um unter Annahme einer</p>

Item ^a	Charakteristikum	Studieninformation
		Standardabweichung von 0,301 einen Therapieunterschied von 0,151 für den sekundären Wirksamkeitsparameter (Veränderung gegenüber der Baseline beim Peak-FEV ₁) in Woche 12 nachzuweisen, lag für beide Vergleiche in Bezug auf die BID-Dosierung von Acridiniumbromid im Vergleich mit Placebo eine statistische Aussagekraft von mindestens 92 % vor (angepasst um die Einflüsse des Mehrfachvergleichsverfahrens), bei einem Gesamt-Signifikanzniveau von 0,05.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Für Teil A der Studie wurde keine Zwischenanalyse durchgeführt. Die Studie wurde nach Ablauf der 12 Therapiewochen entblindet. Alle vorab spezifizierten Analysen, die auf den in den ersten 12 Wochen gesammelten Daten basierten, wurden durchgeführt. Für die Daten aus diesen 12 Wochen wurden alle Fehler vom Typ 1 in den Wirksamkeitsanalysen umfassend untersucht.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Interaktives Voice-Response-System (IVRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	200 µg Acridiniumbromid BID, 400 µg Acridiniumbromid BID oder Placebo BID in einem Randomisierungsverhältnis von 1:1:1
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Patienten erhielten über das IVRS eine Identifikationsnummer, sobald sie die Einverständniserklärung unterschrieben hatten. Wurden alle Eignungskriterien erfüllt, wurde den Patienten über das IVRS eine Randomisierungsnummer zugewiesen.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Patienten erhielten von den Studienmitarbeitern über das IVRS eine Identifikationsnummer. Eine Liste der Patienten-Randomisierungs-codes wurde von der Abteilung für statistische Programmierung im Forest Research Institute Inc. generiert (eine elektronische Version wurde in einem gesperrten Bereich auf einem Sicherheitsserver gespeichert) und von der Abteilung für die Überwachung der Arzneimittelsicherheit des Instituts aufbewahrt. In dieser Liste wurden die Patienten mit ihrer Randomisierungsnummer und der jeweils zugewiesenen Therapieart verzeichnet. Die Randomisierungs-codes und das geplante Set wurden dem IVRS-Anbieter auf gesichertem Weg zur Verfügung gestellt. Die Patienten wurden mithilfe des IVRS randomisiert.

Item ^a	Charakteristikum	Studieninformation
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a.) und b.) treffen zu Die Patienten und Prüfarzte wurden gegenüber der in Teil A verabreichten Studienmedikation (Aclidiniumbromid oder Placebo) verblindet. Um die Sicherheit und das Wohlbefinden des Patienten zu gewährleisten, konnte der Prüfarzt die Verblindung im Notfall über das IVRS aufheben. Im Falle einer Entblindung wurde das Forest Research Institute umgehend informiert und es wurde eine umfassende schriftliche Erklärung vorgelegt. Nachdem der Code am Prüfzentrum offengelegt wurde, wurde der Patient sofort von der Studie ausgeschlossen.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Alle Inhalatoren waren äußerlich identisch und enthielten die gleiche Zusammensetzung; die Placebo-Inhalatoren enthielten allerdings keinen Wirkstoff. Der Wirkstoff war im Hinblick auf Geschmack, Erscheinung, Geruch oder Farbe nicht zu identifizieren, so dass die Verblindung nicht durchbrochen werden konnte.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Wirksamkeit: Die Wirksamkeitsanalysen basierten auf den primären LOCF-Werten der ITT-Population. Der primäre Wirksamkeitsparameter (Veränderung ggü. Baseline beim morgendlichen Ausgangs-FEV ₁ [trough FEV ₁] in Woche 12) wurden mithilfe eines Kovarianzanalyse-Modells (ANCOVA) analysiert; dieses umfasste als Faktoren das Geschlecht und die Behandlungsgruppen und als Kovariaten das Baseline-FEV ₁ und das Alter. Sicherheit: Alle Sicherheitsparameter wurden deskriptiv ausgewertet. Die Sicherheitsanalysen basierten auf der Sicherheitspopulation.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Für die Primärwirksamkeit wurde ein Modell mit gemischten Effekten (Mixed-Effects Model) für wiederholte Messungen auf Basis beobachteter Fälle (OC) eingesetzt und eine Per-Protocol-Analyse (PP) anhand der LOCF-Methode als Empfindlichkeitsanalyse der ITT-Population durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante	<ul style="list-style-type: none"> • Randomisierung: 200 µg Aclidiniumbromid BID (184), 400 µg Aclidiniumbromid BID (178), Placebo BID (182) • Geplante Behandlung erhalten: 200 µg Aclidiniumbromid BID (183), 400 µg Aclidiniumbromid BID (177), Placebo BID (182) • Analyse primärer Ergebnisse: 200 µg Aclidiniumbromid BID

Item ^a	Charakteristikum	Studieninformation
	Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	(182), 400 µg Acclidiniumbromid BID (177), Placebo BID (182)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • 200 µg Acclidiniumbromid BID: 29; Widerruf der Einwilligung: 12; Unerwünschtes Ereignis: 3; unzureichendes therapeutisches Ansprechen: 3; COPD-Exazerbation: 1; andere: 3; Verstoß gegen den Prüfplan: 4; Lost-To-Follow-Up: 3 • 400 µg Acclidiniumbromid BID: 30; Widerruf der Einwilligung: 3; Unerwünschtes Ereignis: 8; unzureichendes therapeutisches Ansprechen: 2; COPD-Exazerbation: 6; andere: 3; Verstoß gegen den Prüfplan: 3; Lost-To-Follow-Up: 2 • Placebo BID: 31; Widerruf der Einwilligung: 8; Unerwünschtes Ereignis: 4; unzureichendes therapeutisches Ansprechen: 6; COPD-Exazerbation: 4; andere: 3; Verstoß gegen den Prüfplan: 3; Lost-To-Follow-Up: 3
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Datum des Studienbeginns (erster Kontakt des ersten Patienten): 21. Dezember 2009 (Teil A) Datum Studienende (letzter Besuch des letzten Patienten): 13. September 2010 (Teil A)
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		

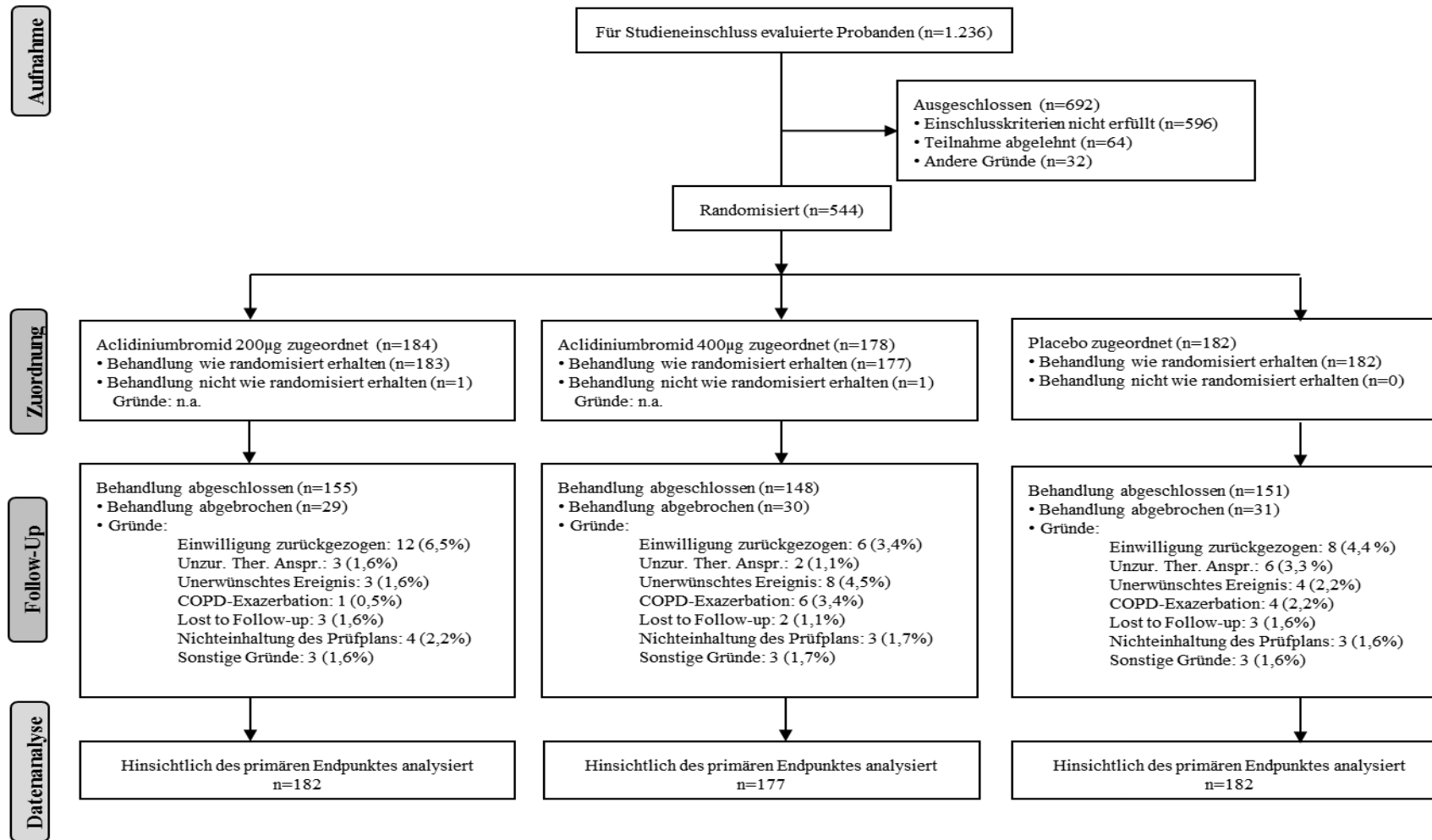


Tabelle 4.7-9 (Anhang): Studiendesign und -methodik für Studie Brusasco 2003

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	• Dokumentation von Exazerbationen und der Nutzung von Gesundheitsressourcen durch COPD-Patienten während einer 6-monatigen Behandlung mit Tiotropium, Salmeterol oder entsprechenden Placebos.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Zwei randomisierte, placebokontrollierte, Doppel-Dummy Studien, Parallelgruppen, Zuteilungsverhältnis n. a.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<ul style="list-style-type: none"> • Relativ stabile Atemwegsobstruktion mit einem forcierten Expirationsvolumen in einer Sekunde (FEV₁) ≤65 % des prognostizierten Normalwerts und ≤70 % der forcierten Vitalkapazität (FVC); über 40 Jahre alt; (ehemalige) Raucher mit mindestens 10 Packungsjahren • Kein Asthma, keine allergische Rhinitis, keine Atopie und/oder keine erhöhte Eosinophilenzahl in der Anamnese • Keine Behandlung mit zusätzlichem Sauerstoff und keine Infektion der oberen Atemwege in den sechs Wochen vor dem Screening
4b	Studienorganisation und Ort der Studiendurchführung	<p>18 Länder. Länder/Regionen in der Publikation nicht angegeben.</p> <p>Sponsor: Boehringer Ingelheim Pharma GmbH & Co. KG</p> <p>Einzelheiten zum publizierenden Autor: Dr. V. Brusasco, Dipartimento di Medicina Interna, Università di Genova, Viale Benedetto XV 6, 16132 Genova, Italien; E-Mail: brusasco@dism.unige.it</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	18 mg Tiotropiumbromid (via HandiHaler) einmal täglich plus Placebo (via Treibgasdosieraerosol (MDI)), 50 mg Salmeterol (via MDI) zweimal täglich plus Placebo (via HandiHaler) oder eine Kombination aus Placebos.
6	Zielkriterien	

Item ^a	Charakteristikum	Studieninformation
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Wirksamkeit:</p> <ul style="list-style-type: none"> • Die spirometrischen Untersuchungen wurden gemäß ATS-Richtlinien vor Beginn der Behandlung 60 und 10 Minuten vor Verabreichung der Dosis beim Randomisierungsbesuch durchgeführt, sowie 30, 60, 120 und 180 Minuten nach der Verabreichung. Die Tests wurden mit den gleichen Zeitintervallen nach 2, 8, 16 und 24 Behandlungswochen wiederholt. Eine der beiden Studien umfasste spirometrische Untersuchungen bis zu 12 Stunden nach der morgendlichen Dosis an den bereits erwähnten Tagen. • Die Fragebögen zur Beurteilung von Dyspnoe und gesundheitsbezogener Lebensqualität (HRQoL) wurden zur Baseline und 8, 16 und 24 Wochen nach der Therapie ausgefüllt. Dyspnoe wurde mithilfe des Baseline Dyspnoea-Index (BDI) und des Transitional Dyspnoea Index (TDI) bewertet. Die HRQoL wurde mithilfe des St. George's Respiratory Questionnaire (SGRQ) beurteilt. • COPD-Exazerbationen, stationäre Aufnahmen, Begleitmedikamente, ungeplante Kontakte mit Ärzten und anderen Dienstleistern im Gesundheitswesen, Arbeitsunfähigkeitstage und Beschäftigungsstatus <p>Sicherheit:</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse, Laboruntersuchungen, Elektrokardiogramme und körperliche Untersuchungen <p>Informationen zur Validierung oder Optimierung der Ergebnismessung wurden nicht berichtet. Für die Ergebnisse wurden jedoch Literaturverweise zur Verfügung gestellt. Diese umfassen u. U. Informationen zur Validierung der Ergebnismessungen.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	n. a.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung,	n. a.

Item ^a	Charakteristikum	Studieninformation
	Stratifizierung)	
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde, Doppel-Dummy Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	18 mg Tiotropiumbromid einmal täglich plus MDI-Placebo, 50 mg Salmeterol zweimal täglich plus HandiHaler-Placebo oder eine Kombination aus Placebos.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Für die Spirometrie-Messungen, den SGRQ-Gesamtpunktwert und den medianen TDI-Wert wurde eine Kovarianzanalyse durchgeführt; bei dieser wurde die Baseline als Kovariate genutzt. Mit Hilfe der beiden angrenzenden Messwerte desselben Patienten an demselben Tag wurde durch lineare Interpolation die zufälligen, mittleren und fehlenden Spirometrie-Messwerte geschätzt. Für Werte, die aufgrund der Einnahme von Notfallmedikation am Ende des Profils fehlten, wurde der an dem Tag beobachtete FEV ₁ -Minimalwert als Schätzwert verwendet. Der letzte verfügbare Wert wurde als Schätzwert für die Daten eingesetzt, die aus Gründen fehlten, die nicht mit dem Ansprechen des Patienten auf die Behandlung in Verbindung standen. Für den SGRQ wurden einzelne fehlende Antworten gemäß der vom

Item ^a	Charakteristikum	Studieninformation
		<p>Entwickler des Fragebogens festgelegten Richtlinien unterstellt. Daten aus fehlenden SGRQ-Fragebögen wurden mithilfe der LOCF-Methode ergänzt. Fehlende TDI-Antworten wurden mittels LOCF-Methode unterstellt, sofern der Patient die Studie nicht aufgrund einer sich verschlechternden COPD abbrach; in diesem Fall wurde der schlechteste Beobachtungswert eingesetzt. Der prozentuale Anteil der Patienten mit einer Verbesserung von mindestens vier Einheiten beim SGRQ-Gesamtpunktwert, einer Verbesserung des medianen TDI von mindestens einer Einheit, unerwünschten Ereignissen, stark erhöhten Steroiddosen aufgrund von COPD-Exazerbationen und einer zusätzlichen Theophyllin-Behandlung aufgrund von COPD-Exazerbationen wurde mithilfe des exakten Fisher-Tests analysiert. Alle Therapiegruppen wurden paarweise separat analysiert. Die Patientenanteile mit mindestens einer COPD-Exazerbation und der Anteil mit mindestens einem Krankenhausaufenthalt aufgrund einer COPD Exazerbation wurden mit logistischer Regression, angepasst an den Grad der Exposition, analysiert. Die Kombination der beiden 6-monatigen Studien zwecks Bewertung der Exazerbationen wurde a priori in den Prüfplänen festgelegt. Die beiden Studien waren bis auf den Endpunkt FEV₁ AUC (nach 3 bzw. 12 Studen) identisch.</p> <p>Die Dauer bis zur ersten COPD-Exazerbation und die Dauer bis zur ersten stationären Aufnahme aufgrund einer COPD-Exazerbation wurden mit einem Log-Rank-Test analysiert. Die Anzahl der Exazerbationen, die Anzahl der Tage mit Exazerbationen, die Anzahl der stationären Aufnahmen aufgrund einer Exazerbation und die Anzahl der Tage im Krankenhaus wurden mit dem Wilcoxon-Mann-Whitney-Test analysiert. In allen diesen Analysen wurden die einzelnen Therapiegruppen paarweise separat analysiert. Fehlende Werte bei stationären Aufnahmen aus beliebigem Grund, ungeplanten Arztbesuchen und Tagen, an denen keine alltäglichen Aktivitäten möglich waren, wurden mithilfe multipler Zuschreibung unter Verwendung der Markov Chain Monte-Carlo Methode geschätzt. Die Anzahl der stationären Aufnahmen und der Tage im Krankenhaus aus beliebigen Gründen, die Anzahl der ungeplanten Arztbesuche und die Anzahl der Tage, an denen keine alltäglichen Aktivitäten durchgeführt werden konnten, wurden mittels einer Varianzanalyse unter Einbeziehung von Behandlungsdauer und Zentrum untersucht. Der Anteil der Patienten mit mindestens einer stationären Aufnahme wurde mithilfe einer normalen Annäherung an die Binomialverteilung analysiert.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	

Item ^a	Charakteristikum	Studieninformation
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Randomisiert: 18 µg Tiotropiumbromid einmal täglich (402), 50 µg Salmeterol BID (405), Placebo (400) • Geplante Behandlung erhalten: n. a. • Im Hinblick auf das Primärergebnis analysierte Patienten: n. a.
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • 18 µg Tiotropiumbromid einmal täglich: 62; Unerwünschtes Ereignis: 29; kein Grund für Abbruch angegeben: 33 • 50 µg Salmeterol BID: 76; Unerwünschtes Ereignis: 60; kein Grund für Abbruch angegeben: 16 • Placebo: 103; Unerwünschtes Ereignis: 64; kein Grund für Abbruch angegeben: 39
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

t

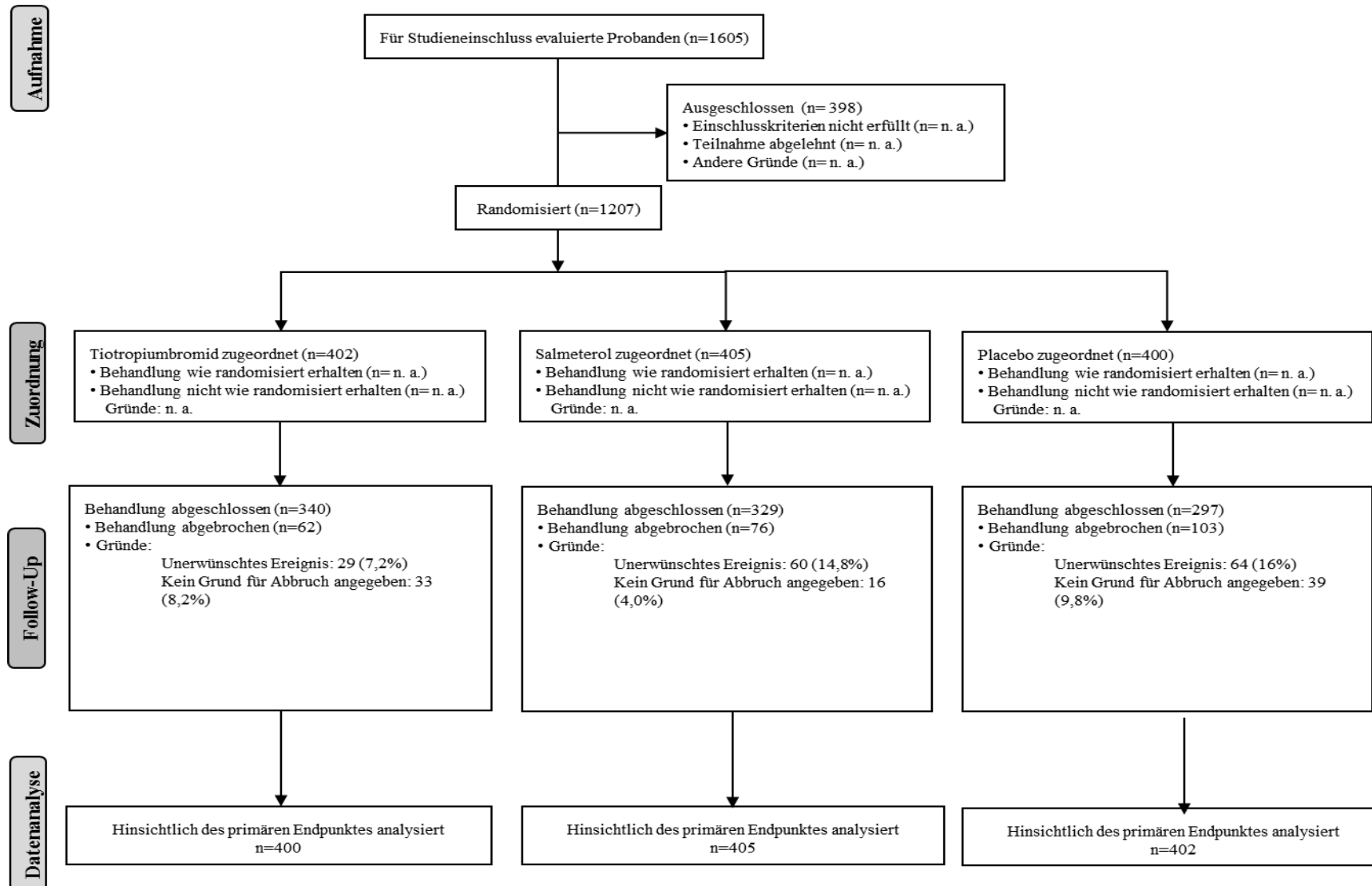


Tabelle 4.7-10 (Anhang): Studiendesign und -methodik für Studie Casaburi 2000

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	• Vergleich der bronchodilatatorischen Wirksamkeit und Sicherheit von Tiotropiumbromid und Placebo.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, doppelblinde, placebokontrollierte Parallelgruppenstudie Randomisierungsverhältnis von Tiotropiumbromid und Placebo 3:2.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<ul style="list-style-type: none"> • Ambulant behandelte Patienten beider Geschlechter im Alter von ≥ 40 Jahren mit einer klinischen COPD-Diagnose gemäß Definition der American Thoracic Society (ATS) • (Ehemalige) Raucher mit >10 Packungsjahren und stabiler Obstruktion der Atemwege, $FEV_1 \leq 65\%$ des prognostizierten Normalwerts und $FEV_1 \leq 70\%$ der FVC • Kein Asthma, keine allergische Rhinitis, keine Atopie und keine erhöhte Eosinophilenzahl $\geq 600/\text{mm}^3$ in der Vergangenheit • Keine regelmäßige Behandlung mit Sauerstoff am Tag oder Einnahme von mehr als dem Äquivalent von 10 mg Prednison pro Tag zur Kontrolle der COPD-Symptome in dem Monat vor Eintritt in die Studie • Kein Myokardinfarkt (≤ 1 Jahr), Herzinsuffizienz (≤ 3 Jahre) oder Herzrhythmusstörungen, die eine medikamentöse Therapie erfordern
4b	Studienorganisation und Ort der Studiendurchführung	Fünfundzwanzig Zentren Sponsor: Boehringer Ingelheim GmbH & co. KG (BIPI) Einzelheiten zum publizierenden Autor: Richard Casaburi, PhD, MD, Harbor-UCLA Medical Center, 1000 W Carson St, Torrance, CA 90509; E-Mail: casaburi@ucla.edu
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Die Probanden nahmen die Studienmedikation durch Inhalation einmal täglich zwischen 8:00 und 10:00 Uhr ein. Der aktive Wirkstoff und Placebo wurden in identisch aussehenden laktosebasierten Inhalatoren zur Verfügung gestellt (HandiHaler BIPI, Ridgefield, CT).

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt: • FEV₁-Ansprechen bei der abschließenden Therapieuntersuchung (92 Tage nach Therapiebeginn)</p> <p>Zusätzliche Wirksamkeitsendpunkte: • FEV₁-Ansprechen in den 3 Stunden nach Verabreichung der Studienmedikation, das Trough-FVC-Ansprechen (wie für FEV₁ definiert) und das FVC-Ansprechen in den ersten drei Stunden nach Verabreichung der Studienmedikation</p> <p>• Das Peak- (maximaler Wert) und Durchschnitts- (Mittel aller Werte) FEV₁- und FVC-Ansprechen (d. h. Veränderung ggü. Baseline) während der ersten 3 Stunden nach Verabreichung der Studienmedikation</p> <p>• Wöchentlicher maximaler expiratorischer Fluss (peak expiratory flow rates - PEFR), COPD-Symptompunktwerte, globaler Bewertungspunktwert des Arztes und ggf. Verwendung von Albuterol</p> <p>Alle Spirometrie-Messungen wurden dreimal ausgeführt; die höchsten FEV₁- und FVC-Werte wurden aufgezeichnet und für die nachfolgenden Analysen verwendet.</p> <p>Die in diesen Studien verwendeten Spirometer mussten die ATS-Vorgaben erfüllen.</p> <p>Um einheitliche Bedingungen für die Tage zu gewährleisten, an denen Lungenfunktionstests durchgeführt wurden, mussten die Probanden die Behandlung mit Theophyllin-Präparaten 24 Stunden vor dem Lungenfunktionstest absetzen.</p> <p>Informationen über die Validierung der Ergebnismessung wurden nicht mitgeteilt. Für die Ergebnisse wurden jedoch Literaturverweise zur Verfügung gestellt. Diese umfassen u. U. Informationen zur Validierung der Ergebnismessungen.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	n. a.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der	n. a.

Item ^a	Charakteristikum	Studieninformation
	zufälligen Zuteilung	
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierungsverhältnis von Tiotropiumbromid und Placebo 3:2.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Nach einer zweiwöchigen Baseline-Periode zum Nachweis und zur Dokumentation der klinischen Stabilität wurden die Patienten innerhalb der einzelnen Prüfzentren (in der Reihenfolge des Eintritts) randomisiert und erhielten entweder Tiotropiumbromid (18 mg) oder Placebo.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Der aktive Wirkstoff und Placebo wurden in identisch aussehenden laktosebasierten Inhalatoren zur Verfügung gestellt (HandiHaler BIPI, Ridgefield, CT).
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Alle statistischen Vergleiche zwischen den beiden Behandlungsgruppen wurden mittels einer Kovarianzanalyse durchgeführt. Wenn die Patienten die Studie abbrechen mussten, weil sich ihre COPD verschlechterte, wurden ihre fehlenden Wirksamkeitsdaten geschätzt. Dafür wurden die am wenigsten positiven Daten verwendet, die vor Beendigung der Studie ermittelt wurden. Für andere Patienten, die Studienbesuche aus anderen Gründen verpassten, wurden fehlende Datenwerte anhand der letzten beim

Item ^a	Charakteristikum	Studieninformation
		Patienten beobachteten Daten geschätzt. Für Patienten, die nicht alle pulmonalen Funktionsmessungen an einem spezifischen Lungenfunktionstest-Tag durchlaufen konnten, wurden fehlende mittlere Spirometrie-Messwerte mittels linearer Interpolation geschätzt. Ebenso wurden die an einem spezifischen Tag beobachteten minimalen Spirometrie-Messwerte verwendet, um die Werte am Ende eines Profils zu schätzen, die aufgrund der Behandlung mit Notfallmedikation fehlten. Schlussendlich wurden die letzten verfügbaren Spirometrie-Messwerte verwendet, um die Werte am Ende eines Profils zu schätzen, die aus Gründen fehlten, die nicht mit dem Ansprechen des Patienten auf die Behandlung in Verbindung standen.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Randomisiert: Tiotropiumbromid 18 µg einmal täglich (279), Placebo (191) • Geplante Behandlung erhalten: n. a. • Im Hinblick auf das Primärergebnis analysierte Patienten: n. a.
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • 18 µg Tiotropiumbromid einmal täglich: 17; unerwünschtes Ereignis: 7; mangelnde Wirksamkeit: 5; kein Grund für Abbruch angegeben: 5 • Placebo: 21; unerwünschtes Ereignis: 12; mangelnde Wirksamkeit: 6; kein Grund für Abbruch angegeben: 3
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie	Geplant

Item ^a	Charakteristikum	Studieninformation
	endete oder beendet wurde	
a: nach CONSORT 2010. n. a. = nicht angegeben		

t

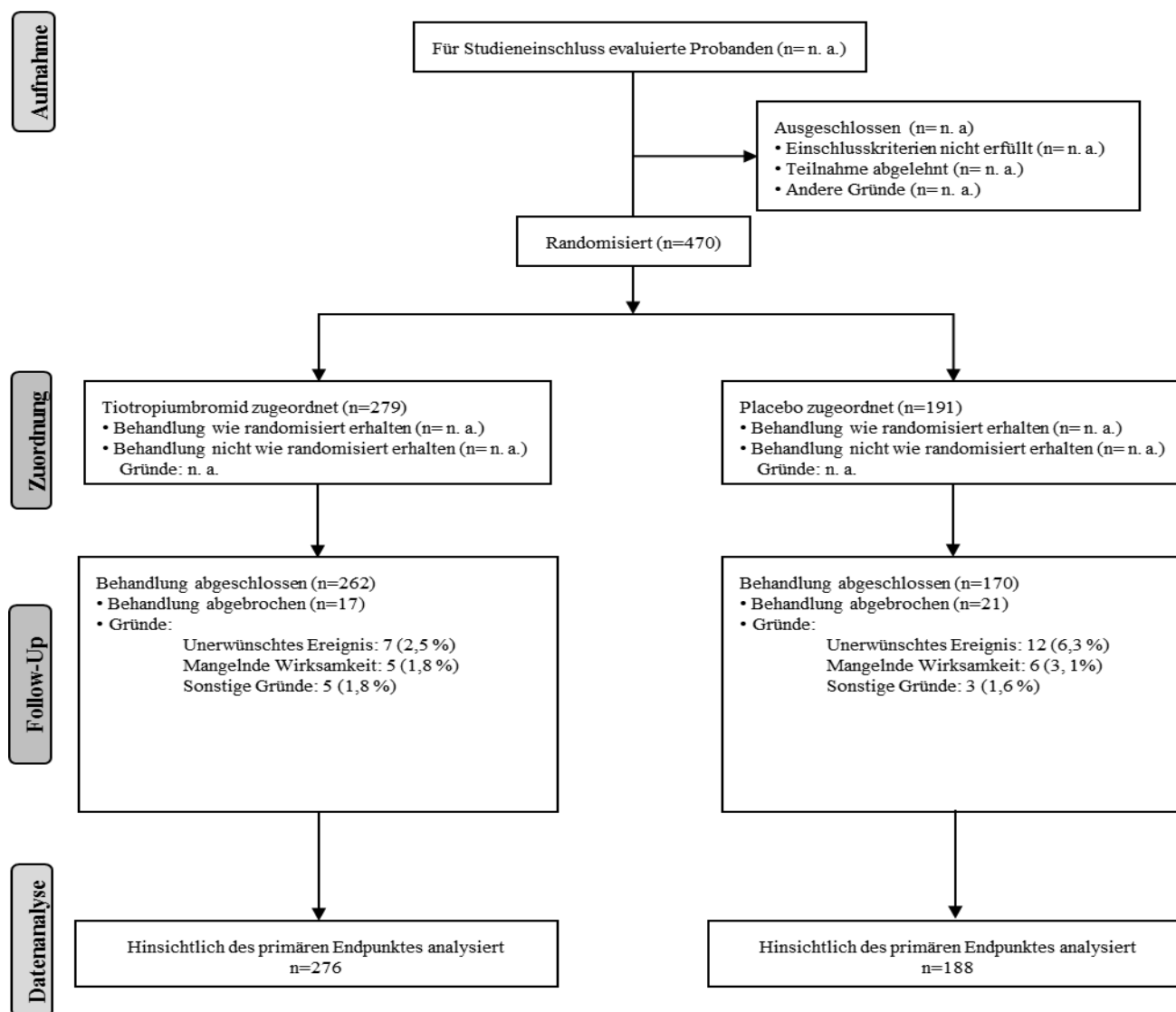


Tabelle 4.7-11 (Anhang): Studiendesign und -methodik für Studie Casaburie 2002

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	• Evaluierung der langfristigen Sicherheit und Wirksamkeit von Tiotropium
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Eine der beiden präsentierten Studien ist die Langzeit-Verlängerungsstudie der von Casaburi 2000 publizierten 13-Wochen-Studie. Doppelblinde, placebokontrollierte Parallelgruppenstudien, randomisiert im Verhältnis 3:2 von Tiotropium oder Placebo.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<ul style="list-style-type: none"> • Ambulant behandelte Patienten beider Geschlechter im Alter von ≥ 40 Jahren mit einer klinischen COPD-Diagnose gemäß Definition der American Thoracic Society (ATS) • Ehemalige und aktuelle Raucher mit >10 Packungsjahren mit stabiler Obstruktion der Atemwege, forciertes Expirationsvolumen/Sekunde (Einsekundenkapazität, FEV₁) ≤ 65 % des prognostizierten Normalwerts und FEV₁ ≤ 70 % der forcierten Vitalkapazität (FVC) • Kein Asthma, keine allergische Rhinitis, keine Atopie und keine erhöhte Eosinophilenzahl von ≥ 600 Zellen/mm³ in der Vergangenheit. • Keine regelmäßige Behandlung mit Sauerstoff am Tag oder Einnahme von mehr als dem Äquivalent von 10 mg Prednison pro Tag in dem Monat vor Eintritt in die Studie • Kein kürzlicher Myokardinfarkt (≤ 1 Jahr), Herzversagen (≤ 3 Jahre) oder Herzrhythmusstörungen, die eine medikamentöse Therapie erfordern
4b	Studienorganisation und Ort der Studiendurchführung	50 Zentren Sponsor: Boehringer Ingelheim Pharma GmbH & Co. KG Einzelheiten zum publizierenden Autor: R. Casaburi, Harbor-UCLA Research and Education Institute, Building RB-2, 1124 W. Carson Street, Torrance, CA 90502, USA. E-Mail: casaburi@ucla.edu
5	Interventionen Präzise Angaben zu den geplanten	Die Probanden inhalierten den aktiven Wirkstoff (Tiotropiumbromid in Laktose) oder Placebo (Laktose) einmal am Morgen mittels identisch aussehender Kapseln durch einen

Item ^a	Charakteristikum	Studieninformation
	Interventionen jeder Gruppe und zur Administration etc.	Trockenpulverinhalator (HandiHaler, Boehringer Ingelheim; Ingelheim am Rhein, Deutschland)
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <ul style="list-style-type: none"> • Trough-FEV₁ <p>Zusätzliche Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • Trough-FEV₁ und FVC an Behandlungstag 1 und nach 1, 7, 13, 25, 37 und 49 Therapiewochen. • Veränderung der Dyspnoe, festgestellt durch den Transitional Dyspnoea Index (TDI) und Veränderung des Gesundheitsstatus, festgestellt durch den krankheitsspezifischen St. George's Respiratory Questionnaire und das generische Short Form 36. <p>Sicherheit:</p> <ul style="list-style-type: none"> • Medikamentöse Behandlung und unerwünschte Ereignisse <p>Alle Spirometrie-Messungen wurden dreimal ausgeführt und die höchsten FEV₁- und FVC-Werte wurden für die nachfolgenden Analysen verwendet. Die Spirometer mussten die ATS-Vorgaben erfüllen.</p> <p>Um die standardisierten Bedingungen an den Tagen der Spirometrie-Messungen zu gewährleisten, mussten die Probanden die Behandlung mit Theophyllin 24 Stunden vor der Spirometrie einstellen.</p> <p>Informationen über die Validierung der Ergebnismessung wurden nicht mitgeteilt. Für die Ergebnisse wurden jedoch Literaturverweise zur Verfügung gestellt. Diese umfassen u. U. Informationen zur Validierung der Ergebnismessungen.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Es wurde eine Gruppengröße von 60 Patienten pro Gruppe berechnet, um bei einer Standardabweichung von 0,215 l einen Unterschied von 0,13 l beim Trough-FEV₁ zu detektieren.</p> <p>Um eine adäquate Sicherheitsexposition zu gewährleisten, wurde für jede Studie eine Gruppengröße von insgesamt 240 Patienten ausgewählt.</p>

Item ^a	Charakteristikum	Studieninformation
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Siehe Bericht über Studie Casaburie 2000 (Zwischenanalyse einer der beiden Studien)
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Patienten wurden mit einem Verhältnis von Tiotropiumbromid: Placebo von 3:2 randomisiert.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Der aktive Wirkstoff (Tiotropiumbromid in Laktose) oder Placebo (Laktose) wurden in Form von äußerlich identischen Kapseln eingenommen
12	Statistische Methoden	

Item ^a	Charakteristikum	Studieninformation
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Das Statistikmodell war eine Kovarianzanalyse mit der Behandlungsdauer sowie klinischen und Baseline-Werten als Kovariaten. Die Abbruchraten und Anteile der Patienten, die klinisch bedeutsame Reaktionen aufwiesen, wurden zwischen den Behandlungsgruppen mittels des exakten Fisher-Tests verglichen. Die Anteile der Patienten, die mindestens eine COPD-Exazerbation und eine stationäre Aufnahme im Zusammenhang mit diesen Exazerbationen aufwiesen, wurden zwischen den Behandlungsgruppen mithilfe logistischer Regressionsanalysen verglichen. Nur COPD-Exazerbationen, die als unerwünschte Ereignisse dokumentiert wurden, wurden für den Vergleich herangezogen; dadurch sollten die täglichen Symptomschwankungen eliminiert werden.</p> <p>Das Intention-To-Treat-Prinzip, alle randomisierten Patienten einzuschließen wurde in allen Wirksamkeitsanalysen angewendet.</p> <p>Die Patienten, für die Baseline-Daten oder Daten aus <2 Wochen für bestimmte Endpunkte fehlten, wurden von den Einjahresvergleichen für diese Endpunkte ausgeschlossen. Patienten, die wegen Ablauf des Verfallsdatums der Studienmedikation nicht der Lage waren, alle ihre Termine wahrzunehmen, beendeten die Studie bei dem Besuch nach neun Monaten, wurden aber trotzdem als bis zum Schluss teilnehmende Patienten angesehen.</p> <p>Brachen Patienten die Studie ab, weil sich ihre Erkrankung verschlimmerte (5 % der Patienten), wurden fehlende Daten durch den am wenigsten positiven Wert ergänzt, der vor dem Abbruch beobachtet wurde. In allen anderen Fällen wurden aufgrund von Studienabbruch fehlende Daten mit dem letzten beobachteten Wert ergänzt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Um die Auswirkungen der Datenschätzungen zu bewerten, wurden die Daten außerdem anhand der Patienten analysiert, die die Studie von sich aus nach 12 Monaten beendeten.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben,	<ul style="list-style-type: none"> • Randomisiert: 18 µg Tiotropiumbromid einmal täglich (550), Placebo (371) • Erhielten die vorgesehene Behandlung: n. a. • Im Hinblick auf das Primärergebnis analysierte Patienten: n. a.

Item ^a	Charakteristikum	Studieninformation
	c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • 18 µg Tiotropiumbromid einmal täglich: 103 Unerwünschtes Ereignis: 53; mangelnde Wirksamkeit: 13, sonstige Gründe: 37 <ul style="list-style-type: none"> • Placebo: 103 Unerwünschtes Ereignis: 51; mangelnde Wirksamkeit: 26, sonstige Gründe: 26 24 % der Patienten brachen die Studie vorzeitig ab, weil das Verfallsdatum der Studienmedikation überschritten war. Es wurde als unnötig erachtet, dem Probanden eine neue Studienmedikation zukommen zu lassen; Grund dafür war, dass die angestrebte Populationsgröße erreicht war, da weniger Patienten als erwartet die Studie vorzeitig abbrachen.
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010. n. a. = nicht angegeben		

t

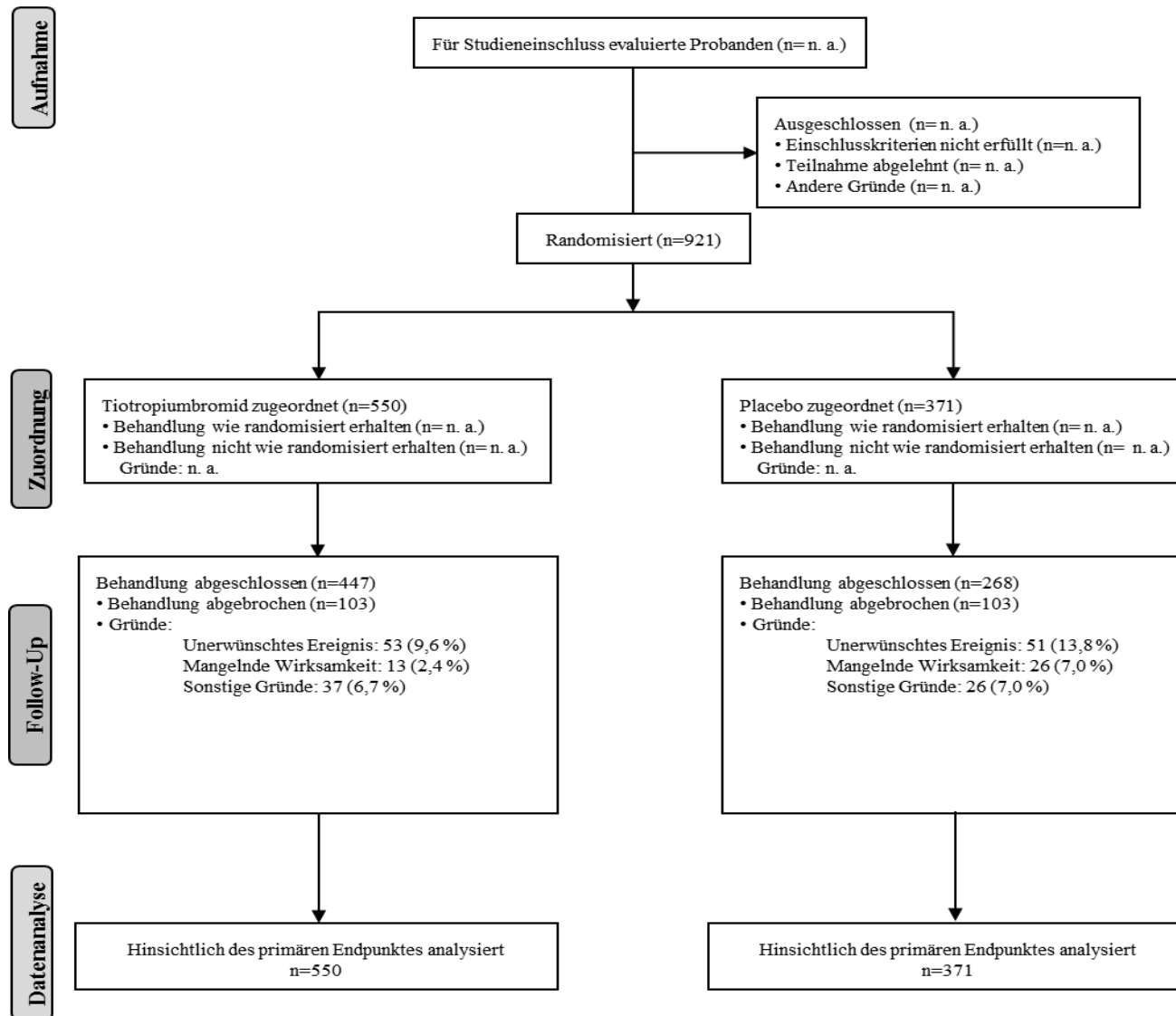


Tabelle 4.7-12 (Anhang): Studiendesign und -methodik für Studie SAFE (Chan 2007)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Beurteilung der Wirkung von Tiotropiumbromid auf den (Trough) FEV ₁ bei COPD-Patienten in Kanada
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, placebokontrollierte Parallelgruppenstudie, Randomisierungsverhältnis von 18 µg Tiotropiumbromid einmal täglich zu Placebo von 2:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Ein Einschlusskriterium wurde verändert: Statt Patienten mit mindestens einer Exazerbation im letzten Jahr (aber nicht in den letzten 6 Wochen) wurden auch Patienten mit mindestens einer Exazerbation während der letzten zwei Jahre zugelassen.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<ul style="list-style-type: none"> • Männliche und weibliche ambulant behandelte Patienten im Alter von mindestens 40 Jahren mit klinischer COPD-Diagnose (FEV₁ maximal 65 % des prognostizierten Wertes und FEV₁/forcierte Vitalkapazität [FVC] maximal 70 %) • Ehemalige und aktuelle Raucher mit mindestens zehn Packungsjahren • Anfangs wurde vorausgesetzt, dass die Patienten im vergangenen Jahr mindestens eine Exazerbation erlitten (Behandlung mit Antibiotika und/oder oralen Steroiden erforderlich), jedoch nicht in den letzten sechs Wochen vor Eintritt in die Studie. Später wurden auch Patienten zugelassen, die in den vergangenen zwei Jahren eine Exazerbation erlitten. • Kein Asthma, keine allergische Rhinitis und keine Atopie in der Vergangenheit • Keine Infektion der tiefen Atemwege und keine Exazerbation (in den letzten sechs Wochen vor Studienbeginn) • Kein kürzlicher Myokardinfarkt (innerhalb der vergangenen sechs Monate) oder Herzrhythmusstörungen, die eine medikamentöse Therapie erfordern • Keine Verabreichung oraler Corticosteroide in instabiler Dosis während der sechs Wochen vor Eintritt in die Studie oder in stabiler Dosis > 10 mg Prednison pro Tag
4b	Studienorganisation und Ort der Studiendurchführung	<p>Die Studie wurde in 101 Zentren in Kanada durchgeführt</p> <p>Sponsoren: Boehringer Ingelheim Canada Ltd und Pfizer Canada. Einzelheiten zum publizierenden Autor: Dr. Charles Chan, University Health Network – Toronto General Hospital, 585 University Avenue, Toronto, Ontario, Kanada, M5G 2N2. E-Mail: charles.chan@uhn.on.ca</p>

Item ^a	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	18 µg Tiotropiumbromid einmal täglich oder Placebo mit dem HandiHaler (Boehringer Ingelheim Deutschland).
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <ul style="list-style-type: none"> • Morgendlicher (Trough) FEV₁ am Studienende (48 Wochen) • Spirometrien wurden beim Screening, an Tag 1 der Behandlung (Randomisierung) und nach 2, 11, 30 und 48 Behandlungswochen durchgeführt. • FEV₁, FVC und FEV₆ (optional für qualifizierte Prüferzentren) 10 Minuten vor der Dosisgabe (Trough) • St. George's Respiratory Questionnaire (SGRQ) (erhoben vor der Dosierung an Tag 1 sowie nach 30 und 48 Behandlungswochen) • Exazerbationen und damit verbundene Klinikaufenthalte, die Anzahl verabreichter Zyklen von oralen Steroiden und Antibiotika zur Behandlung von Exazerbationen <p>Sicherheit:</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse <p>Die Spirometrie wurde gemäß der Vorgaben der American Thoracic Society in der Arztpraxis durchgeführt.</p> <p>Die Gerätetypen waren nicht einheitlich, mussten aber von der American Thoracic Society genehmigt worden sein. Alle spirometrischen Untersuchungen wurden dreimal durchgeführt und die höchsten akzeptablen Messwerte wurden für die nachfolgenden Analysen verwendet. Um einheitliche Bedingungen für die Tage zu gewährleisten, an denen spirometrische Untersuchungen durchgeführt wurden, mussten die Probanden die Behandlung mit kurzwirksamen Beta-2-Agonisten 6 Stunden vor dem Klinikbesuch absetzen, kurzwirksame Theopyllin-Präparate 24 Stunden vor der Untersuchung und lang wirksame Beta-2-Rezeptor-Agonisten (LABA) sowie langwirksame Theopyllin-Präparate 48 Stunden vor der Untersuchung.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Eine Gesamtpopulation von 900 Patienten (Tiotropium: 600 Patienten; Placebo: 300 Patienten) wurde unter der Annahme, bei der Analyse des primären Endpunkts (FEV ₁ Trough-Reaktion) auf Grundlage einer Standardabweichung von 0,215 l mit einer Aussagekraft von 90 % einen Unterschied zwischen den Gruppen von 0,13 l nachzuweisen (bei einem Alpha-

Item ^a	Charakteristikum	Studieninformation
		Konfidenzintervall von 5 %). Die anfängliche Fallzahlplanung umfasste drei Vergleichsgruppen: Starke Raucher, weniger starke Raucher und Ex-Raucher. Aufgrund der geringer als erwarteten Rekrutierung wurde jedoch eine Änderung vorgenommen und die Anzahl der Vergleichsgruppen von drei auf zwei zu reduzieren (Raucher und Ex-Raucher).
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	N. z.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierungsverhältnis von 18 µg Tiotropiumbromid einmal täglich zu Placebo von 2:1
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die	Doppelblinde Studie

Item ^a	Charakteristikum	Studieninformation
	Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	18 µg Tiotropiumbromid einmal täglich oder Placebo mit dem HandiHaler
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Zwischen den Behandlungsgruppen wurden Veränderungen in Bezug auf folgende Punkte verglichen: FEV ₁ , FVC, FEV ₆ , SGRQ-Gesamtpunktwert und Unterpunktwerte, Behandlung mit Notfallmedikation. Dafür wurde eine ANCOVA-Analyse durchgeführt, mit der Behandlung und dem Baseline-Wert als Kovariaten. Die Berechnungen erfolgten in der Full-Analysis-Set-Population, wobei fehlende Werte mit der LOCF-Methode ergänzt wurden. Ausgenommen waren Fälle, in denen Patienten die Studie aufgrund sich verschlechternder COPD abbrachen. In diesen Fällen wurden die am wenigsten positiven Daten übernommen, die vor dem Abbruch erhoben wurden. Die Anzahl (%) der Patienten mit mindestens einer COPD-Exazerbation und die Anzahl (%) der Patienten mit mindestens einem Klinikaufenthalt aufgrund einer COPD-Exazerbation wurde zwischen den Behandlungsgruppen mithilfe des exakten Fisher Tests (zweiseitig) verglichen. Die Anzahl von COPD-Exazerbationen, Tagen mit COPD-Exazerbationen, stationären Aufnahmen aufgrund einer COPD-Exazerbation, Tagen im Krankenhaus aufgrund einer COPD-Exazerbation und die Anzahl der Behandlungszyklen mit Steroiden/Antibiotika während einer COPD-Exazerbationsepisode wurden zwischen den Behandlungsgruppen mithilfe des Poisson-Regressionsmodells unter Einbeziehung des Expositionsumfanges (abzüglich der Dauer der Ereignisse) verglichen.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Wie im Prüfplan festgehalten, wurde eine schrittweise weitere Analyse des Unterschieds beim Trough-FEV ₁ nach 48 Wochen in der Raucher- und Ex-Raucher-Gruppe geplant, sofern der Vergleich der Behandlungsgruppen ein positives Ergebnis ergab. Univariate Analysen, einschließlich einfaktorieller ANOVA-Analysen und exakter Fisher Tests (zweiseitig), wurden jeweils auf kontinuierliche und kategoriale Variablen angewandt, um beim Screening die Merkmale von Rauchern und Ex-Rauchern zu untersuchen. In den Gruppen der Raucher und Ex-Raucher wurde die Veränderung beim FEV ₁ und der SGRQ-Gesamtpunktzahl zwischen den beiden Behandlungsgruppen (Tiotropiumbromid und Placebo) verglichen. Dafür wurde eine ANCOVA-Analyse durchgeführt, mit der Behandlungsdauer und dem Baseline-Wert als Kovariate. Für diese Analysen wurde das Full-Analysis-Set verwendet.
-	Resultate	

Item ^a	Charakteristikum	Studieninformation
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Randomisiert: 18 µg Tiotropiumbromid einmal täglich (608), Placebo (305) • geplante Behandlung erhalten: 18 µg Tiotropiumbromid einmal täglich (608), Placebo (305) • Analyse primärer Ergebnisse: 18 µg Tiotropiumbromid einmal täglich (608), Placebo (305)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • 18 µg Tiotropiumbromid einmal täglich: 135; unerwünschtes Ereignis: 73; administrativ: 52; andere: 10 • Placebo: 84 unerwünschtes Ereignis: 39; administrativ: 42; andere: 3
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		
n. a. = nicht angegeben		

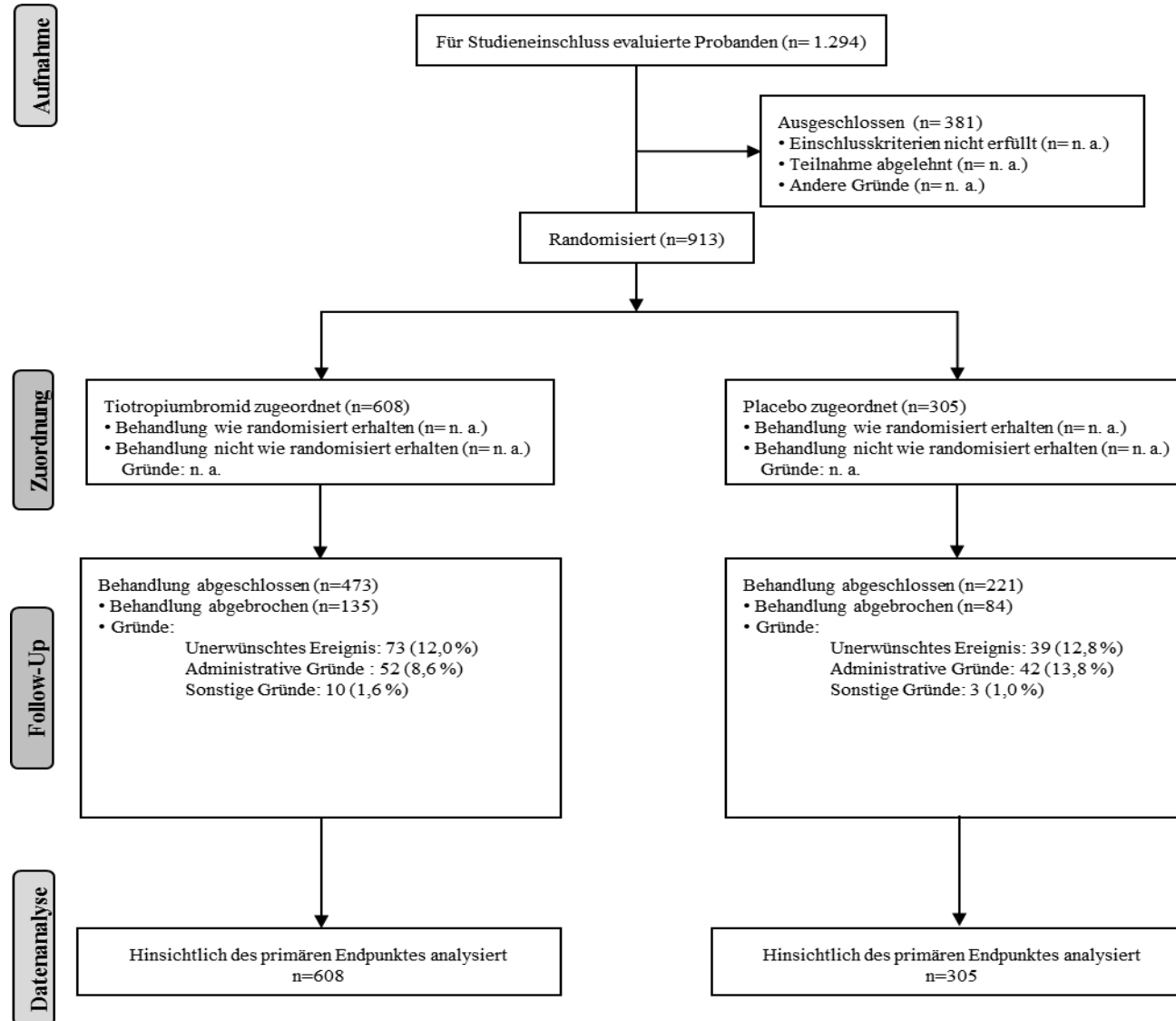


Tabelle 4.7-13 (Anhang): Studiendesign und -methodik für Studie Covelli 2005

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	<ul style="list-style-type: none"> Analyse elektrokardiographischer Befunde nach kurz- und langfristiger Tiotropiumbromid-Therapie bei Patienten mit chronisch obstruktiver Lungenerkrankung (COPD) sowie Nachweis der symptomatischen Wirksamkeit
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, doppelblinde Parallelgruppen-Studie, Gruppenzuteilung im Verhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<ul style="list-style-type: none"> Klinische Diagnose der COPD, mindestens 40 Jahre alt und (ehemalige) Raucher mit mindestens 10 Packungsjahren Ein FEV₁ maximal 60 % des prognostizierten Normalwerts, FVC maximal 70 % Keine frühere Asthma-Erkrankung oder Atopie, keine anormalen Leberenzym-Werte und keine Anzeichen einer chronischen Nierenfunktionsstörung; außerdem keine Infektion der Atemwege oder COPD-Exazerbation innerhalb der 6 Wochen vor der Randomisierung Keine systemischen Corticosteroide in instabiler Dosis, kein Prednison in einer Dosis von 10 mg/Tag oder höher (bzw. Äquivalent) und keine Sauerstoffbehandlung über mehr als 12 Stunden pro Tag Kein Myokardinfarkt innerhalb der letzten sechs Monate, kein Krankenhausaufenthalt aufgrund von Herzinsuffizienz im vergangenen Jahr und keine lebensbedrohenden Arrhythmien im letzten Jahr, die eine Behandlung oder eine Änderung der medikamentösen Therapie erforderten.
4b	Studienorganisation und Ort der Studiendurchführung	<p>Die Studie wurde in 12 Prüfzentren in den USA durchgeführt.</p> <p>Sponsor: Boehringer Ingelheim Pharmaceuticals Inc. Einzelheiten zum publizierenden Autor: Henry Covelli, M.D., 700 Ironwood Drive, Suite 360, Coeur d'Alene, USA, ID 83814;</p>

Item ^a	Charakteristikum	Studieninformation
		E-Mail: hdomc@imbris.net.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	18 µg Tiotropiumbromid oder entsprechendes Placebo, beide verabreicht mit dem HandiHaler-Inhalator (Boehringer Ingelheim GmbH & Co. KG, Ingelheim, Deutschland)
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <ul style="list-style-type: none"> • Morgendliches forciertes Expirationsvolumen in einer Sekunde (FEV₁) vor Verabreichung des Präparats (Trough) nach 12 Wochen (Tag 84) der Behandlung <p>Sekundäre Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • (Trough) FEV₁ an Tag 56 und forcierte Vitalkapazität (FVC) an den Tagen 56 und 84, FEV₁ und FVC nach Verabreichung der Dosis an allen Untersuchungstagen, globale COPD-Bewertungen durch Patient und Arzt, Punktzahl im EQ-5D-Fragebogen und Behandlung mit Notfallmedikation <p>Angewendete Methoden zur Optimierung der Datensammlung und Validierung der Sammlungsmethoden.</p> <p>Die Spirometrie wurde nach adäquatem Washout der Bronchodilatoren mit kalibrierten Spirometern durchgeführt. Die Tests wurden dreimal durchgeführt und der beste dieser drei Versuche wurde gemäß den Kriterien der American Thoracic Society als die höchste gemessene FEV₁ und das höchste FVC bei drei Exhalationen definiert.</p> <p>Die 12-Kanal-EKG wurden zentral von einem qualifizierten Kardiologen ausgewertet, dem der Therapieort nicht bekannt war. Die Daten des 24-Stunden-EKG wurden zentral mithilfe eines elektronischen Holter-Analyse-Systems analysiert und von geschulten Technikern geprüft. Zur Qualitätssicherung wurden 2 % der aufgezeichneten Daten erneut überprüft.</p> <p>Informationen über die Validierung der Ergebnismessung wurden nicht mitgeteilt. Für die Ergebnisse wurden jedoch Literaturverweise zur Verfügung gestellt. Diese umfassen u. U. Informationen zur Validierung der Ergebnismessungen.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	

Item ^a	Charakteristikum	Studieninformation
7a	Wie wurden die Fallzahlen bestimmt?	n. a.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	n. a.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie. Die 12-Kanal-EKG wurden zentral von einem qualifizierten Kardiologen (eResearchTechnology, Inc., Philadelphia, PA) ausgewertet, dem die Behandlung des Patienten nicht bekannt war. Die Daten des 24-Stunden-EKG wurden zentral mithilfe eines elektronischen Holter-Analyse-Systems analysiert und von geschulten Technikern (eResearchTechnology, Inc.) geprüft. Die Bereiche von diagnostischem Interesse wurden von einem qualifizierten Kardiologen ausgewertet, dem die Behandlungszuteilung des Patienten unbekannt war.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	18 µg Tiotropiumbromid oder Placebo, beide verabreicht mit dem HandiHaler-Inhalator (Boehringer Ingelheim GmbH & Co. KG,

Item ^a	Charakteristikum	Studieninformation
		Ingelheim, Deutschland)
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Der primäre Endpunkt des (Trough) FEV₁ nach 12 Wochen wurde mithilfe des ANCOVA-Modells analysiert; dieses umfasste als Faktoren die Therapie und die Zentren und als Kovariaten die Baseline-FEV₁-Werte.</p> <p>Die sekundären Wirksamkeitsendpunkte – alle sekundären Spirometrie-Endpunkte, globale Bewertungen durch Patient und Arzt, Punktzahl auf der visuellen Analogskala (VAS) im EQ-5D-Fragebogen und wöchentliche durchschnittliche Anzahl der Behandlungen mit Notfallmedikation pro Tag – wurden mit dem gleichen ANCOVA-Modell analysiert, das auch für die Analyse des primären Endpunkts verwendet wurde.</p> <p>Fehlten Daten zur Wirksamkeit, weil Patienten die Studie aufgrund einer sich verschlechternden COPD vorzeitig abbrachen, wurde der vor dem Abbruch schlechteste gemessene Wert eingesetzt.</p> <p>In allen anderen Fällen wurden aufgrund von Studienabbruch fehlende Wirksamkeitsdaten mit dem letzten beobachteten Wert ergänzt. Fehlende Werte im EQ-5D-Fragebogen wurden nicht ergänzt.</p> <p>Für die fünf Dimensionen des EQ-5D-Fragebogens wurden für jede Antwort deskriptive Statistiken (ausgedrückt in Zahl und Prozentzahl) angegeben (keine Probleme, einige Probleme und schwerwiegende Probleme).</p> <p>Bei der Analyse wurden die Antworten kategorisiert: „keine Probleme“ oder „irgendwelche Probleme“ (d. h. „einige Probleme“ und „schwerwiegende Probleme“). Um die Behandlungen in den fünf Dimensionen des EQ-5D-Fragebogens zu vergleichen, wurde der Wilcoxon-Rangsummentest angewendet.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die	<ul style="list-style-type: none"> • Randomisiert: 18 µg Tiotropiumbromid einmal täglich (100), Placebo (96)

Item ^a	Charakteristikum	Studieninformation
	a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Erhielten die vorgesehene Behandlung: n. a. • Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 18 µg einmal täglich (94), Placebo (84)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<ul style="list-style-type: none"> • 18 µg Tiotropiumbromid einmal täglich: 10, keine Gründe berichtet • Placebo: 17, keine Gründe berichtet
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

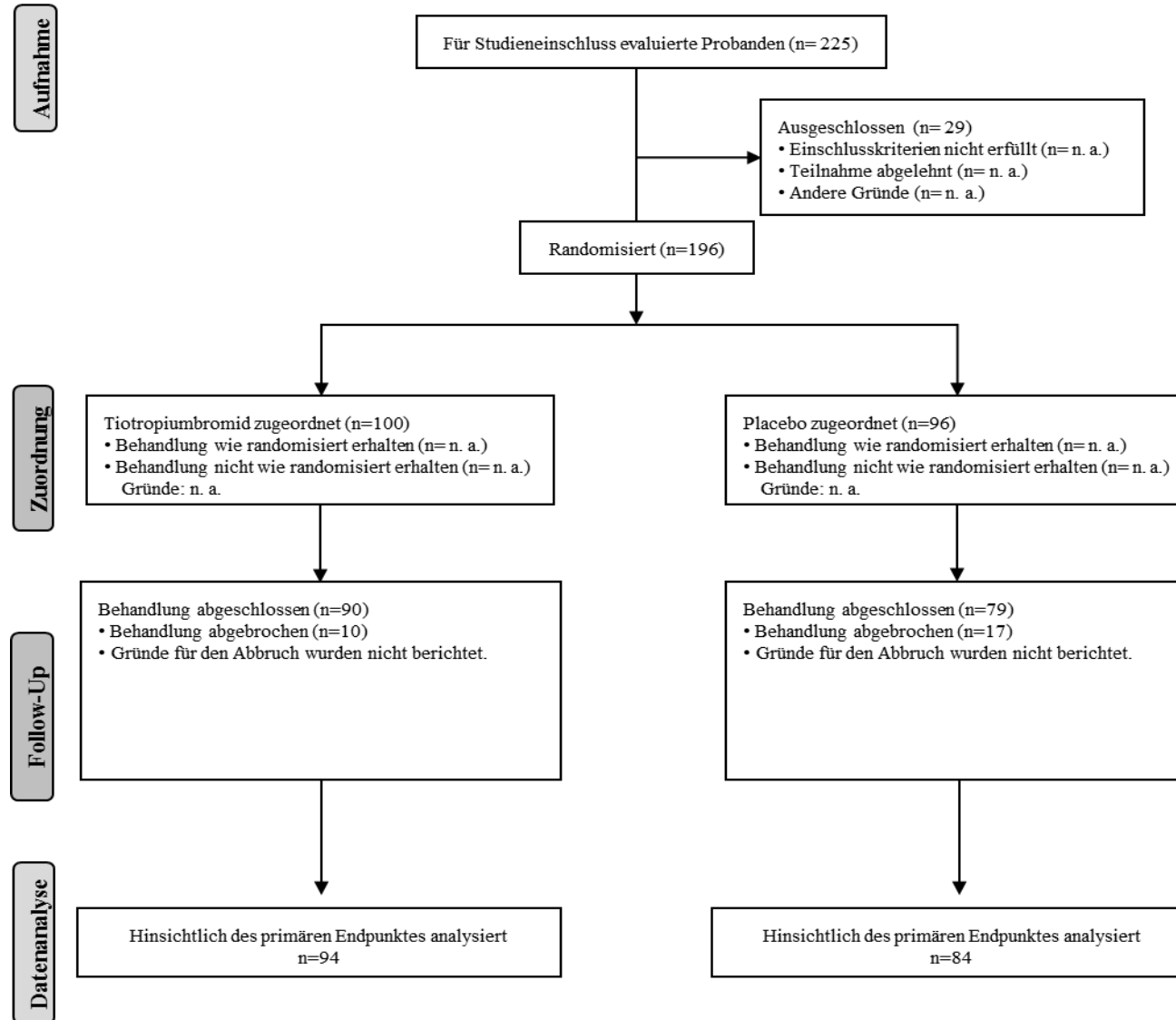


Tabelle 4.7-14 (Anhang): Studiendesign und -methodik für StudieDonohue 2002

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Untersuchung der Wirksamkeit und Sicherheit von Tiotropiumbromid und Salmeterol anhand multipler Ergebnismessungen, darunter Lungenfunktion, Dyspnoe und gesundheitsbezogene Lebensqualität (HRQoL) bei Patienten mit COPD.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, placebo- und aktiv-kontrollierte, doppelblinde klinische Studie, Parallelgruppen
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Eine relativ stabile Atemwegsobstruktion mit $FEV_1 \leq 60\%$ des Sollwerts und $FEV_1 \leq 70\%$ der FVC Mindestens 40 Jahre alt und (ehemalige) Raucher mit mindestens 10 Packungsjahren. Keine Vorgeschichte von Asthma, allergischer Rhinitis oder Atopie, keine erhöhte Eosinophilenzahl in der Vergangenheit oder kürzliche Infektion der oberen Atemwege Keine regelmäßige Gabe von Sauerstoff tagsüber von >1 Std./Tag
4b	Studienorganisation und Ort der Studiendurchführung	Diese Studie wurde in 39 Prüfzentren in 12 Ländern durchgeführt. Sponsor: Boehringer Ingelheim GmbH & Co. KG Einzelheiten zum publizierenden Autor: James F. Donohue, MD, FCCP, Division of Pulmonary Medicine, University of North Carolina, 420 Burnett-Womack Building CB-7020, Chapel Hill, NC E-Mail: jdonohue@med.unc.edu
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Die doppelblinde Verabreichung der Prüfmedikation erfolgte in einem Double-Dummy-Design mit Tiotropiumbromid 18 µg, oder Placebokapseln, die einmal täglich morgens inhaliert wurden, und Salmeterol 50 µg, oder Placebo mittels Metered-Dose-Inhaler (MDI) zweimal täglich (alle 12 Stunden). Die Tiotropiumbromid-Kapseln wurden mit einem Trockenpulverinhalator (HandiHaler; Boehringer Ingelheim; Berkshire, UK) verabreicht. Die Placebo-Gruppe erhielt die Tiotropiumbromid-Placebokapsel einmal

Item ^a	Charakteristikum	Studieninformation
		täglich und das Salmeterol-Placebo mittels MDI zweimal täglich
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Wirksamkeit: <ul style="list-style-type: none"> • 12-Stunden-Spirometrie, Transitional Dyspnoea Index (TDI) und St. George's Respiratory Questionnaire (SGRQ) Sicherheit: <ul style="list-style-type: none"> • Unerwünschte Ereignisse Informationen zur Validierung oder Optimierung der Ergebnismessung wurden nicht berichtet.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	n. a.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	n. a.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.

Item ^a	Charakteristikum	Studieninformation
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Doppel-Dummy-Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Verabreichung dieser Arzneimittel erfolgte in einem Double-Dummy-Design mit Tiotropiumbromid 18 µg, oder Placebokapseln, die einmal täglich morgens eingenommen wurden, und Salmeterol 50 µg, oder Placebo mittels Metered-Dose-Inhaler (MDI) zweimal täglich (alle 12 Stunden). Die Kapseln wurden mit einem Trockenpulverinhalator (HandiHaler; Boehringer Ingelheim; Berkshire, UK) verabreicht. Die Placebo-Gruppe erhielt die Placebokapsel einmal täglich und das Salmeterol-Placebo mittels MDI zweimal täglich.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Kovarianzanalysen wurden mit den Baseline-Werten als Kovariate durchgeführt. Die Baseline-FEV ₁ war definiert als FEV ₁ -Mittelwert der Messungen 60 Minuten und 10 Minuten vor Verabreichung der ersten Dosis der Studienmedikation. Eine ähnliche Berechnung wurde an allen anderen Studientagen durchgeführt. Allerdings werden die an diesen Tagen ermittelten Werte als Trough-Einsekundenkapazität ausgewiesen (d. h. AUC ₁ etwa 23 bis 24 Stunden nach der vorausgehenden Dosis Tiotropiumbromid oder 11 bis 12 Stunden nach der vorausgehenden Dosis Salmeterol). Um zu jedem Zeitpunkt dieselben Patienten in die Zusammenfassung der Spirometrie-Daten einbeziehen zu können, wurden die fehlenden Werte anhand anderer Werte des Patienten geschätzt, die an diesem Untersuchungstag aufgezeichnet wurden. Eine lineare Interpolation zwischen den beiden Nachbarmessungen wurde zur Schätzung der fehlenden Spirometriewerte in der Mitte des Profils verwendet. Für Werte, die aufgrund der Einnahme von Notfallmedikation am Ende des Profils fehlten, wurde der an dem Tag beobachtete FEV ₁ -Minimalwert als Schätzwert verwendet. Der letzte verfügbare Wert wurde als Schätzwert für die Daten

Item ^a	Charakteristikum	Studieninformation
		eingesetzt, die aus Gründen fehlten, die nicht mit dem Ansprechen des Patienten auf die Behandlung in Verbindung standen. Für die Analyse der mittleren wöchentlichen morgendlichen und abendlichen PEFR-Daten (maximaler expiratorischer Fluss) und für die bedarfsmäßige Salbutamol-anwendung wurde der Mittelwert der Aufzeichnungen während der letzten Woche des Baseline-Zeitraums als Kovariate verwendet. Anpassungen für Mehrfachvergleiche wurden nicht durchgeführt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Randomisiert: Tiotropiumbromid 18 µg einmal täglich (209), Salmeterol 50 µg BID (213), Placebo (201) • Erhielten die vorgesehene Behandlung: n. a. • Im Hinblick auf das Primärergebnis analysierte Patienten: n. a.
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Tiotropiumbromid 18 µg einmal täglich (25)</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse: 12; • Kein angegebener Grund: 13 <p>Salmeterol 50 µg BID (36)</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse: 29; • Kein angegebener Grund: 7 <p>Placebo (56)</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse: 39; • Kein angegebener Grund: 17

Item ^a	Charakteristikum	Studieninformation
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

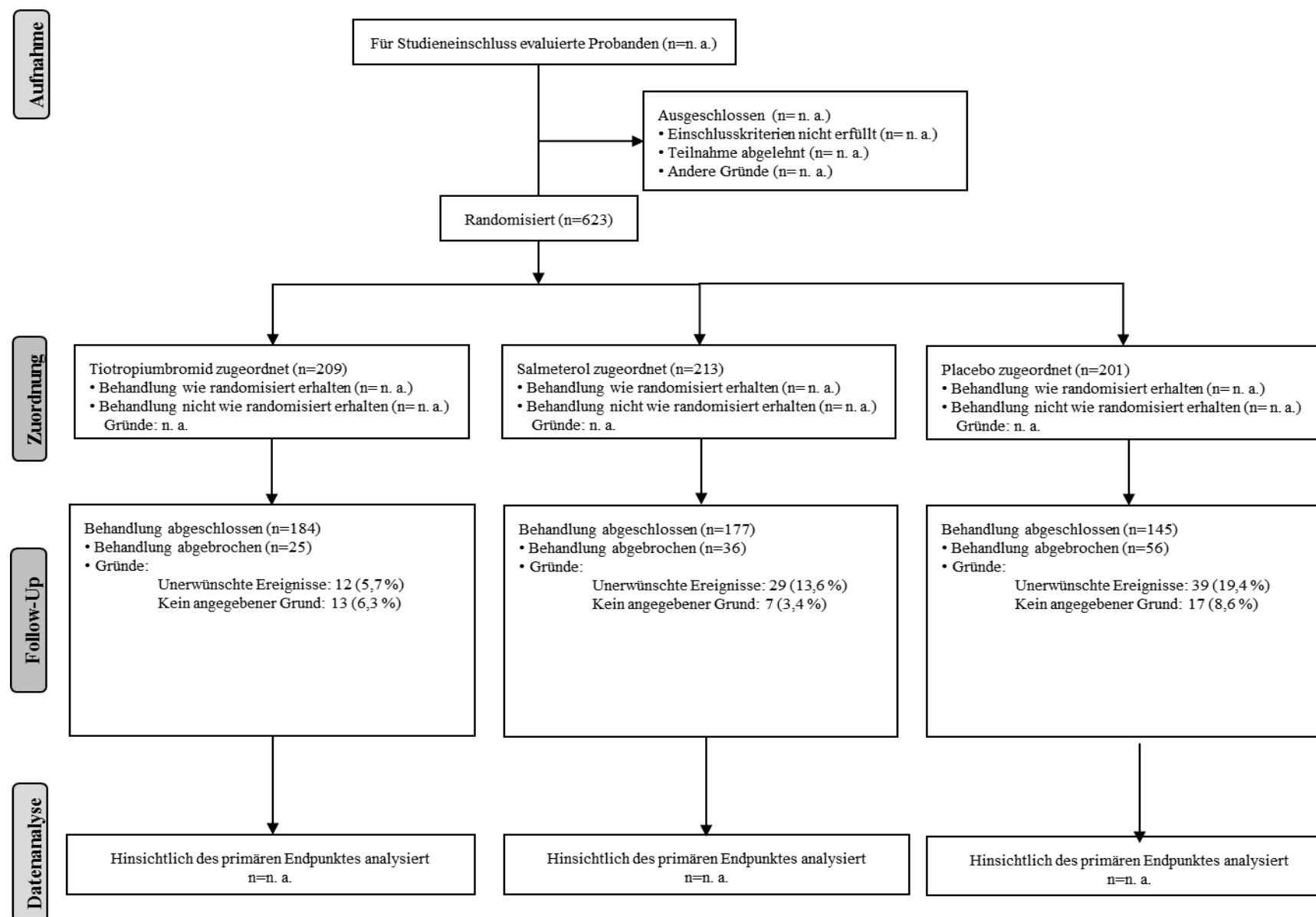


Tabelle 4.7-15 (Anhang): Studiendesign und -methodik für Studie INHANCE (Donohue 2010)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Nachweis der größeren Wirksamkeit von Indacaterol versus Placebo bei forciertem expiratorischem Volumen in 1 Sekunde (Einsekundenkapazität, FEV ₁) 24 Stunden nach Verabreichung der Dosis (Trough) nach 12 Wochen, Vergleich der Wirksamkeit von Placebo und Tiotropiumbromid und Beurteilung der Sicherheit und Verträglichkeit über 26 Wochen
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	RCT, Parallelgruppenstudie, Indacaterol vs. Placebo doppelblind. Die Tiotropiumbromidbehandlung erfolgte offen.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Patienten im Alter von ≥ 40 Jahren und (ehemalige) Raucher von ≥ 20 Packungsjahren und diagnostizierter mittelschwerer bis schwerer COPD Postbronchodilatorisches (innerhalb von 30 Minuten nach Inhalation von Albuterol 360 μ g) forciertes expiratorisches Volumen in 1 Sekunde (Einsekundenkapazität, FEV ₁) < 80 % und ≥ 30 % des Sollwerts und FEV ₁ /forcierte Vitalkapazität (FVC) < 70 % Kein Asthma in der Vorgeschichte
4b	Studienorganisation und Ort der Studiendurchführung	NR (in Publikation) www.clinicaltrials.gov nennt Prüfzentren in den USA, Kanada, Argentinien, Deutschland, Indien, Italien, Korea, Puerto Rico, Spanien, Schweden, Taiwan und der Türkei. ClinicalTrials.gov identifier NCT 00463567 Sponsor: Novartis Pharma AG Einzelheiten zum publizierenden Autor: Dr. Donohue University of North Carolina, 4125 BioInformatics Building, 130 Mason Farm Road, CB 7020,

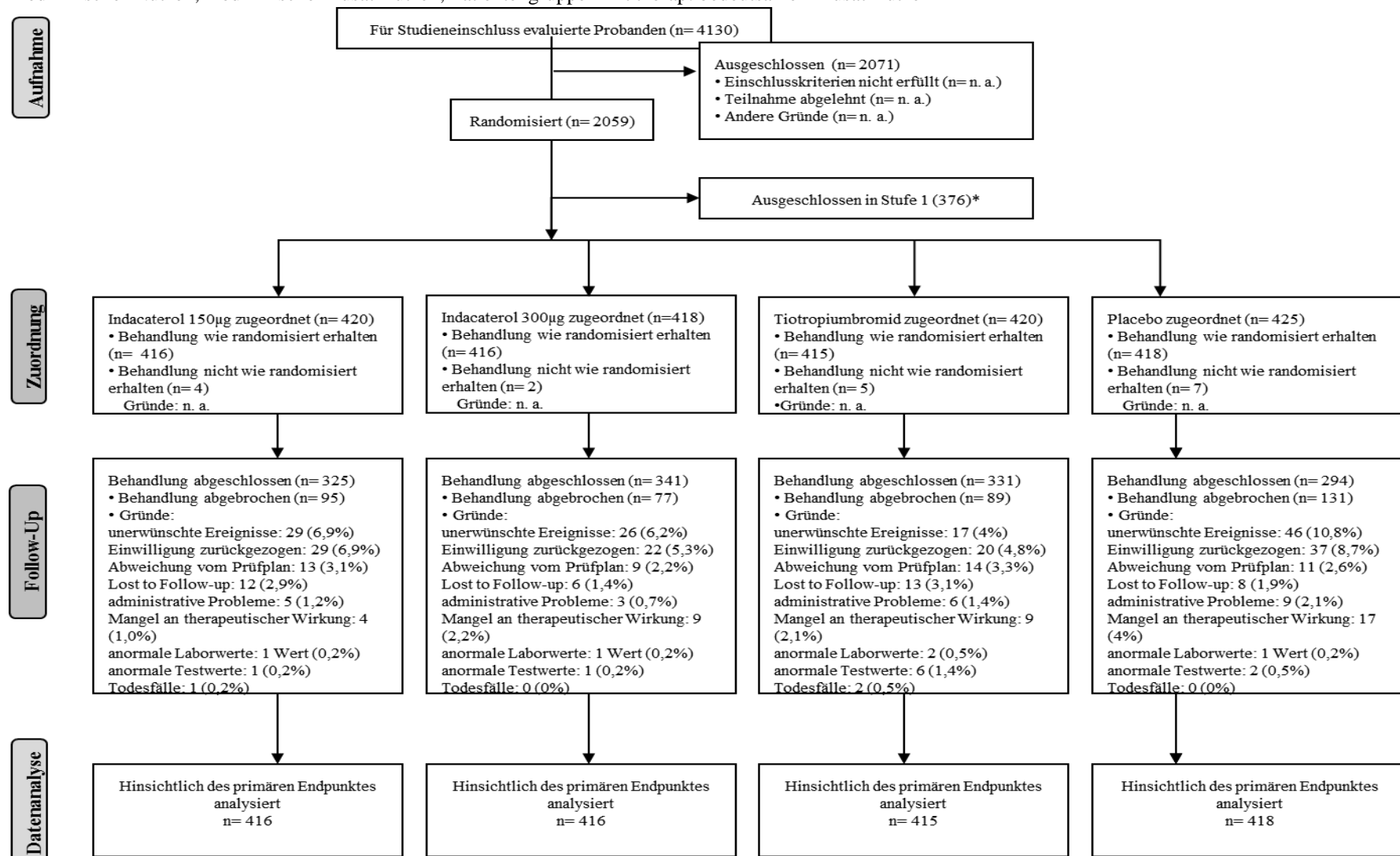
Item ^a	Charakteristikum	Studieninformation
		Chapel Hill, NC, 27599, USA, or james_donohue@med.unc.edu
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	In Stufe 1 wurden die Patienten zu Indacaterol 75, 150, 300 oder 600 µg einmal täglich, Formoterol 12 µg zweimal täglich oder Placebo, alle doppelblind, oder Tiotropiumbromid 18 µg, einmal täglich und offen, randomisiert verabreicht. In Stufe 2 waren die vier Behandlungsgruppen die Gruppen mit den beiden ausgewählten Dosen Indacaterol (150 und 300 µg), Tiotropiumbromid und dem Placebo.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primärer Wirksamkeitsendpunkt: Trough-FEV ₁ nach 12 Wochen Haupt-Sekundär-Endpunkt: Nachweis der Nicht-Unterlegenheit von wenigstens einer Indacaterol-Dosis im Vergleich zu Tiotropiumbromid in Bezug auf den Trough-FEV ₁ nach 12 Wochen. Zusätzliche Wirksamkeitsendpunkte: Transitional Dyspnea Index (TDI), Gesundheitsstatus (St. George's Respiratory Questionnaire, SGRQ) und Exazerbationen Informationen über die Validierung oder Optimierung der Ergebnismessung wurden nicht berichtet.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Beim Primärziel wurde eine Differenz von 120 ml bei der Trough-FEV ₁ für Indacaterol versus Placebo als klinisch bedeutsam präspezifiziert, bei einer angenommenen Standardabweichung von 270 ml. Beim Haupt-Sekundärziel wurde eine Differenz von 55 ml für Indacaterol versus Tiotropiumbromid angenommen, ausgehend von früheren Daten zu Tiotropium. Um mindestens 85 % Aussagekraft für das entscheidende Sekundärziel zu liefern, waren 340 evaluierbare Patienten (400 randomisierte Patienten) pro Indacaterol-Gruppe erforderlich. Diese Patientenzahl führt zu einer Aussagekraft von nahezu 100 % beim Primärziel (Indacaterol vs. Placebo)
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen	Zweistufiges Design: In Stufe 1 wurden die Patienten randomisiert und erhielten Indacaterol 75, 150, 300 oder 600 µg einmal täglich, Formoterol 12 µg zweimal täglich oder Placebo, alle doppelblind, oder Tiotropiumbromid 18 µg, einmal täglich

Item ^a	Charakteristikum	Studieninformation
	Studienabbruch	und offen, verabreicht. Ein unabhängiges Komitee verwendete vorab definierte Wirksamkeitskriterien, um zwei Dosen Indacaterol auf Grundlage von zweiwöchigen Wirksamkeits- und Sicherheitsdaten auszuwählen. Die beiden ausgewählten Dosen Indacaterol waren 150 und 300 µg. In Stufe 2 waren die vier Behandlungsgruppen die Gruppen mit den beiden ausgewählten Dosen Indacaterol, Tiotropiumbromid und Placebo.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Es erfolgte eine Stratifizierung nach Raucherstatus (Ex-Raucher, aktueller Raucher)
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Randomisierung erfolgte mit einem interaktiven Voice-Response-System
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Nach Abschluss von Stufe 1 hatte das für die Dosisauswahl zuständige unabhängige Komitee Zugriff auf die entblindeten Daten. Die einzigen Angaben, die dem Sponsor und den Prüfärzten mitgeteilt wurden, waren die beiden gewählten Indacaterol-Dosen. Die Verblindung der in der andauernden klinischen Studie tätigen Mitarbeiter wurde für die restliche Studiendauer nicht aufgehoben. Die Verblindung von Indacaterol und dem Placebo wurde bis zur Datenbanksperre der Studie am Ende von Stufe 2 aufrecht erhalten.

Item ^a	Charakteristikum	Studieninformation
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Doppelblinde Verabreichung von Indacaterol 150 oder 300 µg oder dem Placebo mittels Trockenpulverinhalator für Einzeldosen oder offene Verabreichung von Tiotropiumbromid 18 µg mittels HandiHaler® (Boehringer Ingelheim). (Verblindetes Tiotropiumbromid war nicht verfügbar.)
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die Wirksamkeit wurde für die ITT-Population evaluiert, die alle randomisierten Patienten umfasste, die mindestens eine Dosis der Studienmedikation erhalten hatten. Die Sicherheitspopulation umfasste alle Patienten, die mindestens eine Dosis des Studienmedikaments erhalten hatten.</p> <p>Primäres und Haupt-Sekundärziel wurden durch ein gemischtes Modell mit Behandlung, Raucherstatus und Land als festen Effekten und dem Baseline-FEV₁-Wert und der Reversibilität als Kovariaten analysiert. Fehlende Werte in Woche 12 wurden durch den vorausgehenden Wert (aus Woche 2 oder später) ersetzt.</p> <p>Die Zeit bis zur ersten COPD-Exazerbation wurde mittels Cox-Regressionsmodell analysiert und die Exazerbationsraten mittels Poisson-Regressionsmodell. Sofern nicht anderweitig angegeben, werden die Mittelwerte der kleinsten Quadrate mit Standardfehlern (SE) oder 95 %-Konfidenzintervallen (KI) angegeben.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Eine explorative Analyse der Primärvariable wurde in Subgruppen nach Baseline-Status von Alter (<65 oder ≥65 Jahre), Raucherstatus (Ex-Raucher oder aktueller Raucher) sowie Anwendung von ICS (inhalativen Glukokortikoiden) durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt	<p>Am Ende von Stufe 1:</p> <ul style="list-style-type: none"> • Randomisiert: Indacaterol 150 µg (420), Indacaterol 300 µg (418), Tiotropiumbromid 18 µg (420), Placebo (425) • Erhielten die vorgesehene Behandlung: Indacaterol 150 µg (416), Indacaterol 300 µg (416), Tiotropiumbromid 18 µg (415), Placebo (418) • Im Hinblick auf das Primärergebnis analysierte Patienten: Indacaterol 150 µg (416), Indacaterol 300 µg (416), Tiotropiumbromid 18 µg (415), Placebo (418)

Item ^a	Charakteristikum	Studieninformation
	wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>In Stufe 2:</p> <ul style="list-style-type: none"> • Indacaterol 150 µg (95) <p>unerwünschte Ereignisse: 29; Einwilligung zurückgezogen: 29; Abweichung vom Prüfplan: 13; Lost to Follow-up: 12; administrative Probleme: 5; Mangel an therapeutischer Wirkung: 4; anormale Laborwerte: 1, anormale Testwerte: 1; Todesfälle: 1</p> <ul style="list-style-type: none"> • Indacaterol 300 µg (77) <p>unerwünschte Ereignisse: 26; Einwilligung zurückgezogen: 22; Abweichung vom Prüfplan: 9; Lost to Follow-up: 6; administrative Probleme: 3; Mangel an therapeutischer Wirkung: 9; anormale Laborwerte: 1, anormale Testwerte: 1; Todesfälle: 0</p> <ul style="list-style-type: none"> • Tiotropiumbromid 18 µg (89) <p>unerwünschte Ereignisse: 17; Einwilligung zurückgezogen: 20; Abweichung vom Prüfplan: 14; Lost to Follow-up: 13; administrative Probleme: 6; Mangel an therapeutischer Wirkung: 9; anormale Laborwerte: 2; anormale Testwerte: 6; Todesfälle: 2</p> <ul style="list-style-type: none"> • Placebo (131) <p>unerwünschte Ereignisse: 46; Einwilligung zurückgezogen: 37; Abweichung vom Prüfplan: 11; Lost to Follow-up: 8; administrative Probleme: 9; Mangel an therapeutischer Wirkung: 17; anormale Laborwerte: 1 Wert; anormale Testwerte: 2; Todesfälle: 0</p>
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	<p>Patienten wurden ab 23. April 2007 rekrutiert, der letzte Patient komplettierte Die Studie am 23. August 2008.</p> <p>2008</p>
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen



* = Zweistufiges Design: In Stufe 1 wurden die Patienten randomisiert und erhielten Indacaterol 75, 150, 300 oder 600 µg einmal täglich, Formoterol 12 µg zweimal täglich oder das Placebo, alle doppelblind, oder Tiotropium 18 µg, einmal täglich und offen, verabreicht. Ein unabhängiges Komitee verwendete vorab definierte Wirksamkeitskriterien, um zwei Dosen Indacaterol auf Grundlage von zweiwöchigen Wirksamkeits- und Sicherheitsdaten auszuwählen. Die beiden ausgewählten Dosen Indacaterol waren 150 und 300 µg. In Stufe 2 waren die vier Behandlungsgruppen die Gruppen mit den beiden ausgewählten Dosen Indacaterol, Tiotropium und dem Placebo.

Tabelle 4.7-16 (Anhang): Studiendesign und -methodik für Studie SPRUCE (Freeman 2007)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Beurteilung der Auswirkung einer Aufnahme von Tiotropiumbromid in die übliche Behandlung bei COPD-Erkrankungen, die typisch sind für die normale klinische Praxis
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Parallelgruppen, doppelblind, Tiotropiumbromid 18 µg oder Placebo (1:1 Blockrandomisierung)
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Diagnose von COPD nach den Kriterien der British Thoracic Society und eine stabile Erkrankung (keine Exazerbation oder Atemwegsinfekte innerhalb der letzten 4 Wochen) mit forciertem expiratorischem Volumen in 1 Sekunde (Einsekundenkapazität, FEV ₁) zwischen 30 % und 65 % des Sollwerts und FEV ₁ /forcierte Vitalkapazität (FVC) ≤70 % präbronchodilatorisch Keine Vorgeschichte von Allergien oder Asthma Mindestens 40 Jahre alt, (ehemalige) Raucher mit mindestens 10 Packungsjahren und Gabe von kurz wirksamen β ₂ -Agonisten (SABA) als Bedarfsmedikation und keine Anticholinergika-Verordnungen im vorausgehenden Jahr
4b	Studienorganisation und Ort der Studiendurchführung	Vierundvierzig Primärversorgungszentren in England, Schottland und Wales Sponsor: Boehringer Ingelheim Pharmaceuticals Einzelheiten zum publizierenden Autor: Daryl Freeman - daryl@respiratoryresearch.org
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Tiotropiumbromid 18 µg einmal täglich und Placebo einmal täglich wurden in identisch aussehenden Inhalatoren auf Lactosebasis (HandiHaler®, Boehringer Ingelheim, Ingelheim am Rhein, Deutschland) abgegeben
6	Zielkriterien	

Item ^a	Charakteristikum	Studieninformation
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <ul style="list-style-type: none"> • Trough-FEV₁-Reaktion am Ende des 12-wöchigen Behandlungszeitraums <p>Sekundäre Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • Trough-FEV₁-Reaktion nach 2 und 6 Wochen und Trough-FVC-Reaktion nach 2, 6 und 12 Wochen <p>Zusätzliche Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • Dyspnoe gemessen durch Oxygen Cost Diagram <p>Sicherheit:</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse <p>Der Spirometer Micro Medical Lab 2000 wurde in allen Prüfzentren verwendet. Die Spirometer und ihre Anwendung, einschließlich Kalibrierung, mussten die Kriterien der American Thoracic Society erfüllen. Eine Spirometrie-Schulung wurde vor Beginn der Studie erteilt und eine Wiederholungsschulung erfolgte beim ersten Besuch. Stichproben der Kalibrierung der Spirometer wurden während der Monitoringbesuche durchgeführt und zufällig ausgewählte Spirographie-Geräte wurden von vier unabhängigen Prüfern des Lenkungsgremiums nach der Studie überprüft, um zu bestätigen, dass Kurven akzeptabler Qualität produziert wurden. Die Akzeptanz/Nicht-Akzeptanz der Lungenfunktionskurven wurde durch Konsens festgelegt.</p> <p>Der höchste FEV₁- und FVC-Wert aus drei technisch akzeptablen Messungen wurde aufgezeichnet.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	

Item ^a	Charakteristikum	Studieninformation
7a	Wie wurden die Fallzahlen bestimmt?	Bei einer angenommenen Standardabweichung von 235 ml wurde eine Gesamtzahl von 348 Patienten (174 pro Gruppe) als ausreichend festgelegt, um eine Differenz von 100 ml der Trough-FEV ₁ -Reaktion zwischen den Behandlungen mit einer Aussagekraft von mindestens 96 % und einem 5 %-igen Signifikanzniveau (zweiseitig) anhand eines zweiseitigen t-Tests für zwei Gruppen festzustellen
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	1:1 Blockrandomisierung
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die	Doppelblinde Studie

Item ^a	Charakteristikum	Studieninformation
	Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid 18 µg und Placebo wurden in identisch aussehenden Inhalatoren auf Lactosebasis (HandiHaler®, Boehringer Ingelheim, Ingelheim am Rhein, Deutschland) abgegeben.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die Analyse aller Endpunkte wurde mit Behandlungsdifferenzen geplant, die anhand einer Kovarianzanalyse bewertet wurden. Aufgrund einer signifikanten Schiefe mit ausgeprägter Heavy-Tailed-Verteilung beim primären Endpunkt (Trough-FEV₁-Reaktion) in beiden Behandlungsgruppen, wurde vom Lenkungsausschuss ein nichtparametrischer Ansatz für den primären Endpunkt als angemessen angesehen. Aus Gründen der Konsistenz wurde dieser Ansatz bei allen Wirksamkeitsendpunkten angewendet. Der Mann-Whitney-Test und Hodges-Lehman-Shift-Parameter für die Schätzung der Effektgrößen und ein 95 % Konfidenzintervall (KI) wurden daher zum Vergleich der Behandlungen herangezogen.</p> <p>Anzahl und Prozentsatz der Patienten mit ≥ 1 COPD-Exazerbation wurde mit dem Chi-Quadrat-Test verglichen.</p> <p>Um so viele Patienten wie möglich in die Auswertung einzuschließen, wurden die Wirksamkeitsanalysen für Lungenfunktionstests und Tagebucheintragen (DIARY) nach dem ITT-Prinzip mit dem Full-Analysis-Set (FAS) durchgeführt, d. h. randomisierte Patienten mit Baseline-Daten und Daten nach multiplen Dosen der zugewiesenen Behandlung. Fehlende Daten aufgrund einer sich verschlechternden COPD wurden durch den ungünstigsten Wert ersetzt. Daten, die aus anderen Gründen fehlten, wurden durch den letzten beobachteten Wert ersetzt</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten	<p>Randomisiert: Tiotropiumbromid 18 µg einmal täglich (200), Placebo (195)</p> <p>Erhielten die vorgesehene Behandlung: n. a.</p> <p>Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 18 µg einmal täglich (191), Placebo (183)</p>

Item ^a	Charakteristikum	Studieninformation
	haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Tiotropiumbromid 18 µg einmal täglich (18)</p> <p>Krankheitsverschlechterung während der Studie: 2; Verschlechterung einer anderen vorbestehenden Erkrankung: 1; sonstige unerwünschte Ereignisse: 8; fehlende Compliance: 3; Lost to Follow-up: 0; Einwilligung zurückgezogen: 3; sonstige Gründe: 1</p> <p>Placebo (33)</p> <p>Krankheitsverschlechterung während der Studie: 8; Verschlechterung einer anderen vorbestehenden Erkrankung: 0; sonstige unerwünschte Ereignisse: 14; fehlende Compliance: 2; Lost to Follow-up: 2; Einwilligung zurückgezogen: 5; sonstige Gründe: 2</p>
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Von Oktober 2002 bis Oktober 2003
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

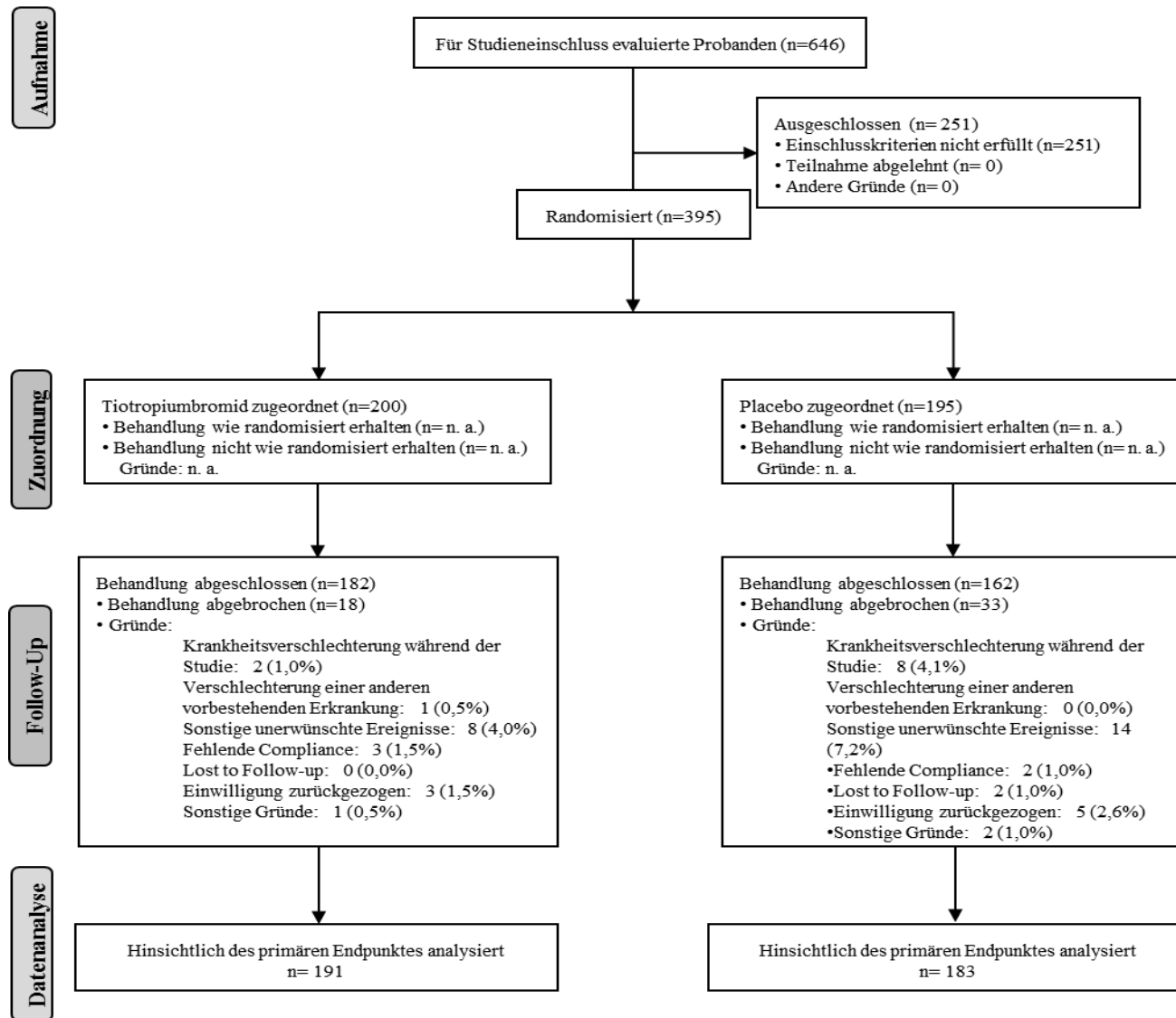


Tabelle 4.7-17 (Anhang): Studiendesign und -methodik für Studie SAFE Portugal (Moita 2008)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Untersuchung, ob die Wirkung von Tiotropiumbromid auf den Trough-FEV ₁ bei Patienten mit COPD durch den Raucherstatus beeinflusst wird
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte doppelblinde klinische Studie, Parallelgruppen
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Männer oder Frauen im Alter von ≥ 40 Jahren mit diagnostizierter COPD (FEV ₁ ≤ 70 % des Sollwerts und FEV ₁ /forcierte Vitalkapazität (FVC) ≤ 70 %) und (ehemalige) Raucher mit ≥ 10 Packungsjahren Keine Vorgeschichte von Asthma, allergischer Rhinitis, Atopie, Myokardinfarkt, instabiler Arrhythmie Keine ≥ 3 Exazerbationen der COPD im vorhergehenden Jahr oder eine Exazerbation oder Infektion der unteren Atemwege innerhalb der letzten sechs Wochen vor der Randomisierung
4b	Studienorganisation und Ort der Studiendurchführung	31 Prüfzentren in Portugal Sponsor: Boehringer Ingelheim Study Coordinator B.I. Portugal Lda. Einzelheiten zum publizierenden Autor: E-Mail-Adresse: joaquimmoita@chc.min-saude.pt (J. Moita). Tel.: +351 239 800 195; Fax: +351 239 800 149.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Tiotropiumbromid 18 mg einmal täglich, verabreicht mit dem HandiHaler-Inhalator (Boehringer Ingelheim, Ingelheim am Rhein, Deutschland) oder Placebo in Pulverkapseln zur oralen Inhalation mittels HandiHaler.

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <p>Trough-FEV₁ nach 12 Behandlungswochen</p> <p>Sekundäre Wirksamkeitsendpunkte:</p> <p>Trough-FEV₁ nach 6 Behandlungswochen, Trough-FVC nach 6 und 12 Behandlungswochen, Beurteilung der COPD-Symptome, Physician's Global Evaluation, Fragebogen zur Lebensqualität (EQ-5D) und Einsatz der Bedarfsmedikation tagsüber und nachts</p> <p>Sicherheit:</p> <p>Unerwünschte Ereignisse</p> <p>FEV₁ und FVC-Messungen wurden dreifach mit dem Dataspir 120C Spirometer (Sibelmed, Barcelona, Spanien) gemäß den Kriterien der American Thoracic Society durchgeführt. Es liegen keine weiteren Informationen über die Methoden vor, die zur Optimierung der Datenerfassung und zur Validierung der Erfassungsmethoden angewendet wurden.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Für die Feststellung einer Differenz von 130 ml mit 90 % Aussagekraft auf Grundlage einer Standardabweichung von 215 ml waren für jede Behandlungsgruppe 59 Patienten erforderlich. Um 59 ehemalige Raucher in der Placebo-Gruppe zu erhalten – ausgehend von einem angenommenen Verhältnis von Rauchern zu ehemaligen Rauchern von 3:2 – waren für jede Behandlungsgruppe 148 Patienten erforderlich
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	n. a.

Item ^a	Charakteristikum	Studieninformation
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid 18 mg einmal täglich, verabreicht mit dem HandiHaler-Inhalator (Boehringer Ingelheim, Ingelheim am Rhein, Deutschland) oder Placebo in Pulverkapseln zur oralen Inhalation mittels HandiHaler.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Für den primären Endpunkt wurde der Vergleich zwischen Tiotropiumbromid und dem Placebo mit einer Kovarianzanalyse (ANCOVA) durchgeführt, unter Einbeziehung von Behandlungsdauer und Zentrum als festen Effekten und dem Mittelwert des Trough-FEV ₁ bei Baseline als Kovariate. Der primäre Endpunkt wurde auch nach Raucherstatus stratifiziert (Raucher und ehemalige Raucher), der dem ANCOVA-Modell als fester Effekt hinzugefügt wurde. Der Mittelwert der kleinsten Quadrate (LSM) für die FEV ₁ -Reaktion wurde berechnet und mit jeder Behandlungsgruppe ein Vergleich durchgeführt, stratifiziert nach Raucherstatus. Die sekundären Spirometriemessungen wurden ebenfalls mit einem ähnlichen ANCOVA-Modell analysiert. Die COPD-Symptome

Item ^a	Charakteristikum	Studieninformation
		<p>und die Gesamtbeurteilung durch den Prüfarzt wurden nur deskriptiv analysiert. Der EQ-5D Fragebogen wurde nach den Konventionen der EuroQol Group Enterprise ausgewertet. Deskriptive statistische Methoden wurden für die Sicherheitsvariablen angewendet.</p> <p>Wirksamkeitsanalysen wurden unter Einbeziehung des Full-Analysis-Set (FAS) durchgeführt, welches alle Patienten umfasst, die mindestens eine Dosis des Medikaments erhielten, mindestens einen Baseline-Wert aufwiesen und ihren Raucherstatus nicht während der Studie geändert haben. Das Safety-Analysis-Set (SAS) umfasste alle Patienten, die mindestens eine Dosis des Medikaments erhalten und eine Sicherheitsbeurteilung nach der Behandlung hatten</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	<p>Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die</p> <p>a) randomisiert wurden,</p> <p>b) tatsächlich die geplante Behandlung/Intervention erhalten haben,</p> <p>c) in der Analyse des primären Zielkriteriums berücksichtigt wurden</p>	<p>Randomisiert: Tiotropiumbromid 18 µg einmal täglich (147), Placebo (164)</p> <p>Erhielten die vorgesehene Behandlung: Tiotropiumbromid 18 µg einmal täglich (147), Placebo (164)</p> <p>Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 18 µg einmal täglich (144), Placebo (160)</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Tiotropiumbromid 18 µg einmal täglich (11)</p> <p>unerwartete Verschlechterung der COPD: 1; unerwartete Verschlechterung einer anderen vorbestehenden Erkrankung: 1; Lost to Follow-up: 5; Einwilligung zurückgezogen: 0; sonstige unerwünschte Ereignisse: 3; sonstige Gründe: 1</p> <p>Placebo (11)</p> <p>unerwartete Verschlechterung der COPD: 2; unerwartete Verschlechterung einer anderen vorbestehenden Erkrankung: 0; Lost to Follow-up: 4; Einwilligung zurückgezogen: 1; sonstige unerwünschte Ereignisse: 2; sonstige Gründe: 2</p>

Item ^a	Charakteristikum	Studieninformation
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

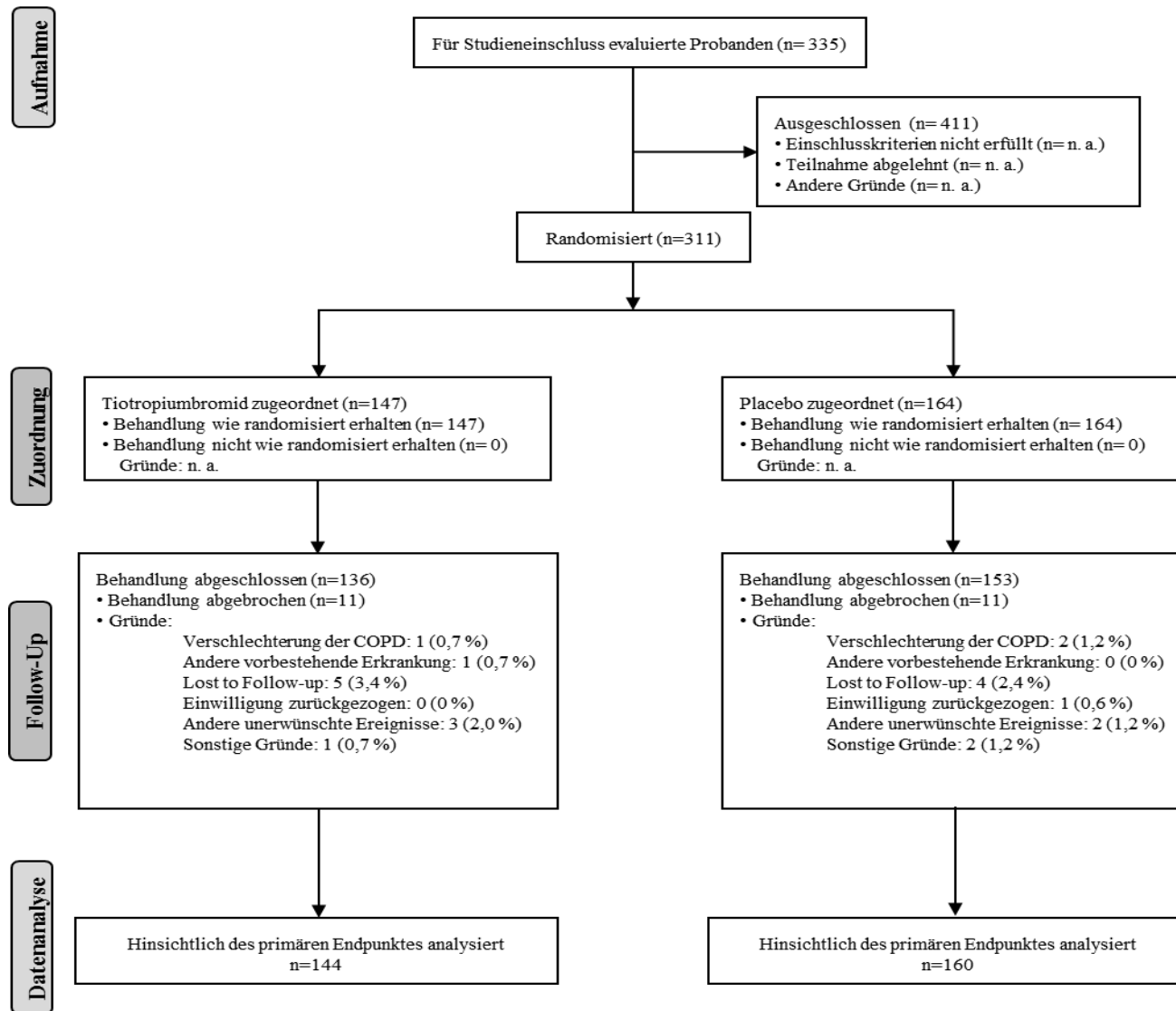


Tabelle 4.7-18 (Anhang): Studiendesign und -methodik für Studie Niewoehner 2005

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Untersuchung, ob Tiotropiumbromid COPD-Exazerbationen und Krankenhausaufenthalte aufgrund von Exazerbationen reduziert.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde RCT, Parallelgruppen, Randomisierungsverhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<p>Alter ab 40 Jahre, (ehemalige) Raucher mit mindestens 10 Packungsjahren, klinische Diagnose der COPD und FEV₁ von höchstens 60 % des Sollwerts und höchstens 70 % der FVC</p> <p>Keine klinische Diagnose von Asthma, Myokardinfarkt in den vorausgehenden 6 Monaten, schwere kardiale Arrhythmie oder Krankenhausaufenthalt aufgrund von Herzinsuffizienz im vorausgehenden Jahr.</p> <p>Bekannte mittelschwere bis schwere Niereninsuffizienz, moderate bis schwere symptomatische Prostatahypertrophie oder Blasenhalsostruktion, Engwinkelglaukom, aktuelle Strahlen- oder Chemotherapie aufgrund einer malignen Erkrankung oder Unfähigkeit zur Einwilligungserklärung</p> <p>Keine systemischen Corticosteroide in instabilen Dosen oder in regelmäßigen Tagesdosen > 20 mg Prednison (oder Äquivalent)</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>26 medizinische Zentren für Veteranen in den Vereinigten Staaten</p> <p>Name und Adresse des Studienorganisationsleiters liegen nicht vor. Einzelheiten zum publizierenden Autor: Dennis E. Niewoehner, MD, Pulmonary Section (111N), Veterans Affairs Medical Center, One Veterans Drive, Minneapolis, MN 55417; E-mail, niewo001@umn.edu.</p> <p>Die Studie wurde durch einen Grant von Boehringer Ingelheim und Pfizer unterstützt.</p>
5	Interventionen	Während der Studie haben die Patienten jeden Morgen eine Kapsel Tiotropiumbromid (18 µg) oder eine identisch aussehende

Item ^a	Charakteristikum	Studieninformation
	Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Placebokapsel inhaliert.
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <p>Prozentsatz von Patienten mit einer COPD-Exazerbation</p> <p>Prozentsatz von Patienten mit Krankenhausaufenthalt aufgrund einer COPD-Exazerbation</p> <p>Sekundärer Wirksamkeitsendpunkt:</p> <p>Zeit bis zur ersten COPD-Exazerbation</p> <p>Zeit bis zum ersten Krankenhausaufenthalt aufgrund einer COPD-Exazerbation</p> <p>Häufigkeit der Exazerbationen, exazerbationsbezogene Nutzung von Ressourcen des Gesundheitssystems, Häufigkeit von Krankenhausaufenthalten aufgrund aller Ursachen und Krankenhaustage</p> <p>Spirometrie:</p> <p>Die Prüfzentren führten die Spirometrie durch Anwendung eines allgemeinen prädiktiven Nomogramms mit Geräten und Methoden gemäß den Empfehlungen der American Thoracic Society durch. Es liegen keine weiteren Informationen über die Methoden vor, die zur Optimierung der Datenerfassung und zur Validierung der Erfassungsmethoden angewendet wurden.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Es wurde berechnet, dass ein Stichprobenumfang von 1.800 Patienten eine Aussagekraft von 80 % zur Ermittlung einer 18 %-igen relativen Reduktion des Prozentsatzes von Patienten mit mindestens einer Exazerbation hat.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	

Item ^a	Charakteristikum	Studieninformation
8a	Methode zur Generierung der zufälligen Zuteilung	Zentral erstellte Liste für Blockrandomisierung.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Geeignete Patienten wurden in gleicher Anzahl den Gruppen zugewiesen und erhielten Tiotropiumbromid oder das Placebo. Eine einzige Randomisierung wurde vorgenommen und die Blocks den Zentren zugeteilt.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Patienten wurden in den Studienzentren in die Studie aufgenommen.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie. Die Verblindung der Materialien wurde bei Boehringer Ingelheim vor der Auslieferung an die Prüfbzentren durchgeführt. Die Doppelverblindung wurde aufrechterhalten, bis alle Patienten die klinischen Untersuchungen vollständig abgeschlossen hatten oder die Entblindung aufgrund eines schwerwiegenden unerwünschten Ereignisses erforderlich war.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Patienten haben eine Kapsel Tiotropiumbromid (18 µg) oder eine identisch aussehende Placebokapsel inhaliert.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Daten wurden anhand eines modifizierten ITT-Prinzips analysiert: Alle verfügbaren Daten der Patienten mit einem beliebigen Follow-up-Kontakt, die mindestens 1 Kapsel des Studienmedikaments genommen hatten, wurden in die Analyse bis zum ersten Ereignis (für die Endpunkte der Zeit bis zum Ereignis) oder zum Ausscheiden aus der Studie eingeschlossen. Fehlten Daten, wurden Längsschnittanalysemethoden angewendet

Item ^a	Charakteristikum	Studieninformation
		<p>(Spirometrie), ausschließlich die beobachteten Daten (Anzahl von Ereignissen) analysiert oder Analysemethoden für zensierte Daten (Daten der Zeit bis zum Ereignis) angewendet.</p> <p>Die primären Co-Endpunkte wurden mit einem Cochran-Mantel-Haenszel-Test mit dem Zentrum als Stratum analysiert.</p> <p>Exazerbationen: Im ersten Schritt wurden die Prozentsätze der Patienten mit einer Exazerbation getestet. Bei Zurückweisung wurde der Prozentsatz von Patienten mit Krankenhausaufenthalt aufgrund einer COPD-Exazerbation analysiert. Die Testung erfolgte jeweils mit einem zweiseitigen Signifikanz-Niveau von 0,05. Wegen des präspezifizierten geschlossenen Testverfahrens war keine Anpassung bezüglich Multiplizität erforderlich.</p> <p>Die Daten der Zeit bis zum Ereignis wurden mit einem proportionalen Cox-Hazard-Modell, stratifiziert nach Zentrum, analysiert, wobei die Behandlung der einzige Term war. Kaplan-Meier-Kurven stellen die Überlebensfunktionen dar.</p> <p>Die Anzahl der Ereignisse (z. B. Exazerbationen, Exazerbationstage und Krankenhaustage) wurden mit einem stratifizierten Wilcoxon-Mann-Whitney-Test analysiert, bei dem beobachtete (keine abgeleiteten) Daten verwendet wurden. Bei den analysierten Daten handelte es sich um die Anzahl der Ereignisse eines Patienten geteilt durch die Exposition des Patienten stratifiziert nach Zentren.</p> <p>Die Spirometriedaten wurden analysiert mit einem linearen Mixed-Effektmodell mit Termen für Zentrum, Behandlung, Zeit und die Behandlung-Zeit-Interaktion als feste kategorische Effekte und dem Baseline-Wert als kontinuierliche Kovariate. Eine unstrukturierte Varianz-Kovarianz-Matrix wurde eingesetzt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Der Breslow-Day-Test wurde zur Beurteilung der zentrenübergreifenden Homogenität der Odds Ratios verwendet und zur Beurteilung der Konsistenz des subgruppenübergreifenden Prozentsatzes von Patienten mit einer Exazerbation oder einem Krankenhausaufenthalt (beispielsweise aktueller Raucher vs. ehemaliger Raucher).</p> <p>Eine zusätzliche Analyse wurde zur Schätzung der Behandlungsdifferenzen und der 95 %-igen KI, gewichtet nach Zentrum, verwendet.</p>
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung	Randomisiert: Tiotropiumbromid 18 µg einmal täglich (914),

Item ^a	Charakteristikum	Studieninformation
	gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Placebo (915) Erhielten die vorgesehene Behandlung: Tiotropiumbromid 18 µg einmal täglich (914), Placebo (915) Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 18 µg einmal täglich (914), Placebo (915)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Tiotropiumbromid 18 µg einmal täglich (75) Verschlechterung der COPD: 19; neue oder sich verschlechternde Komorbidität: 31; Nichteinhaltung des Prüfplans: 6; Lost to Follow-up: 4; Ablehnung der weiteren Studienteilnahme: 15 Placebo (111) Verschlechterung der COPD: 37; neue oder sich verschlechternde Komorbidität: 26; Nichteinhaltung des Prüfplans: 16; Lost to Follow-up: 7; Ablehnung der weiteren Studienteilnahme: 23; sonstige Gründe: 2
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Die Aufnahme in die Studie begann im September 2001, die Nachbeobachtung wurde im Februar 2003 abgeschlossen.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		
n. a. = nicht angegeben		

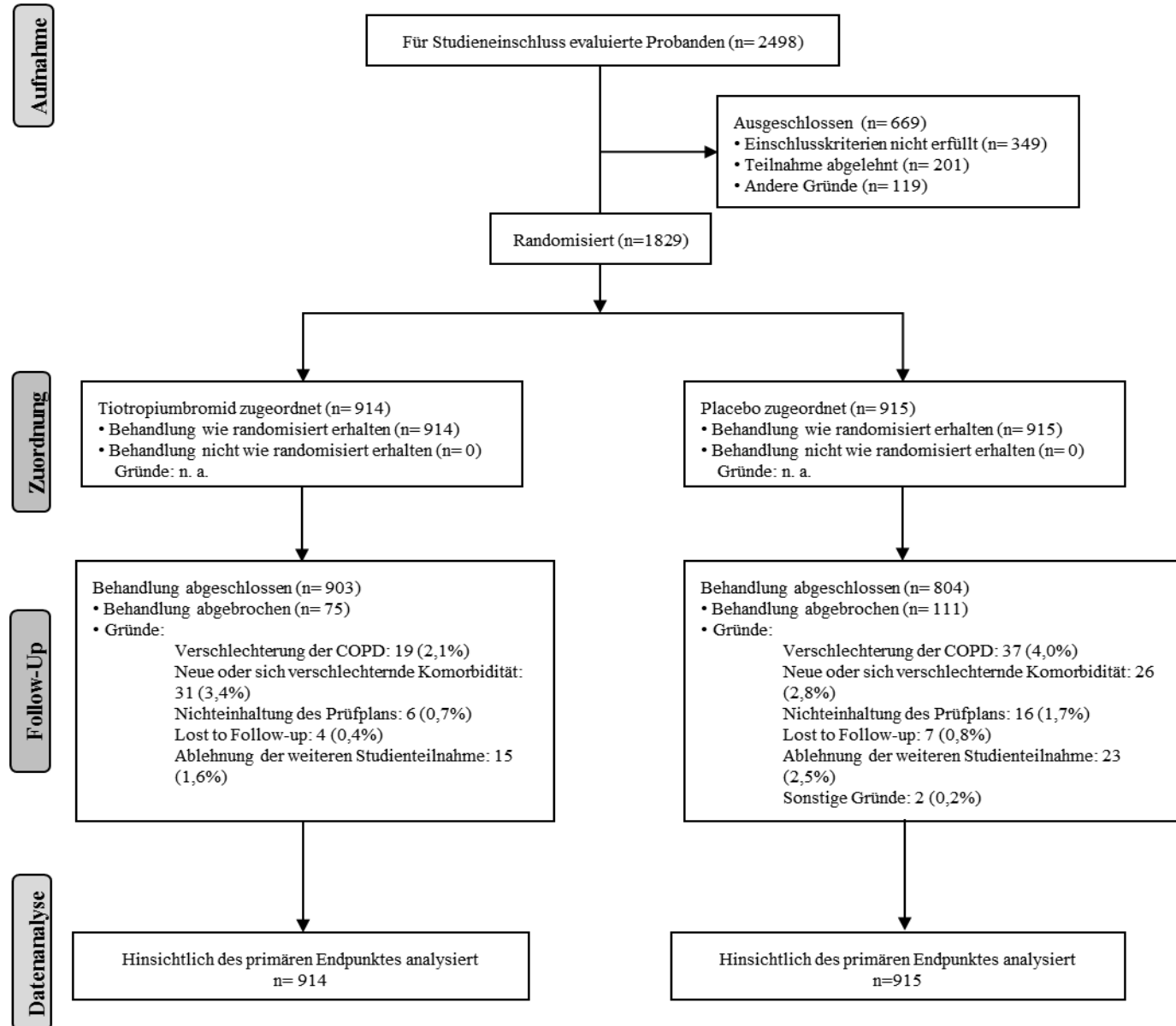


Tabelle 4.7-19 (Anhang): Studiendesign und -methodik für Studie UPLIFT (Tashkin 2008)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Untersuchung des Langzeiteffekts von Tiotropiumbromid (18 µg) als Inhalationskapsel mit dem HandiHaler® auf klinisch relevante Outcome-Parameter von COPD-Patienten (FEV ₁ , gesundheitsbezogene Lebensqualität, Exazerbationen, assoziierte Hospitalisationen und Mortalität).
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblinde, placebokontrollierte RCT, Parallelgruppen, Randomisierung im Verhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<p>Einschlusskriterien</p> <p>Ambulante Patienten beiderlei Geschlechts</p> <p>mindestens 40 Jahre alt</p> <p>(Ehemalige) Raucher mit ≥ 10 Packungsjahren</p> <p>Diagnose der COPD (postbronchodilatorische FEV₁ ≤70 % des Sollwerts und FEV₁/FVC ≤70 %)</p> <p>Wichtigste Ausschlusskriterien:</p> <p>Asthma-Anamnese</p> <p>Exazerbation oder Atemwegsinfektion in den letzten 4 Wochen vor der Studie</p> <p>Sauerstoff-Gabe für mehr als 12 Stunden pro Tag</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>490 Prüfzentren in 37 Ländern</p> <p>Sponsoren: Boehringer Ingelheim, Pfizer</p> <p>Deborah Burkhart Boehringer Ingelheim Pharmaceuticals, Inc. 900 Ridgebury Rd., P.O. Box 368 Ridgefield, Connecticut 06877-0368 Telephone: (203) 798-4363 Fax:(203) 791-6269 Einzelheiten zum publizierenden Autor: Dr. Tashkin von der David Geffen School of Medicine, University of California, Los Angeles, 10833 Le Conte Ave., Los Angeles, CA 90024</p>

Item ^a	Charakteristikum	Studieninformation
		Angeles, CA 90095-1690, oder unter dtashkin@mednet.ucla.edu.
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p>Prüfprodukt: Tiotropiumbromid Inhalationskapseln</p> <p>Dosis: 18 µg einmal täglich</p> <p>Art der Verabreichung: Orale Inhalation mittels HandiHaler</p> <p>Referenztherapie: Placebo-Inhalationskapseln</p> <p>Dosis: Eine Kapsel täglich</p> <p>Art der Verabreichung: Orale Inhalation mittels HandiHaler</p> <p>Die Patienten wurden angewiesen, die Studienmedikation immer zur ungefähr gleichen Zeit morgens anzuwenden (zwischen 7:00 und 10:00 Uhr).</p>
6	Zielkriterien	
6a	<p>Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten</p>	<p>Primäre Wirksamkeitsendpunkte:</p> <p>Jährliche Rückgangsrate des Trough-FEV₁ ab Tag 30 (stabiler Status) bis zum Ende der doppelblinden Behandlung</p> <p>Jährliche Rückgangsrate der FEV₁ 90 Minuten nach Verabreichung der Studienmedikation und von Ipratropium (einschließlich 30 Minuten nach Verabreichung von Salbutamol) ab Tag 30 (stabiler Status) bis zum Ende der doppelblinden Behandlung</p> <p>Sekundäre Wirksamkeitsendpunkte:</p> <p>Mittlere jährliche Rückgangsrate von FEV₁, der FVC und SVC 90 Minuten nach Verabreichung von Ipratropium, einschließlich 30 Minuten nach der Inhalation von Salbutamol, ab Tag 1 bis zum Ende der Studie (30 Tage nach der Behandlung mit der Studienmedikation).</p> <p>Mittlere jährliche Rückgangsrate von FVC und SVC 90 Minuten nach Verabreichung der Studienmedikation und von Ipratropium (einschließlich 30 Minuten nach Verabreichung von Salbutamol) ab Tag 30 bis zum Ende der doppelblinden Behandlung.</p> <p>Mittlere jährliche Rückgangsrate von FEV₁, der FVC und SVC vor Inhalation von Ipratropium und Salbutamol ab Tag 1 bis zum Ende der Studie (30 Tage nach der Behandlung mit der Studienmedikation)</p> <p>Mittlere jährliche Rückgangsrate der Trough-FVC und der SVC ab Tag 30 bis zum Ende der doppelblinden Behandlung</p> <p>Häufigkeit der COPD-Exazerbationen</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Zeit bis zur ersten Exazerbation.</p> <p>Anzahl der Patienten mit COPD-Exazerbationen.</p> <p>Anzahl der Exazerbationstage.</p> <p>Anzahl von Tagen zwischen den Exazerbationen.</p> <p>Anzahl der Patienten mit COPD-Exazerbationen, die zu Krankenhausaufenthalten führten.</p> <p>Zeit bis zur Exazerbation, die zu einem Krankenhausaufenthalt führte.</p> <p>Anzahl von Krankenhausaufenthalten aufgrund von Exazerbationen.</p> <p>Anzahl von Krankenhaustagen aufgrund von COPD-Exazerbationen.</p> <p>Rückgang des SGRQ-Gesamtscore</p> <p>Sicherheitsendpunkte:</p> <p>Alle unerwünschten Ereignisse</p> <p>Gesamtmortalität (beliebige Ursache)</p> <p>Todesfälle durch Atemwegserkrankungen</p> <p>Die Spirometrie wurde nach den Richtlinien der American Thoracic Society durchgeführt. Die Prüfbzentren erhielten identische Spirometriegeräte und studienspezifische Software. Eine zentrale Qualitätssicherungsprüfung aller Spirometriedaten wurde während der Studie durchgeführt. Weitere Informationen über die Validierung der Ergebnismessung wurden nicht berichtet. Die Literaturhinweise werden nach Ergebnis bereitgestellt und können Informationen über die Validität der Ergebnismessungen enthalten.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	<p>Während des Studienverlaufs erfolgten 3 umfangreiche Ergänzungen des Prüfplans. Amendment 1 ermöglichte die Erfassung der Lebensqualität mit dem EuroQol Five-Dimension Fragebogen.</p> <p>Amendments 2 und 3 etablierten Verfahren zur Erfassung der Informationen zum Vitalstatus für bis zu 4 Jahre nach der Randomisierung bei Patienten, die vorzeitig aus der Studie ausgeschieden sind, und setzten ein Entscheidungskomitee in Bezug auf die Mortalität ein, um die primäre Todesursache anhand verblindeter Daten zu beurteilen.</p>

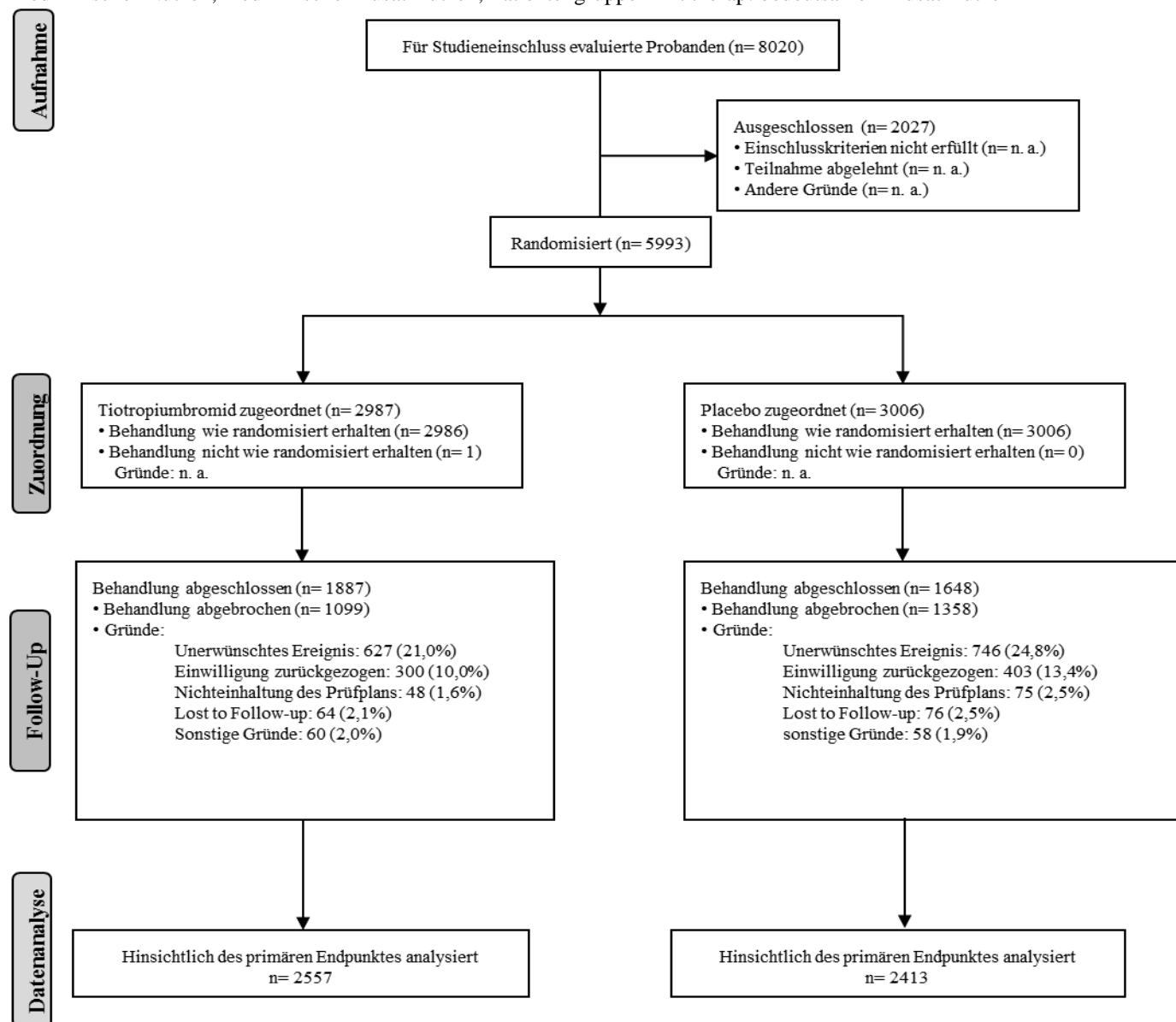
Item ^a	Charakteristikum	Studieninformation
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Auf Grundlage von Daten der früheren langfristigen Studien mit COPD-Patienten, wie den Studien zur Lungengesundheit, ISOLDE und BRONCHUS, sowie von Daten der einjährigen und halbjährigen Studien zu Tiotropiumbromid wurden als Standardabweichung der Rückgangsrate pro Jahr 90 ml angenommen. Daher wurde zur Ermittlung einer Differenz von 15 ml bei der Rückgangsrate zwischen Tiotropiumbromid und Placebo bei einem 5 %-igen Signifikanzniveau und einer Aussagekraft von 90 % ein Stichprobenumfang von 758 Patienten pro Gruppe benötigt. Für die Primäranalyse mussten die Patienten mindestens drei Beobachtungstermine ab Tag 30 absolvieren. Ausgehend von der Annahme, dass 35 % dieser Patienten die Studie vorzeitig ohne Vorliegen ausreichender Daten abbrechen würden (mindestens drei PFT-Messungen nach Tag 30, einschließlich Tag 30) und einer entsprechenden Anpassung diesbezüglich, wurde festgelegt, dass 1.166 Patienten pro Behandlungsgruppe benötigt werden. Auf Grundlage der Tiotropiumbromid-Studien 205.130 und 205.137, die in 18 Ländern in Europa, Nordamerika, Südafrika und Australien durchgeführt wurden, waren etwa 33 % der Patienten schwerwiegende Fälle nach ATS-Kriterien ($FEV_1 < 35\%$ des Sollwerts), etwa 40 % waren aktuelle Raucher und etwa 67 % wendeten inhalative Corticosteroide an. Um Subgruppenanalysen bei 40 % der Raucher mit ausreichender Aussagekraft durchführen zu können (1.166 Patienten pro Behandlungsgruppe), wurde der für die Studie benötigte Stichprobenumfang mit 2.916 Patienten pro Behandlungsgruppe oder 5.832 Patienten insgesamt bestimmt. Insgesamt wurde der Einschluss von 6.000 Patienten geplant.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Unabhängige Interims-Sicherheitsanalysen wurden durchgeführt und vom DSMB (Data and Safety Monitoring Board) etwa 12, 24, 36, 48 und 54 Monate nach Beginn der Studie überprüft</p> <p>Abbruchregeln:</p> <p>Interimsanalysen wurden etwa 24 und 36 Monate nach Beginn der Studie durchgeführt. Es war möglich, die Studie frühzeitig aufgrund eines Risiko-/Nutzen-Vorteils zugunsten von Tiotropiumbromid aus einem der beiden folgenden Gründe abbrechen, abgesehen von der Beobachtung signifikant weniger Todesfälle in der Tiotropium-Gruppe gegenüber der Placebo-Gruppe (nominaler p-Wert weniger als 0,05 zum Zeitpunkt der Interimsanalyse).</p> <p>1) Signifikant weniger Krankenhausaufenthalte aufgrund von COPD-Exazerbationen in der Tiotropium-Gruppe gegenüber Placebo (einseitige Wahrscheinlichkeit des Typ-I-Fehlers mit Kontrolle auf einem Niveau von 0,0005 am Ende der Studie unter Anwendung der O'Brien-Fleming-Spending-Funktion für die</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Interimsanalysen), die Untergrenze des Konfidenzintervalls für die Reduktion der Anzahl derartiger Krankenhausaufenthalte übersteigt 50 % oder</p> <p>2) Signifikant höhere Reduktion in der Tiotropiumbromid-Gruppe gegenüber Placebo bei der Rückgangsrates der Lungenfunktion bei beiden koprimary Endpunkten (einseitige Wahrscheinlichkeit des Typ-I-Fehlers mit auf einem Niveau von 0,0245 am Ende der Studie unter Anwendung der O'Brien-Fleming-Spending-Funktion für die Interimsanalysen und ebenso Anwendung des schrittweisen Verfahrens für die beiden koprimary Endpunkte, wie im Prüfplan angegeben); die Größe der mittleren Reduktion übersteigt 33 ml/Jahr bei der Interimsanalyse für beide koprimary Endpunkte im Vergleich zum Placebo</p>
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Interaktives Sprach-Response-System (IVRS)
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Geeignete Patienten wurden zufällig einer der beiden Behandlungen, Tiotropiumbromid oder Placebo, in Blöcken mit gleicher Verteilung der Behandlungen innerhalb eines Blocks zugeteilt. Die Randomisierungsliste wurde mit einem validierten System erstellt, das einen Pseudozufallszahlengenerator beinhaltet, so dass die resultierende Behandlungssequenz sowohl reproduzierbar als auch nicht vorhersagbar war
9	<p>Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment)</p> <p>Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war</p>	Jedes Prüfzentrum erhielt eine Telefonnummer (mit 24-stündiger Erreichbarkeit) und ein Passwort, das das Abrufen einer Reihe von Anweisungen ermöglichte, wie einem Patienten ein Medikamentenset zuzuweisen war. Medikamentensets wurden den Patienten alle sechs Monate zugeteilt. Ein Informationshandbuch mit Beschreibung der Schritte zur Randomisierung eines Patienten, zur Zuweisung eines Medikamentensets und zur Bestätigung des Empfangs der Studienmedikation wurde jedem Prüfzentrum im Prüfarzt-Studienfile am Standort zur Verfügung gestellt.
10	<p>Randomisierung, Durchführung</p> <p>Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?</p>	Die Randomisierung, Verpackung und Etikettierung der Studienmedikation wurde von BIPI vorgenommen. Ein interaktives Beantwortungssystem (IVRS) wurde für die Randomisierung der Patienten und zur Verwaltung des Arzneimittelbedarfs vom Prüfzentrum verwendet.
11	Verblindung	

Item ^a	Charakteristikum	Studieninformation
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	<p>Die Verabreichung von Tiotropiumbromid in Kapseln mit Inhalationspulver (18 µg) erfolgte mit dem HandiHaler®, der von Boehringer Ingelheim gestellt wurde. Die Verabreichung des Placebos in Kapseln mit Inhalationspulver erfolgte mit dem HandiHaler®, der von Boehringer Ingelheim gestellt wurde.</p> <p>Die Verblindung der Studienmedikamente erfolgte so, dass die Substanzen nicht zu erkennen waren.</p>
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Der Vergleich der primären Endpunkte zwischen den beiden Behandlungsgruppen erfolgte mit Zufalls-Koeffizient-Regressionsanalysen. Die Auswirkungen von Baseline-FEV₁, Alter, Geschlecht, Größe und Raucherstatus auf die jährliche Rückgangsrate der Einsekundenkapazität (FEV₁) wurden untersucht.</p> <p>In Sekundäranalysen wurden die folgenden sekundären Endpunkte verglichen:</p> <p>a) Die mittlere jährliche Rückgangsrate der Trough-FVC und der SVC von Tag 30 bis zum Abschluss der doppelblinden Behandlung wurde zwischen den beiden Behandlungsgruppen durch dieselbe Analyse wie bei den primären Endpunkten verglichen.</p> <p>b) Die mittlere jährliche Rückgangsrate des FEV₁, FVC und SVC vor und 90 Minuten nach Verabreichung von Ipratropium (30 Minuten nach Verabreichung von Salbutamol) von Tag 1 bis zum Ende der Studie (30 Tage nach Verabreichung der Studienmedikation) wurde berechnet und zwischen den beiden Behandlungsgruppen mittels Varianzanalyse verglichen.</p> <p>c) Die mittlere jährliche Rückgangsrate der FVC und der SVC 90 Minuten nach Verabreichung des Studienmedikaments und von Ipratropium (einschließlich 30 Minuten nach Verabreichung von Salbutamol) von Tag 30 bis Abschluss der doppelblinden Behandlung wurde zwischen den beiden Behandlungsgruppen durch dieselbe Analyse wie beim primären Endpunkt verglichen.</p> <p>d) Die Anzahl der Exazerbationen pro Patient wurde berechnet und zwischen den beiden Behandlungsgruppen mittels Wilcoxon-</p>

Item ^a	Charakteristikum	Studieninformation
		<p>(Mann-Whitney)-Rangsummentest verglichen.</p> <p>e) Die Zeit bis zur ersten Exazerbation wurde behandlungsgruppenübergreifend mittels Log-Rank-Test verglichen.</p> <p>f) Die Anzahl der Patienten mit COPD-Exazerbationen wurde mit dem Fisher-Exakt-Test verglichen.</p> <p>g) Bei der Anzahl der Exazerbationstage wurde zuerst die Anzahl pro Patient errechnet und dann die beiden Behandlungsgruppen mittels Wilcoxon-(Mann-Whitney)-Rangsummentest verglichen.</p> <p>h) Die Anzahl der Tage zwischen einzelnen Exazerbationen wurde errechnet und zwischen den beiden Behandlungsgruppen mittels Wilcoxon-Rangsummentest verglichen.</p> <p>i) Die Anzahl der Patienten mit COPD-Exazerbationen, die zu Krankenhausaufenthalten führten, wurde mit dem Fisher-Exact-Test verglichen.</p> <p>j) Die Zeit bis zur ersten Exazerbation, die zu einem Krankenhausaufenthalt führte, wurde zwischen den beiden Behandlungsgruppen mittels Log-Rank-Test verglichen.</p> <p>k) Bei Krankenhausaufenthalten aufgrund von Exazerbationen wurde die Anzahl pro Patient errechnet und danach die beiden Behandlungsgruppen mittels Wilcoxon-(Mann-Whitney)-Rangsummentest verglichen.</p> <p>l) Bei der Anzahl von Krankenhaustagen aufgrund von COPD-Exazerbationen wurde zuerst die entsprechende Anzahl pro Patient berechnet und dann die beiden Behandlungsgruppen mittels Wilcoxon-(Mann-Whitney)-Rangsummentest verglichen.</p> <p>m) Der Rückgang beim SGRQ-Gesamtscore von Tag 30 bis zum Behandlungsende wurde bei beiden Behandlungsgruppen mit derselben Analyse verglichen, wie sie bei den primären Endpunkten verwendet wurde.</p> <p>Alle randomisierten Patienten mit drei PFT-Messungen nach Tag 30 (einschließlich Tag 30) wurden in die Wirksamkeitsanalysen aufgenommen. Fehlende Daten wurden nicht abgeleitet.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Die Primäranalyse wurde für verschiedene Subgruppen, einschließlich Alter, Geschlecht, Raucherstatus, GOLD-Stadium, Region, Reversibilität, BMI und Begleitmedikation durchgeführt.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im	

Item ^a	Charakteristikum	Studieninformation
	Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Randomisiert: Tiotropiumbromid 18 µg einmal täglich (2.987), Placebo (3.006) Erhielten die vorgesehene Behandlung: Tiotropiumbromid 18 µg einmal täglich (2.986), Placebo (3.006) Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 18 µg einmal täglich (2.557), Placebo (2.413)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Tiotropiumbromid 18 µg einmal täglich (1.099) unerwünschtes Ereignis: 627; Einwilligung zurückgezogen: 300; Nichteinhaltung des Prüfplans: 48; Lost to Follow-up: 64; sonstige Gründe: 60 Placebo (1.358) unerwünschtes Ereignis: 746; Einwilligung zurückgezogen: 403; Nichteinhaltung des Prüfplans: 75; Lost to Follow-up: 76; sonstige Gründe: 58
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Die Patienten wurden von Januar 2003 bis einschließlich März 2004 für die Studie rekrutiert. Die Studie endete im Februar 2008.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		
n. a. = nicht angegeben		



Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4.7-20 (Anhang): Studiendesign und -methodik für Studie TIPHON (Tonnel 2008)

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Untersuchung der Wirkung einer neunmonatigen Behandlung mit Tiotropiumbromid 18 µg auf die HRQoL bei Patienten mit mittelschwerer bis schwerer COPD, unter Anwendung des SGRQ
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblind, Parallelgruppen, Randomisierungsverhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Männliche und weibliche ambulante Patienten im Alter von ≥ 40 Jahren mit klinischer COPD-Diagnose (prä- und postbronchodilatorische FEV ₁ 20 % – 70 % des Sollwerts und FEV ₁ /langsame Vitalkapazität [SVC] ≤ 70 %) entsprechend einer leichten, mittelschweren oder schweren COPD gemäß den Kriterien von 1995 der American Thoracic Society sowie eine Raucheranamnese von >10 Packungsjahren Kein Asthma, allergische Rhinitis oder Atopie in der Anamnese, regelmäßige Sauerstofftherapie tagsüber, Atemwegsinfektionen (in den letzten 6 Wochen), Myokardinfarkt innerhalb der letzten 6 Monate, behandlungspflichtige Herzrhythmusstörungen im letzten Jahr oder Krankenhausaufenthalt aufgrund von Herzversagen oder Lungenödem während der letzten 3 Jahre
4b	Studienorganisation und Ort der Studiendurchführung	123 Prüfzentren in Frankreich Die Studiennummer 205.256 wurde bei clinicaltrials.gov registriert. Sponsoren: Boehringer Ingelheim GmbH & Co. KG, Pfizer Einzelheiten zum publizierenden Autor: A.B. Tonnel, Centre Hospitalier Regional Universitaire (CHRU) de Lille, Service de Pneumologie et Immuno-Allergologie, Boulevard J. Leclercq, 59037 Lille Cedex, Frankreich. E-Mail: ab-tonnel@chru-lille.fr
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Tiotropiumbromid 18 µg einmal täglich oder Placebo verabreicht mit dem HandiHaler-Inhalator (Boehringer Ingelheim GmbH Deutschland)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <p>Anteil der Patienten, die bei Studienende (Monat 9) eine Reduktion von mindestens 4 Einheiten beim SGRQ-Score erreichen</p> <p>Zusätzliche Wirksamkeitsendpunkte:</p> <p>Beurteilungen wurden an Tag 1 und nach 3, 6 und 9 Behandlungsmonaten durchgeführt.</p> <p>SGRQ-Score</p> <p>VSRQ-Gesamtscore</p> <p>FEV₁, FVC, inspiratorische Kapazität, SVC und forciertes inspiratorisches Volumen in 1 Sekunde (Einsekundenkapazität, FEV₁) 30 Minuten vor der Verabreichung der Dosis und 24 Stunden nach der vorausgehenden Dosis</p> <p>COPD-Exazerbationen</p> <p>Alle spirometrischen Tests wurden dreifach durchgeführt und der höchste Messwert wurde in den nachfolgenden Analysen verwendet. Die Validierung der VSRQ liegt nicht vor, aber laut Artikel wird diese Information „an anderer Stelle publiziert“. Es liegen keine weiteren Informationen über die Methoden vor, die zur Optimierung der Datenerfassung und zur Validierung der Erfassungsmethoden angewendet wurden</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Ein Stichprobenumfang von 234 Patienten pro Behandlungsgruppe war erforderlich, um beim primären Endpunkt eine Aussagekraft von 80 % zur Ermittlung einer Differenz von 13 % zwischen den Gruppen zu liefern.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Patienten wurden im Verhältnis 1:1 randomisiert und erhielten entweder Tiotropiumbromid oder das Placebo. Die Patienten wurden unter Anwendung eines computererstellten Randomisierungsschemas ohne

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Item ^a	Charakteristikum	Studieninformation
		Stratifizierung zugewiesen (Blockgröße 4).
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid 18 µg einmal täglich oder das Placebo verabreicht mit dem HandiHaler-Inhalator (Boehringer Ingelheim GmbH Deutschland)
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Der primäre Endpunkt wurde mit dem Cochrane Mantel-Haenszel-Test, stratifiziert nach Zentrum, analysiert. Die Kovarianzanalyse mit Termen für die Behandlung wurde für alle HRQoL- und spirometrischen Endpunkte mit den Baseline-Daten als Kovariaten verwendet. Der exakte Test nach Fisher wurde zum Vergleich der Exazerbationshäufigkeit eingesetzt. Der Wilcoxon-Mann-Whitney-Test wurde zum Vergleich der Anzahl von Exazerbationstagen verwendet. Paarweise Behandlungsvergleiche der Zeit bis zur ersten COPD-Exazerbation anhand von Kaplan-Meier-Schätzungen wurden unter Anwendung des Log-Rank-Tests durchgeführt.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Es wurden vier Subgruppen definiert. Schwere der FEV ₁ (unter und über 50 % des Sollwerts) und Reversibilität bei Baseline, Anwendung begleitender inhalativer Corticosteroide (ja/nein) und Vorkommen von Exazerbationen (ja/nein) während der Studie. Die Subgruppenanalyse der SGRQ-Gesamtscore-Veränderungen gegenüber Baseline erfolgte mit einem Kovarianzanalyse-Modell mit den Faktoren Behandlung, Subgruppe und Subgruppen-Behandlung-Interaktion und mit dem Baseline-Wert als

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Item ^a	Charakteristikum	Studieninformation
		Kovariate.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Randomisiert: Tiotropiumbromid 18 µg einmal täglich (266), Placebo (288) 1 Patient erhielt die Behandlung nicht Analysiert in Bezug auf das Primäresultat: Tiotropiumbromid 18 µg einmal täglich (247), Placebo (245)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Tiotropiumbromid 18 µg einmal täglich (39) Unerwünschte Ereignisse: 13; Nichteinhaltung des Prüfplans: 11; Lost to Follow-up: 3; Einwilligung zurückgezogen: 5; sonstige Gründe: 7 Placebo (74) Unerwünschte Ereignisse: 31; Nichteinhaltung des Prüfplans: 10; Lost to Follow-up: 6; Einwilligung zurückgezogen: 20; sonstige Gründe: 7
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Die Patienten wurden von Mai 2002 bis Juni 2003 rekrutiert. Die Nachbeobachtung fand von August 2002 bis April 2004 statt.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		
n. a. = nicht angegeben		

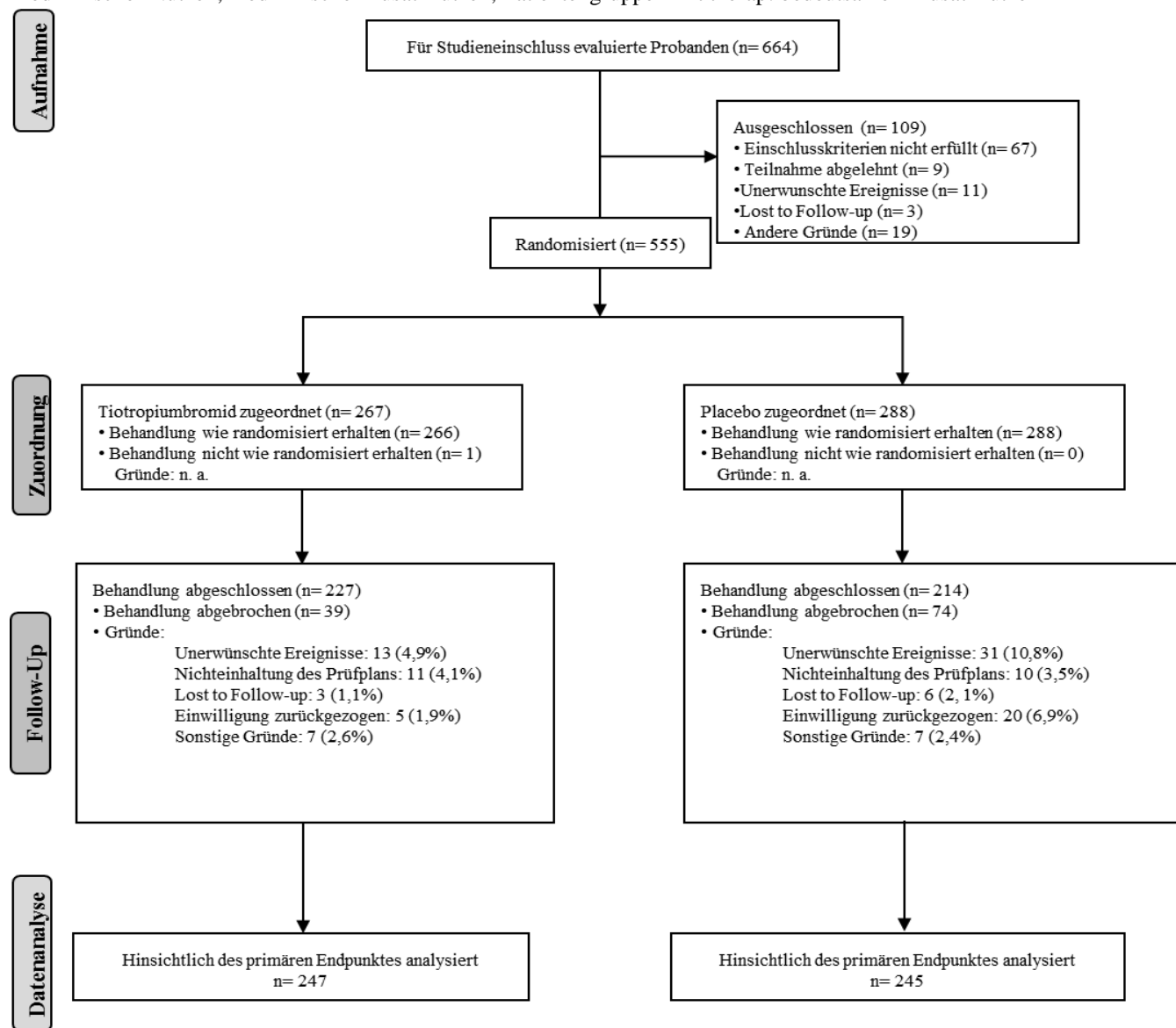


Tabelle 4.7-21 (Anhang): Studiendesign und -methodik für Studie Verkindre 2006

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Vergleich der Wirkung von Tiotropiumbromid mit Placebo auf die forcierte Vitalkapazität (FVC) bei Patienten mit mittelschwerer bis schwerer COPD und Hyperinflation der Lunge mit Ausdauer bei Belastung, Dyspnoe und gesundheitsbezogener Lebensqualität (HRQoL) als sekundäre Endpunkte
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblind, Parallelgruppen, Patienten wurden im Verhältnis 1:1 zu Tiotropiumbromid 18 µg einmal täglich oder Placebo randomisiert
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<p>Einschlusskriterien: Männliche und weibliche ambulante Patienten im Alter von ≥ 40 Jahren, Raucheranamnese von mindestens 10 Packungsjahren und mittelschwerer bis schwerer COPD ($FEV_1 \leq 50\%$ des Sollwerts und FEV_1/langsame Vitalkapazität, SVC, $\leq 70\%$), mit Hyperinflation der Lunge</p> <p><u>Ausschlusskriterien</u> Asthma, allergische Rhinitis oder Atopie in der Anamnese.</p> <p>Eosinophilenzahl ≥ 600 Zellen/mm³; kein Myokardinfarkt innerhalb des letzten Jahres, kongestives Herzversagen innerhalb der letzten 3 Jahre oder medikamentös behandlungspflichtige kardiale Arrhythmie, kürzliche Infektion der unteren Atemwege; regelmäßige Sauerstoffgabe; Anwendung oraler Corticosteroide in instabilen Dosen innerhalb von 6 Wochen vor Aufnahme in die Studie oder in einer stabilen Dosis > 10 mg Prednisolon Äquivalent täglich.</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>10 Prüfzentren in Frankreich</p> <p>Sponsoren: Boehringer Ingelheim, Pfizer Einzelheiten zum publizierenden Autor: Gerard Huchon, Pneumologie et Réanimation Hôtel-Dieu, 1 Place du Parvis Notre-Dame, FR-75181 Paris Cedex 4 (Frankreich) Tel.: +33 1 4234 8482, Fax: +33 1 4234 8448, E-Mail: gerard.huchon@htd.ap-hop-paris.fr</p>

Item ^a	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Tiotropiumbromid 18 µg einmal täglich oder Placebo verabreicht mit dem HandiHaler®-Inhalator (Boehringer Ingelheim)
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäre Wirksamkeit:</p> <p>Trough-FVC</p> <p>Zusätzliche Wirksamkeit:</p> <p>Spirometrie erfolgte beim Screening und an den Studientagen 1, 42 und 84</p> <p>FVC, IC, SVC und FEV 1. Alle spirometrischen Daten wurden 30 Minuten vor Verabreichung und 2 Stunden nach Verabreichung des Arzneimittels aufgezeichnet</p> <p>Täglicher maximaler expiratorischer Fluss (PEFR)</p> <p>Die Belastungskapazität wurde anhand des inkrementellen SWT (Shuttle-Walk-Test) beurteilt</p> <p>Die Belastungsdyspnoe wurde mit der modifizierten Borg-Skala, eine Dyspnoe bei Alltagsbelastung mit dem Transitional Dyspnoea Index (TDI) an Tag 42 und 84 beurteilt</p> <p>HRQoL anhand des St. George's Respiratory Questionnaire an Tag 1, 42 und 84</p> <p>Sicherheit:</p> <p>Unerwünschte Ereignisse</p> <p>Alle spirometrischen Tests wurden dreifach durchgeführt und der höchste Messwert wurde in den nachfolgenden Analysen verwendet. Um an den Tagen, an denen Spirometrietests durchgeführt wurden, standardisierte Bedingungen zu gewährleisten, wurde die Gabe von Salbutamol und von kurz- und langwirkendem Theophyllin mindestens 12, 24 bzw. 48 Stunden vor dem Test abgesetzt. Die Patienten trugen den höchsten der je drei täglich, morgens und abends gemessenen expiratorischen Spitzenflusses-Werte (PEFR) in ein Tagebuch ein</p> <p>Es liegen keine weiteren Informationen über die Methoden vor, die zur Optimierung der Datenerfassung und zur Validierung der Erfassungsmethoden angewendet wurden</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit	n. a.

Item ^a	Charakteristikum	Studieninformation
	Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Ein Stichprobenumfang von 98 Patienten (49 pro Gruppe) war erforderlich, um beim primären Endpunkt (FVC) eine Aussagekraft von 80 % zur Ermittlung der Differenz von 0,23 l zu liefern, ausgehend von einer Standardabweichung von 0,4 l bei einem 5 %-igen Konfidenzniveau
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	N. z.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung 1:1 zu Tiotropiumbromid 18 µg einmal täglich oder Placebo
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen	Doppelblinde Studie

Item ^a	Charakteristikum	Studieninformation
	beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid 18 µg einmal täglich oder Placebo verabreicht mit dem HandiHaler®-Inhalator (Boehringer Ingelheim)
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Eine ITT-Methode unter Verwendung des letzten beobachteten Wertes (LOCF-Methode) wurde für alle Analysen verwendet, außer bei Patienten, die aufgrund einer sich verschlechternden COPD aus der Studie ausschieden. In letzterem Fall wurde der ungünstigste Wert vor dem Ausscheiden verwendet.</p> <p>Die Differenz zwischen den beiden Behandlungsgruppen wurde mit einer Kovarianzanalyse (ANCOVA) unter Einbeziehung von Behandlungsdauer und Zentrum untersucht. Die Baseline-Daten dienten als Kovariate.</p> <p>Assoziationen zwischen Variablen wurden anhand der Pearson-Korrelation bewertet. Die statistische Signifikanz wurde bei $p < 0,05$ festgelegt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p>Randomisiert: Tiotropiumbromid 18 µg einmal täglich (46), Placebo (54)</p> <p>Erhielten die vorgesehene Behandlung: n. a.</p> <p>Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 18 µg einmal täglich (45), Placebo (46)</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von	<p>Tiotropiumbromid 18 µg einmal täglich</p> <p>Unerwünschte Ereignisse (1); Exazerbation (0)</p>

Item ^a	Charakteristikum	Studieninformation
	Gründen	Placebo Unerwünschte Ereignisse (6); Exazerbation (2)
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

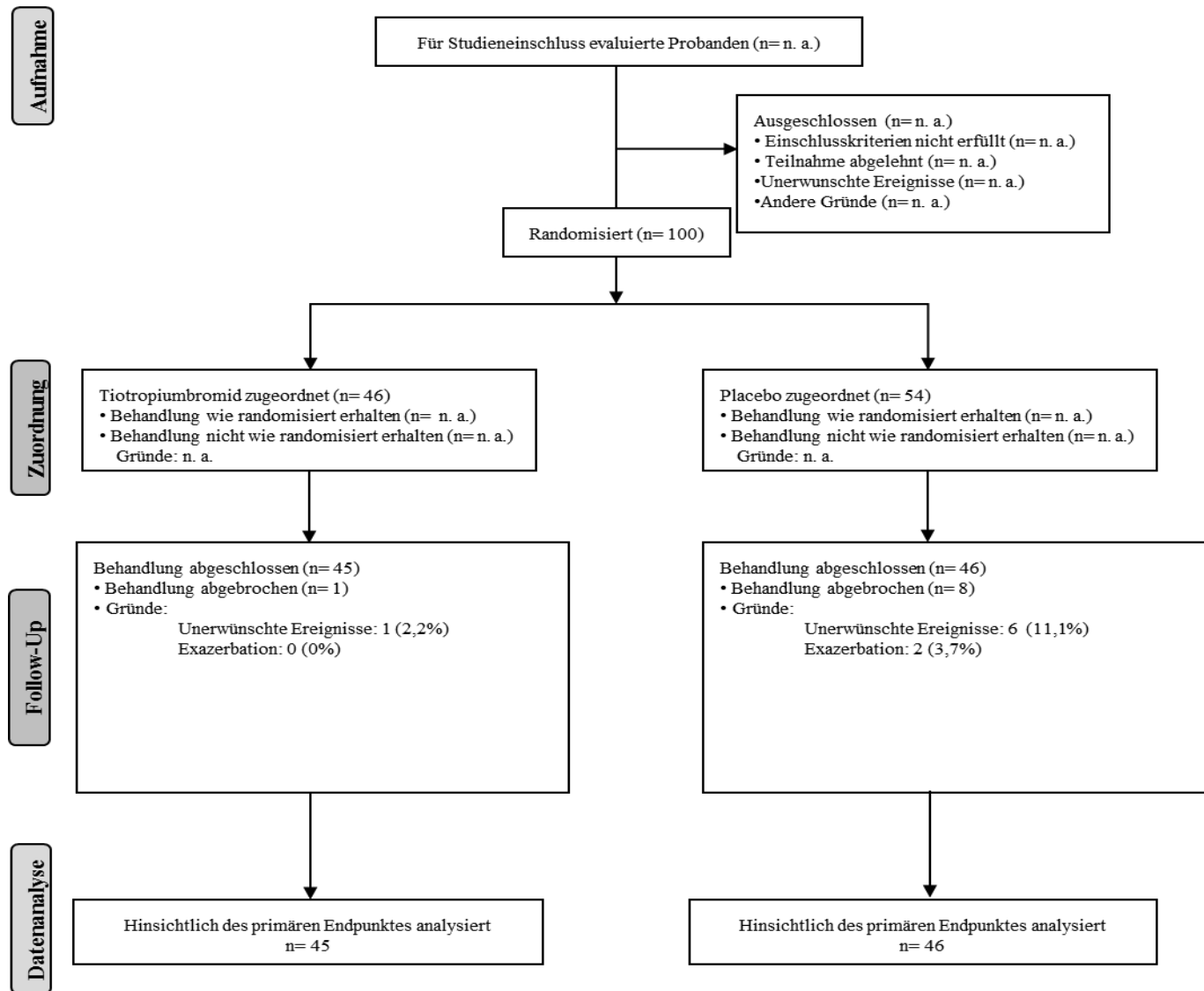


Tabelle 4.7-22 (Anhang): Studiendesign und -methodik für Studie Vogelmeier 2008

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Beurteilung der Wirksamkeit und Sicherheit von Formoterol, Tiotropium und der Kombination der beiden Substanzen bei COPD über 24 Wochen. Untersuchung, ob die Kombination von Formoterol zweimal täglich mit Tiotropium einmal täglich zu einem Nutzen führen würde bei einer Reihe von Wirksamkeitsvariablen über einen 24-wöchigen Zeitraum. Beurteilung der Sicherheitsprofile dieser Behandlungsansätze.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Partiell geblindet (Formoterol und Placebo), Parallelgruppen, Randomisierungsverhältnis 1:1:1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Männer und Frauen mit stabiler COPD im Alter von ≥ 40 Jahren bei COPD-Ausbruch und Raucheranamnese ≥ 10 Packungsjahren, forciertes expiratorisches Volumen in 1 Sekunde (Einsekundenkapazität, FEV ₁) < 70 % des Sollwerts für den Patienten (und $\geq 1,00$ L) und FEV ₁ /forcierte Vitalkapazität (FVC) < 70 % symptomatisch an mindestens 4 von 7 Tagen vor der Randomisierung (Symptomscore >0 auf der Tagebuchkarte). Keine Atemwegsinfektion und kein Krankenhausaufenthalt aufgrund einer akuten Exazerbation der COPD im Monat vor dem Screening.
4b	Studienorganisation und Ort der Studiendurchführung	86 Zentren: Arztpraxen und Spezialkliniken in Deutschland (30), Italien (19), den Niederlanden (9), der Russischen Föderation (9), Polen (7), der Tschechischen Republik (4), Spanien (4) und Ungarn (4) Sponsor: Novartis Pharma AG Einzelheiten zum publizierenden Autor: Claus Vogelmeier, claus.vogelmeier@med.uni-marburg.de Tel.: +49 6421 286 6451;

Item ^a	Charakteristikum	Studieninformation
		Fax: +49 6421 286 8987
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	(1) Formoterol 10 µg zweimal täglich (BID) mit dem Mehrfachdosen-Trockenpulverinhalator (MDDPI); (2) Placebo BID mit dem MDDPI; (3) Tiotropiumbromid 18 µg einmal täglich mit dem HandiHaler + Formoterol 10 mg BID mit dem MDDPI und (4) Tiotropiumbromid 18 mg einmal täglich (morgens) mit dem HandiHaler + Placebo BID mit dem MDDPI
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt:</p> <p>FEV₁ gemessen 2 Stunden nach Verabreichung der Dosis nach 24 Behandlungswochen</p> <p>Sekundäre Wirksamkeitsendpunkte:</p> <p>FEV₁ und FVC zu anderen Zeitpunkten während der Studie (5 Minuten, 2 und 3 Stunden nach Verabreichung der ersten Dosis des Medikaments und nach 12 und 24 Behandlungswochen); COPD-Exazerbationen; Symptom-Scores, Einsatz von Bedarfsmedikation, expiratorischer Spitzenfluss (PEF); Lebensqualität und 6-Minuten-Gehtest. Der St. George's Respiratory Questionnaire (SGRQ) wurde bei Baseline und nach 12 und nach 24 Behandlungswochen erhoben. Der 6-Minuten-Gehtest wurde bei Baseline und nach 24 Behandlungswochen bewertet.</p> <p>In Bezug auf die FEV₁ und die FVC wurde der höchste der jeweils drei akzeptablen Werte für jeden Zeitpunkt aufgezeichnet. Sofern möglich, wurde ein Patient im gesamten Studienverlauf immer von derselben Person beurteilt. Die Spirometer wurden stets vor den Messungen bei jedem Besuch kalibriert und die Kalibrierungsberichte wurden gespeichert. Der 6-Minuten-Gehtest wurde nach den Angaben der ATS durchgeführt: Richtlinien für den 6-Minuten-Gehtest. (Am J Respir Crit Care Med 2002;166:111e7.) Es liegen keine weiteren Informationen über die Methoden vor, die zur Optimierung der Datenerfassung und zur Validierung der Erfassungsmethoden angewendet wurden</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Zwei frühere Formoterol-Studien bei Patienten mit COPD wurden herangezogen, um die FEV ₁ -Standardabweichung auf 0,30 l bei Untersuchungen 2 Stunden nach Verabreichung des Medikaments zu schätzen. Um eine Aussagekraft von 80 % zur Ermittlung der klinisch relevanten Differenz von 0,12 l zwischen Formoterol und Placebo statistisch signifikant auf 5-%igen

Item ^a	Charakteristikum	Studieninformation
		Niveau bei zweiseitiger Testung zu erhalten, waren etwa 100 Patienten pro Gruppe erforderlich. Eine separate Zielsetzung war die Erfassung von Sicherheitsdaten für Patienten im Alter ab 65 Jahren, die (davon wurde ausgegangen) etwa die Hälfte der Patientenpopulation ausmachen würden. Auf dieser Basis war ein Stichprobenumfang von etwa 200 Patienten pro Gruppe gewählt, um eine ausreichende Aussagekraft für den Test des primären Ziels und die Analysen der sekundären Ziele bei gleichzeitiger zufriedenstellender Generierung von Sicherheitsdaten für die Anwendung bei älteren Patienten zu haben.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgte ohne Stratifizierung.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Anzahl der je Prüfzentrum randomisierten Patienten lag durchschnittlich bei 10 (Bereich von 1 bis 32). Keine weiteren Informationen über die Durchführung.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c)	Die Studie war teilweise doppelblind (Formoterol, Placebo)

Item ^a	Charakteristikum	Studieninformation
	diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	<p>Ein Placebo für Tiotropiumbromid war nicht verfügbar. Der Vergleich zwischen Formoterol und Tiotropiumbromid erfolgte daher unverblindet.</p> <p>Die Patienten wussten, ob sie Tiotropiumbromid nahmen, wussten aber nicht, ob sie Formoterol oder Placebo mit dem MDDPI anwendeten. Nur der Vergleich Formoterol versus Placebo war doppelblind.</p>
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>FEV₁ und FVC wurden mit der Kovarianzanalyse (ANCOVA) analysiert, mit Land, Geschlecht, Reversibilität (<15 % oder ≥15 %), Raucherstatus bei Baseline (Raucher oder ehemaliger Raucher) und Behandlung als Faktoren und den FEV₁-Wert bei Baseline als Kovariate.</p> <p>Der PEF-(expiratorischer Spitzenfluss)-Durchschnittswert im Verlauf der ersten 6 Behandlungswochen wurde mit einem ähnlichen ANCOVA-Modell analysiert, wobei der mittlere PEF der letzten 7 Tage der Screening-Phase als Baseline-Wert diente.</p> <p>Die Prozentsätze der belastenden COPD-Tage und der Exazerbationstage wurden mit einem ähnlichen ANCOVA-Modell analysiert, wobei als Baseline-Wert der Durchschnitt der 7 Tage vor der Randomisierung herangezogen wurde. Die Anzahl der Patienten mit COPD-Exazerbationen, die eine zusätzliche Therapie benötigten und der Patienten, bei denen Krankenhausaufenthalte erforderlich waren, wurden mit dem Cochran-Mantel-Haenszel-Test, stratifiziert nach Land, analysiert. Der mittlere tägliche Gesamtsymptomscore (Summe der Scores für die fünf Symptome) und der mittlere Einsatz von Bedarfsmedikation wurden mit dem Van-Elteren-Test, stratifiziert nach Baseline-Score, analysiert.</p> <p>Bei der Lebensqualität wurden die Scores für drei Domänen (Symptome, Aktivität und Auswirkungen) und der Gesamtscore errechnet und mit einem ANCOVA-Modell analysiert, das dem bei der primären Variablen verwendeten ähnlich war.</p> <p>Der 6-Minuten-Gehtest wurde mit einem ähnlichen Modell analysiert.</p> <p>Die ITT-Population umfasste alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten hatten. Diese Population wurde für die Wirksamkeits- und Sicherheitsanalysen verwendet. Alle Analysen erfolgten mit einem 5%-igen Signifikanzniveau (zweiseitig). Bei den</p>

Item ^a	Charakteristikum	Studieninformation
		Wahrscheinlichkeitswerten erfolgte keine Anpassung in Bezug auf Mehrfachvergleiche.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Randomisiert: Formoterol (210), Tiotropiumbromid (221), Tiotropiumbromid + Formoterol (207), Placebo (209) Erhielten die vorgesehene Behandlung: Formoterol (210), Tiotropiumbromid (221), Tiotropiumbromid + Formoterol (207), Placebo (209) Analysiert in Bezug auf das Primärergebnis: Formoterol (210), Tiotropiumbromid (221), Tiotropiumbromid + Formoterol (207), Placebo (209)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Formoterol (25) Einwilligung vom Probanden zurückgezogen: 12; unerwünschte Ereignisse: 6, Nichteinhaltung des Prüfplans: 4; Lost to Follow-up: 2; administrative Probleme: 1; unbefriedigende therapeutische Wirkung: 0; Tod: 0 Tiotropiumbromid (29) Einwilligung vom Probanden zurückgezogen: 11; unerwünschte Ereignisse: 13, Nichteinhaltung des Prüfplans: 3; Lost to Follow-up: 0; administrative Probleme: 0; unbefriedigende therapeutische Wirkung: 2; Tod: 0 Tiotropiumbromid + Formoterol (25) Einwilligung vom Probanden zurückgezogen: 11; unerwünschte Ereignisse: 8, Nichteinhaltung des Prüfplans: 3; Lost to Follow-up: 0; administrative Probleme: 3; unbefriedigende therapeutische Wirkung: 0; Tod: 0 Placebo (30) Einwilligung vom Probanden zurückgezogen: 11; unerwünschte Ereignisse: 8, Nichteinhaltung des Prüfplans: 6; Lost to Follow-up: 0; administrative Probleme: 1; unbefriedigende therapeutische

Item ^a	Charakteristikum	Studieninformation
		Wirkung: 3; Tod: 1
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Der erste Patient wurde im Oktober 2004 rekrutiert und alle Patienten hatten die Studie bis November 2005 abgeschlossen.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

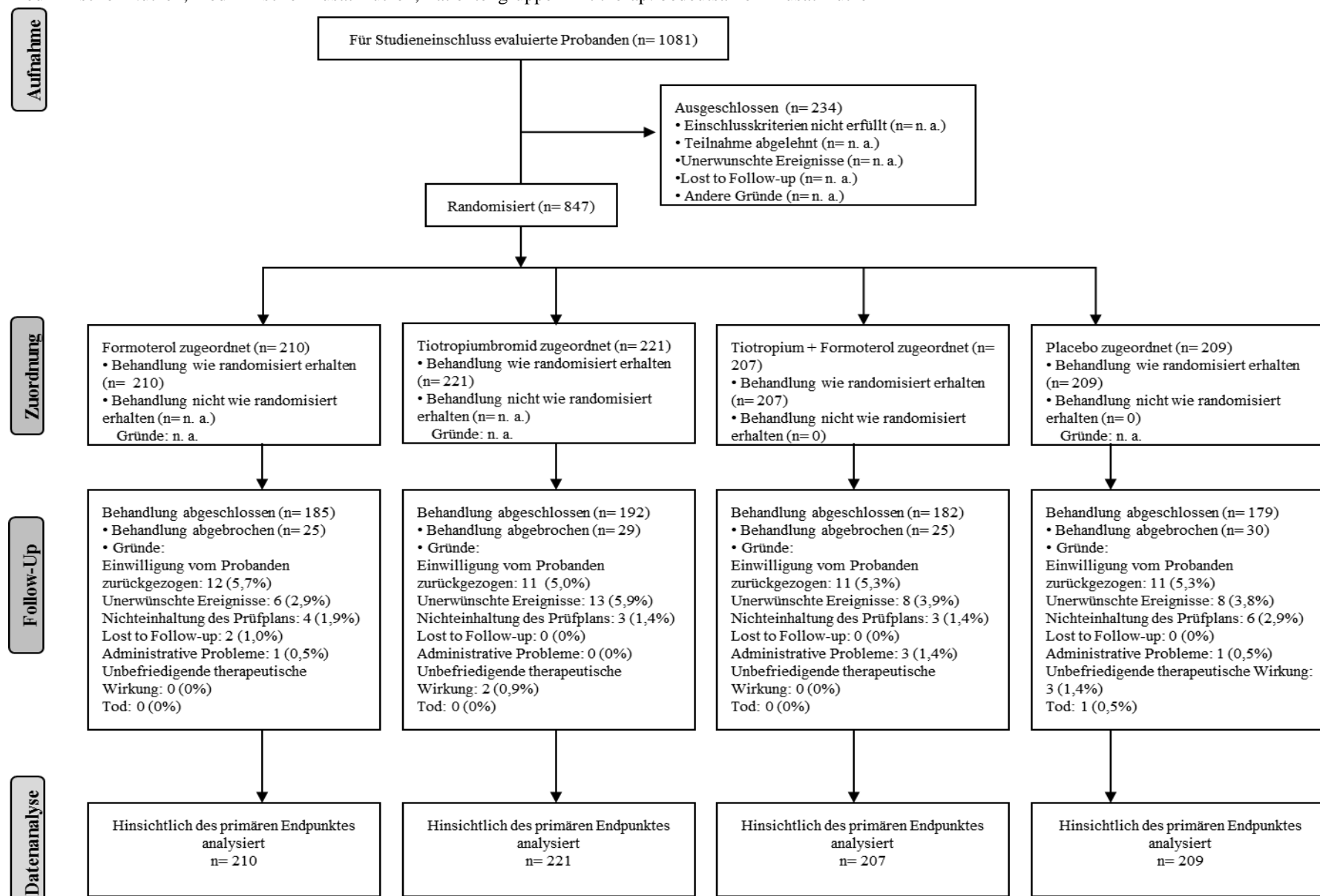


Tabelle 4.7-23 (Anhang): Studiendesign und -methodik für Studie Dusser 2006

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Untersuchung der Auswirkungen von Tiotropiumbromid auf das Vorkommen von Exazerbationen, die Inanspruchnahme von Gesundheitsressourcen (HRU) und die Einschränkung des Luftstroms bei Patienten mit COPD in Frankreich
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblind, Parallelgruppen, Randomisierung 1:1 zu Tiotropiumbromid 18 µg einmal täglich oder Placebo
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<p>Männliche und weibliche Patienten im Alter von >40 Jahren mit klinischer Diagnose von COPD (präbronchodilatorische FEV₁ 30-65 % des Sollwerts und FEV₁/langsame Vitalkapazität (SVC) <70 % des Sollwerts)</p> <p>Raucheranamnese ≥10 Packungsjahre und eine oder mehrere Exazerbationen im letzten Jahr, aber nicht innerhalb der letzten 6 Wochen vor Aufnahme in die Studie.</p> <p>Keine Vorgeschichte von Asthma, allergischer Rhinitis oder Atopie; keine Infektion der unteren Atemwege oder Exazerbation innerhalb der letzten 6 Wochen.</p> <p>Kein regelmäßiger Einsatz von Sauerstofftherapie tagsüber; kein Einsatz oraler Corticosteroide in instabilen Dosen 6 Wochen vor Aufnahme in die Studie oder in einer Dosis > 10 µg Prednison-Äquivalent täglich</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>177 Prüfzentren in Frankreich</p> <p>Sponsoren: Boehringer Ingelheim, Frankreich: Pfizer</p> <p>Einzelheiten zum publizierenden Autor:</p> <p>D. Dusser, Service de Pneumologie, Hopital Cochin, 27 Rue du Faubourg Saint Jacques, 75679 Paris Cedex 14 France daniel.dusser@cch.aphp.fr</p>

Item ^a	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Tiotropiumbromid 18 µg einmal täglich oder das Placebo, verabreicht mit dem HandiHaler (Boehringer Ingelheim, Ingelheim am Rhein, Deutschland)
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Wirksamkeitsendpunkte: <ul style="list-style-type: none"> • Exazerbationen der COPD, Krankenhauseinweisungen aufgrund einer COPD-Exazerbation, Begleitmedikationen und nicht anberaumte Arztbesuche werden bei Randomisierung (Tag 1) erfasst sowie nach 6, 12, 24, 36 und 48 Behandlungswochen. • Verbrauch von Gesundheitsressourcen • Expiratorischer Spitzenfluss (PEF) <ul style="list-style-type: none"> • Tägliche Messungen des expiratorischen Spitzenflusses (PEF) morgens vor Inhalation der Studienmedikation Die Patienten dokumentierten den höchsten der drei gemessenen Werte grafisch in ihrer Tagebuchkarte. Spirometrien wurden bei jedem Klinikbesuch durchgeführt. FEV₁, FVC, SVC und inspiratorische Kapazität (IC) wurden 30 Minuten vor Verabreichung der Dosis (im trough) aufgezeichnet. Alle Spirometrien wurden dreifach durchgeführt und der höchste Messwert wurde in den nachfolgenden Analysen verwendet. Sicherheit: <ul style="list-style-type: none"> • Unerwünschte Ereignisse Weitere Informationen über die Validierung oder Optimierung der Ergebnismessung wurden nicht berichtet.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Ein Stichprobenumfang von 459 Patienten pro Gruppe wurde als ausreichend eingestuft, um eine Differenz von 15 l/min. im mittleren wöchentlichen PEF am Morgen bei einer Standardabweichung von 70 l/min. zu detektieren. Um eine ausreichende Sicherheitsexposition sicherzustellen, wurde ein Stichprobenumfang von 1.000 Patienten (500 pro Gruppe) gewählt.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und	n. a.

Item ^a	Charakteristikum	Studieninformation
	Kriterien für einen vorzeitigen Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Randomisierung 1:1 zu Tiotropiumbromid 18 µg einmal täglich oder Placebo
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid 18 µg einmal täglich oder das Placebo, verabreicht mit dem HandiHaler (Boehringer Ingelheim, Ingelheim am Rhein, Deutschland)
12	Statistische Methoden	
12a	Statistische Methoden zur	Die Kovarianzanalyse mit Termen für die Behandlung wurde für

Item ^a	Charakteristikum	Studieninformation
	Bewertung der primären und sekundären Zielkriterien	<p>den PEF und die spirometrischen Endpunkte durchgeführt, die Baseline-Daten wurden als Kovariaten verwendet. Beim PEF wurden die Mittelwerte der Messungen, die in der Woche vor dem Behandlungszeitraum durchgeführt wurden, als Baseline-Daten verwendet. Eine ITT-Methode unter Verwendung des letzten beobachteten Wertes (LOCF-Methode) wurde für die Analysen der spirometrischen Endpunkte verwendet, außer bei Patienten, die aufgrund einer sich verschlechternden COPD aus der Studie ausschieden. Im letzteren Fall wurden die ungünstigsten Daten vor dem Ausscheiden verwendet. Ein ITT-Ansatz mit geschätzten fehlenden Daten wurde für die Analyse des PEF verwendet, wobei die lineare Interpolation zur Schätzung fehlender Messungen diente. Waren jedoch Daten nicht für Nachbarwochen verfügbar, kam der LOCF-Ansatz zur Anwendung. Wenn Patienten aufgrund einer COPD-Verschlechterung aus der Studie ausschieden, wurden die ungünstigsten Werte vor dem Abbruch verwendet.</p> <p>Die Prozentsätze der Patienten mit einer oder mehreren COPD-Exazerbationen oder Krankenhauseinweisungen aufgrund einer COPD-Exazerbation im Verlauf der einjährigen Studie wurden mit dem Fisher-Exact-Test analysiert.</p> <p>Die Anzahl der COPD-Exazerbationen, die Exazerbationstage, Krankeneinweisungen, Krankenhaustage, Arztbesuche, Telefon-anrufe, Behandlungszyklen mit oralen Steroiden und Antibiotika wurden mit dem Wilcoxon-Mann-Whitney-Test analysiert. Die Dauer der Gabe von oralen Steroiden und Antibiotika wurde ebenfalls mit dem Wilcoxon-Mann-Whitney-Test analysiert. Die Zeit bis zur ersten COPD-Exazerbation wurde mit dem Log-Rank-Test analysiert. Der Prozentsatz von Patienten, die zusätzliche Medikamente für die Atemwege während einer Exazerbation benötigten, wurde mit dem Chi-Quadrat-Test analysiert.</p> <p>Die statistische Signifikanz wurde bei $p < 0,05$ festgelegt.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Um die Ergebnisse dieser Studie unmittelbarer mit denen früherer Exazerbationsstudien vergleichen zu können, wurde eine Post-hoc-Analyse durchgeführt, in der eine allgemeinere Klassifizierung der Exazerbationsschwere auf Grundlage des Verbrauchs von Gesundheitsressourcen und des Einsatzes von Medikamenten angewendet wurde
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe,	Randomisiert: Tiotropiumbromid 18 µg einmal täglich (500), Placebo (510)

Item ^a	Charakteristikum	Studieninformation
	<p>die</p> <p>a) randomisiert wurden,</p> <p>b) tatsächlich die geplante Behandlung/Intervention erhalten haben,</p> <p>c) in der Analyse des primären Zielkriteriums berücksichtigt wurden</p>	<p>Erhielten die vorgesehene Behandlung: n. a.</p> <p>Analysiert in Bezug auf das Primärergebnis: n. a.</p>
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Tiotropiumbromid 18 µg einmal täglich (117)</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse: 15 • Sonstige Gründe: 102 <p>Placebo (147)</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse: 18 • Sonstige Gründe: 129
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

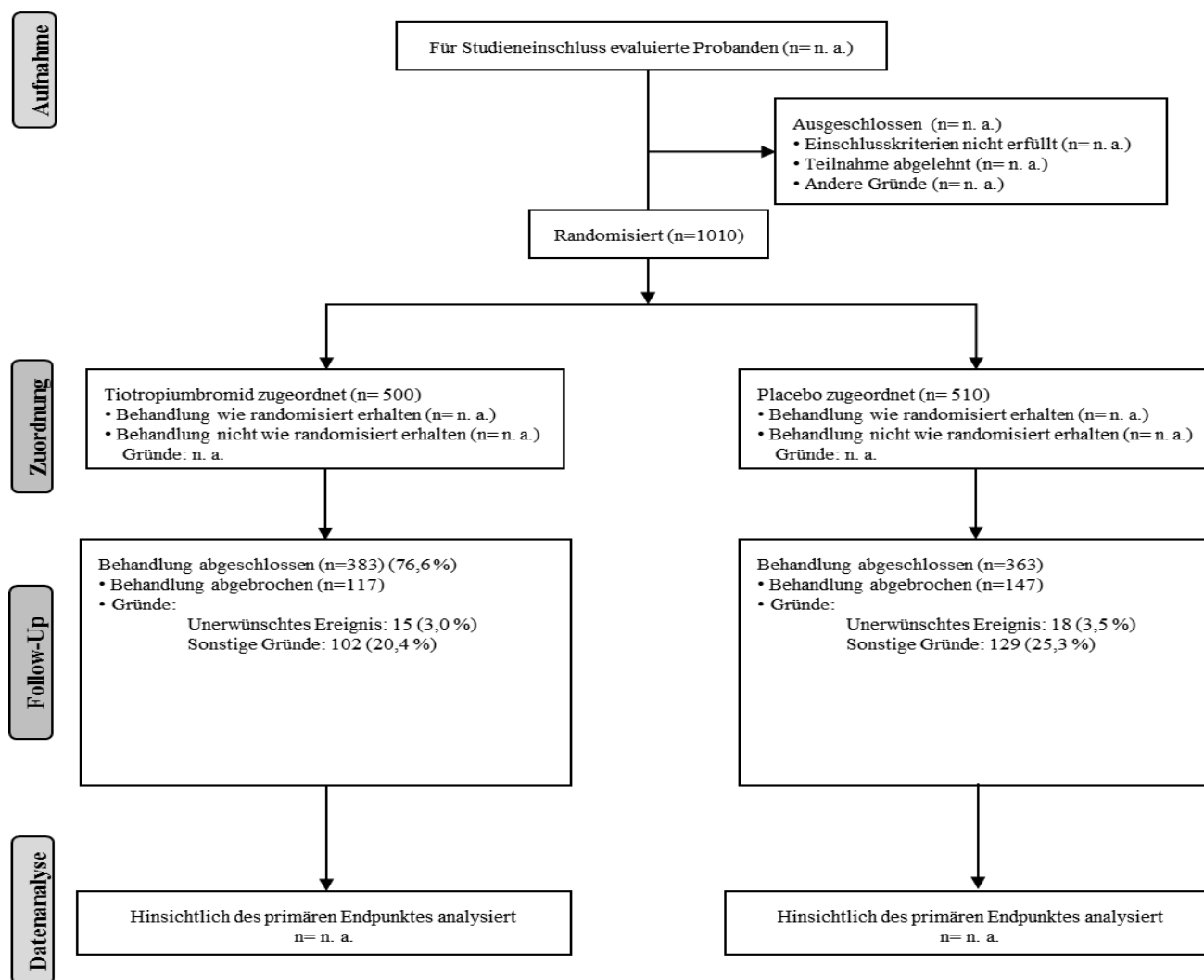


Tabelle 4.7-24 (Anhang): Studiendesign und -methodik für Studie Powrie 2007

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Erstmalige Untersuchung der Wirkung von inhaliertem Tiotropiumbromid auf die Atemwege und die systemischen Entzündungsmarker bei COPD in einer einjährigen, randomisierten, placebo-kontrollierten Studie
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Doppelblind, Parallelgruppen, RCT, Randomisierung im Verhältnis 1:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Patienten im Alter von ≥ 40 Jahren mit diagnostizierter COPD ($FEV_1 < 80\%$ des Sollwerts und $FEV_1/FVC < 70\%$), Raucheranamnese ≥ 10 Packungsjahren. Keine Vorgeschichte von Asthma oder Atopie oder Langzeit-Sauerstofftherapie.
4b	Studienorganisation und Ort der Studiendurchführung	Allgemeinarztpraxen oder ambulante Abteilung des London Chest Hospital Die Studie ist bei clinicaltrials.gov unter NCT00405236 registriert (laut clinicaltrials.gov waren die Royal Free and University College Medical School und Boehringer Ingelheim Pharmaceuticals die Sponsoren und führten die Studie gemeinsam durch). Einzelheiten zum Korrespondenzautor: J.A. Wedzicha Academic Unit of Respiratory Medicine Royal Free and University College Medical School, Rowland Hill Street London, NW3 2PF, UK E-Mail: j.a.wedzicha@medsch.ucl.ac.uk
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Tiotropiumbromid 18 μ g einmal täglich oder das Placebo mit dem HandiHaler-Inhalator

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt: Konzentration von Interleukin-6 (IL-6) im Sputum</p> <p>Sekundäre Wirksamkeit: Sputum-Interleukin-8 (IL-8) und Myeloperoxidase (MPO), Serum-IL-6 und C-reaktives Protein (CRP), Bakterienkolonisation im Sputum, FEV₁ und Exazerbationshäufigkeit</p> <p>Informationen über die Validierung oder Optimierung der Ergebnismessung wurden nicht berichtet.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	n. a.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	n. a.
9	<p>Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment)</p> <p>Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war</p>	n. a.
10	Randomisierung, Durchführung	n. a.

Item ^a	Charakteristikum	Studieninformation
	Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid 18 µg einmal täglich oder Placebo mit dem HandiHaler-Inhalator
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die Analysen wurden mit der Full-Analysis-Set-Population (alle randomisierten, behandelten Patienten mit Wirksamkeitsdaten) mit einer Kovarianzanalyse durchgeführt, mit Anpassungen nach Raucherstatus und Exazerbationen im vergangenen Jahr (<3 oder ≥3), die bei Rekrutierung aufgezeichnet wurden.</p> <p>Bei Sputum-Markern wurde die Fläche unter der Kurve (AUC) für jeden Patienten berechnet, wobei fehlende Daten durch Interpolation ersetzt wurden oder der letzte gemessene Wert verwendet wurde. Die Analyse verwendete die Entzündungsmarkerspiegel bei Baseline als Kovariate. Die AUC für IL-6 und MPO wiesen Schiefen auf und wurden logarithmisch transformiert (log10).</p> <p>Bei der Lungenfunktion wurden die Veränderungen zwischen Beginn und Ende der Studie verglichen. Die präspezifizierte AUC-Datenanalyse wurde aufgrund ihres Potenzials gewählt, die Schätzung des Gesamteffekts von Tiotropiumbromid auf das Sputum und die Serumzytokine im Verlauf eines Jahres zu ermöglichen, statt zu einem einzelnen Zeitpunkt.</p> <p>Für systemische Entzündungsmarker erfolgten in Woche 4 und 16 keine Probenahmen, daher erfolgten die Vergleiche durch den</p> <p>Wilcoxon Signed-Rank-Test für die Veränderungen von Baseline bis zur letzten Probe.</p> <p>Die Wirkung von Tiotropiumbromid auf individuelle jährliche Exazerbationsraten wurde mit einem Wilcoxon-Test überprüft. Differenzen bei der Zeit bis zur nächsten Exazerbation wurden</p>

Item ^a	Charakteristikum	Studieninformation
		mit einem Log-Rank-Test überprüft.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Bei der Lungenfunktion wurde eine zusätzliche Analyse mit Hilfe von Querschnitt-Regressionsmodellen durchgeführt
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Randomisiert: Tiotropiumbromid 18 µg einmal täglich (69), Placebo (73) Erhielten die vorgesehene Behandlung: Tiotropiumbromid 18 µg einmal täglich (69), Placebo (73) Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 18 µg einmal täglich (55), Placebo (58)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Tiotropiumbromid 18 µg einmal täglich (n=21) unerwünschtes Ereignis 7, sonstige Gründe 14 Placebo (n=21), unerwünschtes Ereignis 14, sonstige Gründe 7
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010. n. a. = nicht angegeben		

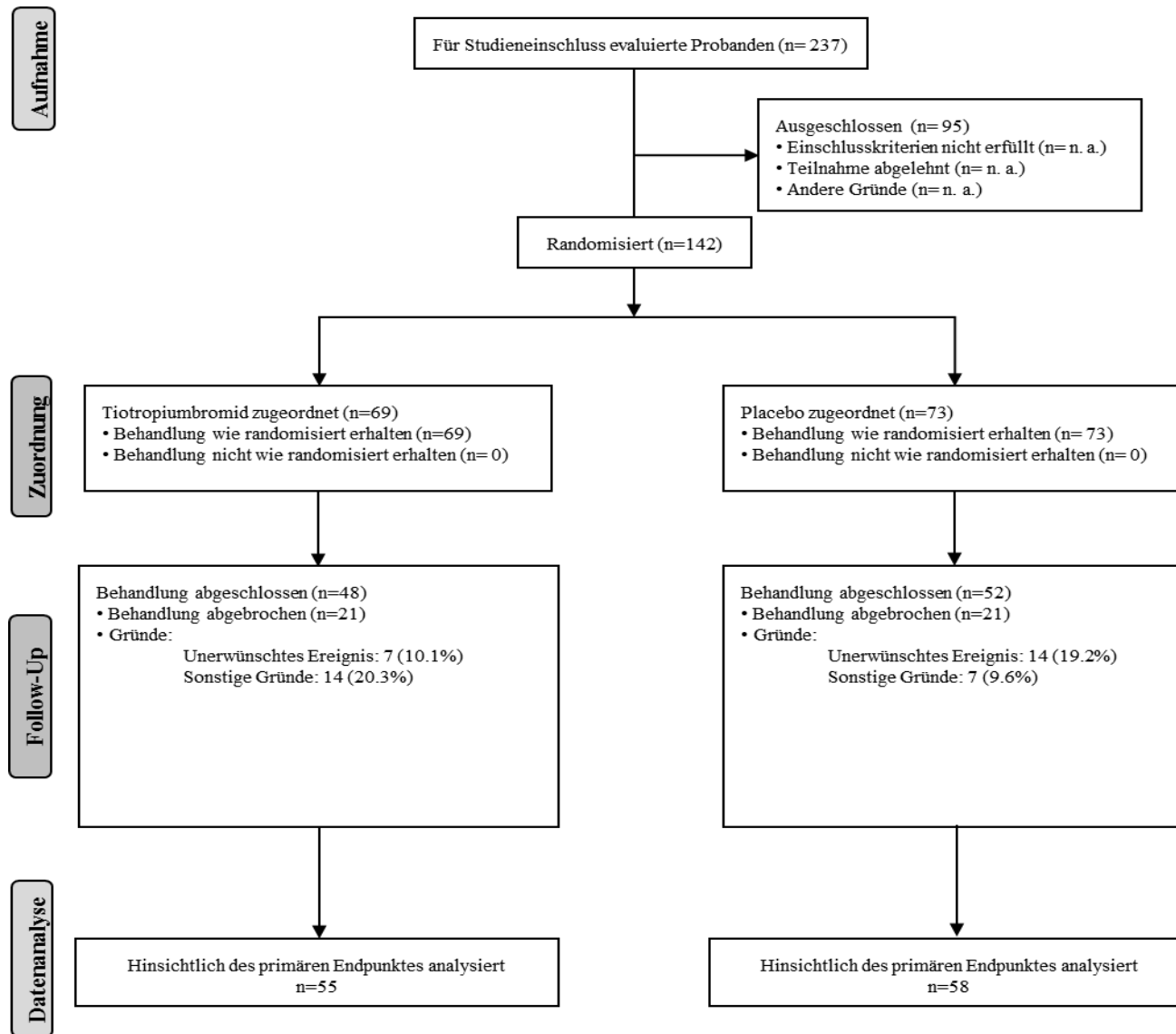


Tabelle 4.7-25 (Anhang): Studiendesign und -methodik für Studie Batemen 2010a

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Beurteilung der langfristigen Wirksamkeit und Sicherheit von zwei Dosen Tiotropiumbromid einmal täglich versus Placebo, verabreicht mit dem Respimat® Soft-Mist-Inhaler (SMI)
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Zwei identisch designte doppelblinde Parallelgruppenstudien
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Männer und Frauen im Alter von ≥ 40 Jahren mit diagnostizierter COPD und stabiler, mittelschwerer bis schwerer Atemwegsobstruktion (präbronchodilatorisches forciertes expiratorisches Volumen in 1 Sekunde [FEV ₁] ≤ 60 % des Sollwerts und FEV ₁ ≤ 70 % der forcierten Vitalkapazität [FVC]) und Raucheranamnese ≥ 10 Packungsjahren. <u>Ausschlusskriterien:</u> Bekanntes Hypersensitivität gegen Anticholinergika oder einen Bestandteil der Respimat® Inhalationslösung; Arzneimittelkontraindikationen ggü. Anticholinergika; vorherige Anwendung von Spiriva® HandiHaler®; regelmäßige Sauerstofftherapie tagsüber, orale β -Adrenergika oder langwirkende β -Adrenergika oder signifikanter Alkohol- oder Drogen-/Medikamentenmissbrauch
4b	Studienorganisation und Ort der Studiendurchführung	Multizentrisch, multinational Korrespondenz-Anschrift: Lesley Towse Clinical Research Department, Medical Division, Boehringer Ingelheim Ltd., Ellesfield Avenue, Bracknell, Berkshire, RG12 8YS, UK Tel.: +44 (0) 1344 741264 Fax: +44 (0) 1344 741657 lesley.towse@bra.boehringeringelheim.com

Item ^a	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Tiotropiumbromid Respimat® SMI 5 µg (2 Hübe mit 2,5 µg Tiotropiumbromid [Tio R5]), Tiotropiumbromid Respimat® SMI 10 µg (2 Hübe mit 5 µg Tiotropiumbromid [Tio R10]) oder Respimat® Placebo (2 Hübe der Placebo-Inhalationslösung) einmal täglich morgens
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Co-primäre Wirksamkeitsendpunkte:</p> <p>Trough-FEV₁-Reaktion in Woche 48 (FEV₁ 24 Stunden nach Verabreichung (im Trough) ausgedrückt als Veränderung gegenüber dem Baseline-FEV₁ vor Verabreichung der Substanz)</p> <ul style="list-style-type: none"> • St George's Respiratory Questionnaire (SGRQ) Gesamtscore bei Ende des 48-wöchigen Behandlungszeitraums • Fokal-Score des Mahler Transitional Dyspnoea Index (TDI) nach 48 Behandlungswochen • COPD-Exazerbationen pro Patientenjahr <p>Sekundäre Wirksamkeitsendpunkte:</p> <p>FVC, maximaler expiratorischer Fluss (PEFR) und mittlere wöchentliche Anzahl der Situationen, in denen Bedarfsmedikamente angewendet wurden.</p> <p>Physician's Global Evaluation (PGE) und Patient's Global Rating (PGR).</p> <p>Detaillierte Informationen zu den COPD-Exazerbationen und exazerbationsbezogenen Krankenhausaufenthalten.</p> <p>Spirometrie, gesundheitsbezogene Lebensqualität (HRQoL) und Tagebuchaufzeichnungen der Patienten aus dem gesamten 48-wöchigen Behandlungszeitraum.</p> <p>Sicherheit:</p> <p>Unerwünschte Ereignisse und unerwünschte Ereignisse mit Todesfolge während der Anlaufphase, dem 48-wöchigen Behandlungszeitraum und der 30-tägigen Nachbeobachtungsphase.</p> <p>Veränderungen bei Studienende (Woche 48) bei Vitalzeichen, standardmäßigem Routine-Labor und körperliche Untersuchung.</p> <p>Informationen über die Validierung oder Optimierung der Ergebnismessung wurden nicht berichtet.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit	n. a.

Item ^a	Charakteristikum	Studieninformation
	Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Jede Studie verfügte über eine ausreichende Aussagekraft zur Ermittlung von Differenzen bei den drei koprimary Endpunkten (Trough-FEV ₁ -Reaktion, SGRQ-Gesamtscore und Mahler TDI-Fokal-Score), hatten jedoch keine ausreichende Aussagekraft für Tests in Bezug auf Differenzen bei Exazerbationen zwischen den Gruppen. Dies wurde in den Prüfplänen beschrieben und präspezifiziert. Die FDA genehmigte die Kombination der Studien. In beiden Einzelstudien wurde ein Stichprobenumfang von 810 Patienten als ausreichend angesehen, um eine Differenz von 0,13 l der mittleren Trough-FEV ₁ (5 %-iges Signifikanzniveau) zu detektieren. Die kombinierte Analyse verfügte über eine Aussagekraft von mindestens 95 % zur Detektion einer Differenz von 0,05 l bei der mittleren Trough-FEV ₁ -Reaktion (5 %-iges Signifikanzniveau). Die kombinierte Analyse war ausreichend zur Detektion einer Differenz von 0,4 Exazerbationen pro Jahr mit einer Aussagekraft von 76 % (5 %-iges Signifikanzniveau) sowie einer mittleren Differenz beim Mahler TDI-Fokal-Score von 1 Einheit mit einer Aussagekraft von 90 % und einer mittleren Differenz von 4 Einheiten beim Gesamtscore des SGRQ mit einer Aussagekraft von 96 % (5 %-iges Signifikanzniveau).
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	n. a.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.

Item ^a	Charakteristikum	Studieninformation
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblinde Studie.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid Respimat® SMI 5 µg (2 Hübe mit 2,5 µg Tiotropiumbromid [Tio R5]), Tiotropiumbromid Respimat® SMI 10 µg (2 Hübe mit 5 µg Tiotropiumbromid [Tio R10]) oder Respimat® Placebo (2 Hübe der Placebo-Inhalationslösung) einmal täglich morgens Die Interventionen erfolgten mit denselben Inhalatoren (Aussehen und Zusammensetzung nicht unterscheidbar)
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Drei der koprimary Endpunkte (Trough-FEV₁-Reaktion, Gesamtscore des SGRQ und Mahler-TDI-Fokal-Score nach 48 Wochen) wurden mit einer Kovarianzanalyse (ANCOVA) mit festen Effekten analysiert, wobei der Raucherstatus bei Studieneintritt, das Zentrum und die Behandlungsgruppe als Terme dienten und der Baseline-Wert als kontinuierliche Kovariate.</p> <p>Das LSD-Verfahren nach Fisher (least significant difference procedure) wurde angewendet, wenn paarweise Vergleiche zwischen den Behandlungsgruppen vorgenommen wurden.</p> <p>Der verbleibende koprimary Endpunkt, COPD-Exazerbationen, wurde mit Wilcoxon-Mann-Whitney-Tests für paarweise Vergleiche bewertet. Kaplan-Meier-Schätzungen der Wahrscheinlichkeit keiner COPD-Exazerbation an einem beliebigen Testtag wurden durchgeführt.</p> <p>Die Analyse der COPD-Exazerbationen wurde mit einer gepoolten Analyse durchgeführt und da beide Studien in Design und Ausführung ähnlich waren, wurden die Studien nicht als Random-Effekt in dieses Modell aufgenommen. Paarweise Behandlungsvergleiche der Zeit bis zur ersten COPD-</p>

Item ^a	Charakteristikum	Studieninformation
		<p>Exazerbation wurden unter Anwendung des Log-Rank-Tests durchgeführt. Ähnliche Analysen erfolgten in Bezug auf Krankenhausaufenthalte.</p> <p>Ein ANCOVA-Modell wurde ebenfalls verwendet, um die sekundären Wirksamkeitsvariablen (FVC, PGE, COPD-Symptomscore, PEFR und Bedarfsmedikation) zu analysieren. Ein Varianzanalysemodell, das Terme für Raucherstatus bei Studieneintritt, Zentrum und Behandlungsgruppe beinhaltet, wurde bei den PGR-Daten (Patient's Global Rating) verwendet.</p> <p>Nur die Patienten mit Baseline-Daten und Daten des Behandlungszeitraums bei mindestens einem primären Endpunkt wurden in die Wirksamkeitsanalysen aufgenommen. Randomisierte Patienten, die ≥ 1 Dosis der Studienmedikation erhielten, wurden in die Sicherheitsanalyse aufgenommen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Post-hoc-Subgruppenanalysen wurden zur Trough-FEV₁-Reaktion nach 48 Wochen durchgeführt, um herauszufinden, ob Patienten, die bei Aufnahme in die Studie inhalative Corticosteroide eingenommen haben, unterschiedlich angesprochen haben als jene Patienten, die keine inhalativen Corticosteroide einnahmen. Das Modell, das zur Analyse des koprimary Endpunkts der Trough-FEV₁-Reaktion nach 48 Wochen angewendet wurde, wurde auch bei dieser Subgruppenanalyse verwendet.</p>
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	<p>Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die</p> <p>a) randomisiert wurden,</p> <p>b) tatsächlich die geplante Behandlung/Intervention erhalten haben,</p> <p>c) in der Analyse des primären Zielkriteriums berücksichtigt wurden</p>	<p>Randomisiert: Tiotropiumbromid 5 µg einmal täglich (670), Tiotropiumbromid 10 µg einmal täglich (667), Placebo (653)</p> <p>Erhielten die vorgesehene Behandlung: Tiotropiumbromid 5 µg einmal täglich (670), Tiotropiumbromid 10 µg einmal täglich (667), Placebo (653)</p> <p>Analysiert in Bezug auf das Primärergebnis:</p> <p>Tiotropiumbromid 5 µg einmal täglich (Trough-FEV₁: 650, Exazerbation: 670);</p> <p>Tiotropiumbromid 10 µg einmal täglich (Trough-FEV₁: 644, Exazerbation: 667);</p> <p>Placebo (Trough-FEV₁: 603, Exazerbation: 653)</p>

Item ^a	Charakteristikum	Studieninformation
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Tiotropiumbromid 5 µg einmal täglich (115)</p> <p>unerwünschte Ereignisse: 67; Krankheitsverschlechterung während der Studie: 31, kein angegebener Grund: 17</p> <p>Tiotropiumbromid 10 µg einmal täglich (136)</p> <p>unerwünschte Ereignisse: 79; Krankheitsverschlechterung während der Studie: 34, kein angegebener Grund: 23</p> <p>Placebo (205)</p> <p>unerwünschte Ereignisse: 122; Krankheitsverschlechterung während der Studie: 92</p>
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	<p>Die Rekrutierung für die Studien fand von Januar 2003 bis Dezember 2005 statt.</p> <p>Eine zusätzliche retrospektive Studie wurde von März 2007 bis Januar 2008 durchgeführt, um Daten aller vorzeitig ausgeschiedenen Patienten für den beabsichtigten Behandlungszeitraum von 48 Wochen zum 30-tägigen Nachbeobachtungszeitraum zu erfassen, einschließlich Informationen über den Vitalstatus und die Veränderungen der Behandlung nach Absetzen der Studienmedikation.</p>
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

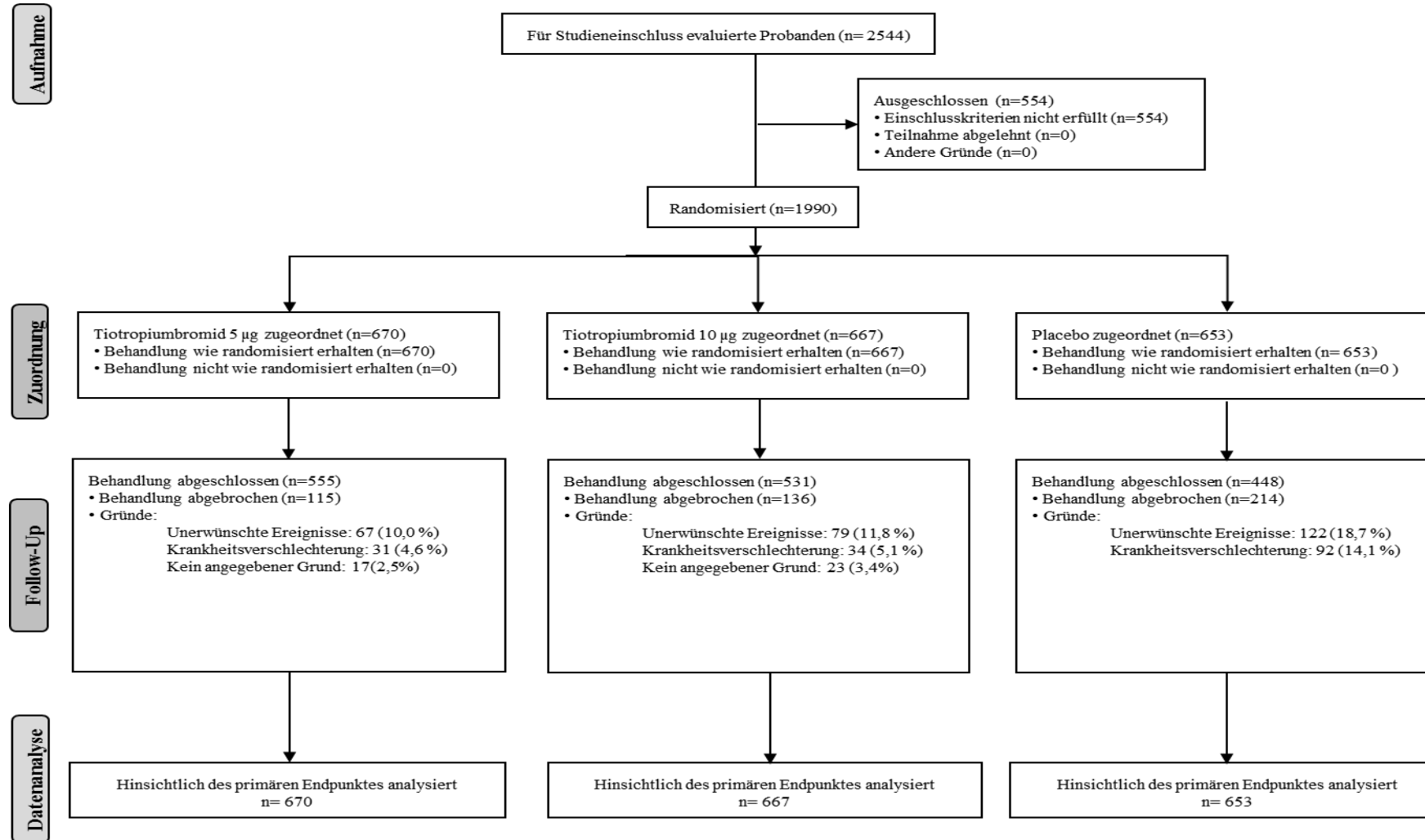


Tabelle 4.7-26 (Anhang): Studiendesign und -methodik für Studie Bateman 2010b

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Beurteilung der klinischen Wirksamkeit und Sicherheit einer einjährigen Behandlung mit Tiotropiumbromid 5 µg, verabreicht mit dem Respimat® Inhalator, bei COPD-Patienten, insbesondere zur Beurteilung der Wirkung auf Exazerbationen
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisiert, doppelblind, Parallelgruppendesign Zuteilungsverhältnis 1:1 zu Tiotropiumbromid 5 µg oder zum Placebo
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	COPD-Patienten beiderlei Geschlechts Im Alter von ≥40 Jahren, präbronchodilatorische FEV ₁ ≤60 % des Sollwerts und ein Verhältnis von FEV ₁ zur FVC von ≤70 % und Raucheranamnese ≥10 Packungsjahre Keine abnormen Ergebnisse bei Hämatologie, Urinanalyse, Blutchemieuntersuchungen, Vorgeschichte von Asthma, allergische Rhinitis oder Eosinophilenzahl von ≥600/mm ³ Keine vorausgehende Lungenresektion, Teilnahme an pulmonaren Rehabilitationsprogrammen in den letzten 6 Wochen oder regelmäßige Sauerstoffgabe tagsüber Keine instabile Arrhythmie-Anamnese, kein Myokardinfarkt in den letzten 6 Monaten oder Herzinsuffizienz mit erforderlichem Krankenhausaufenthalt in den letzten 12 Monaten
4b	Studienorganisation und Ort der Studiendurchführung	336 ambulante Zentren auf fünf Kontinenten in 31 Ländern Sponsor: Boehringer Ingelheim (Studiennummer 205.372). Einzelheiten zum publizierenden Autor: Eric Bateman, eric.bateman@uct.ac.za; Tel: +27 21 406 6901
5	Interventionen	Tiotropiumbromid 5 µg (zwei Hübe mit 2,5 µg) oder Placebo (zwei Hübe), beide Substanzen einmal täglich morgens inhaliert

Item ^a	Charakteristikum	Studieninformation
	Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	mit dem Respimat®-Inhalator
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primäre Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • Trough-FEV₁-Reaktion nach 48 Behandlungswochen • Zeit bis zur ersten COPD-Exazerbation <p>Sekundäre Wirksamkeitsendpunkte:</p> <ul style="list-style-type: none"> • Trough-FEV₁-Reaktion nach 4 und 24 Behandlungswochen • Trough-FVC-Reaktion nach 4, 24 und 48 Behandlungswochen • Anzahl von Exazerbationen pro Patient • Anzahl von Patienten mit mindestens einer Exazerbation • Zeit bis zur ersten Exazerbation, die einen Krankenhausaufenthalt erforderte • HRQoL-Veränderung versus Baseline unter Anwendung des SGRQ nach 24 und 48 Wochen <p>Sicherheit:</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse • Elektrokardiographie • Mortalität <p>Informationen über die Validierung oder Optimierung der Ergebnismessung wurden nicht berichtet.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Bei einer angenommenen Standardabweichung der Trough-FEV ₁ von 0,229 l würden 1.500 abgeschlossene Patienten pro Behandlungsgruppe die Ermittlung einer Differenz zwischen den

Item ^a	Charakteristikum	Studieninformation
		Studienmedikationen von 0,13 l bei der mittleren Trough-FEV ₁ -Reaktion, mit einem Signifikanzniveau von 5 % (zweiseitig) und einer statistischen Aussagekraft von mindestens 99 % ermöglichen. Diese Zahl von Patienten ermöglicht ebenfalls die Ermittlung einer proportionalen Hazardrate von 0,85 für die Zeit bis zur ersten Exazerbation in einem Log-Rank-Test, mit einem Signifikanzniveau von 5 % und einer statistischen Aussagekraft von 80 %, ausgehend von der Annahme, dass der Anteil von Patienten, bei denen es zu ≥ 1 Exazerbation kommt, 37,2 % und 42,1 % mit Tiotropiumbromid bzw. Placebo beträgt
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Zuteilung der Behandlung erfolgte mit einem computer-generierten Randomisierungscode.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgte stratifiziert nach Prüfzentrum und in den Zentren und wurde in Blöcken durchgeführt, um zu jeder Zeit eine ausgewogene Verteilung bei den Behandlungsgruppen sicherzustellen.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Personen, die direkt mit der Durchführung und Analyse der Studie befasst waren, hatten vor Abschluss der Studie keinen Zugriff auf die Zuteilungssequenz.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Zuteilung zur Behandlung erfolgte mit einem von Boehringer Ingelheim bereitgestellten und computer-generierten Randomisierungscode.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung	Prüfärzte, Bewerter und Patienten waren in Bezug auf die Identität der Medikamente verblindet

Item ^a	Charakteristikum	Studieninformation
	durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid 5 µg (zwei Hübe mit 2,5 µg) oder Placebo (zwei Hübe), beide Substanzen einmal täglich morgens inhaliert mit dem RespiMat®-Inhalator. Die Inhalatoren waren identisch und für Patienten und Prüfer nicht unterscheidbar.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Veränderungen von Trough-FEV₁ und FVC wurden mit einer Kovarianzanalyse (ANCOVA) mit Termen für Prüfzentrum, LABA-Einsatz und Behandlung als feste Effekte analysiert, der Baseline-FEV₁-Wert und der Baseline-FVC-Wert dienten als Kovariaten.</p> <p>Die Zeit bis zur ersten Exazerbation wurde mit dem proportionalen Cox-Hazard-Modell mit Termen für Zentrum, Behandlung und LABA-Einsatz analysiert. Die zwei primären Co-Endpunkte wurden sequenziell getestet, um den Fehler erster Art bei 0,05 zu halten. Die primären Co-Endpunkte wurden bei allen Patienten analysiert, die randomisiert den Behandlungen zugeteilt und behandelt wurden und eine Trough-FEV₁-Messung bei Baseline und mindestens einen nachfolgenden Besuch aufwiesen.</p> <p>Die Analyse weiterer Exazerbations-Endpunkte wurde mit den folgenden Tests durchgeführt: Log-Rank-Test für die Zeit bis zur Exazerbation, Wilcoxon-Mann-Whitney-Test und generalisiertes lineares Regressionsmodell für die Anzahl von Exazerbationen pro Patient und logistische Regression für den Anteil von Patienten, bei denen es zu mindestens einer Exazerbation kam. Die SGRQ-Scores wurden mit einem ANCOVA-Modell analysiert.</p> <p>In Bezug auf tödliche Ereignisse wurde die Cox-Regressionsanalyse zur Schätzung der Hazard-Ratio und des 95 %-igen Konfidenzintervalls für die Zeit bis zum Tod verwendet. Die Überlebenswahrscheinlichkeit wurde mittels Kaplan-Meier-Analyse und Log-Rank-Test analysiert.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	

Item ^a	Charakteristikum	Studieninformation
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Randomisiert: Tiotropiumbromid 5 µg einmal täglich (1.989), Placebo (2.002) • Erhielten die vorgesehene Behandlung: Tiotropiumbromid 5 µg einmal täglich (1.952), Placebo (1.965) • Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 5 µg einmal täglich (1.939), Placebo (1.953)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Tiotropiumbromid 5 µg einmal täglich (318)</p> <p>unerwünschte Ereignisse: 143; mangelnde Wirksamkeit: 28; Nichteinhaltung des Prüfplans: 47; Lost to Follow-up: 22; Einwilligung zurückgezogen: 20; sonstige Gründe: 58</p> <p>Placebo (373)</p> <p>unerwünschte Ereignisse: 156; mangelnde Wirksamkeit: 67; Nichteinhaltung des Prüfplans: 35; Lost to Follow-up: 28; Einwilligung zurückgezogen: 35; sonstige Gründe: 52</p>
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Die Rekrutierung begann im Oktober 2006 und endete im Dezember 2007
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

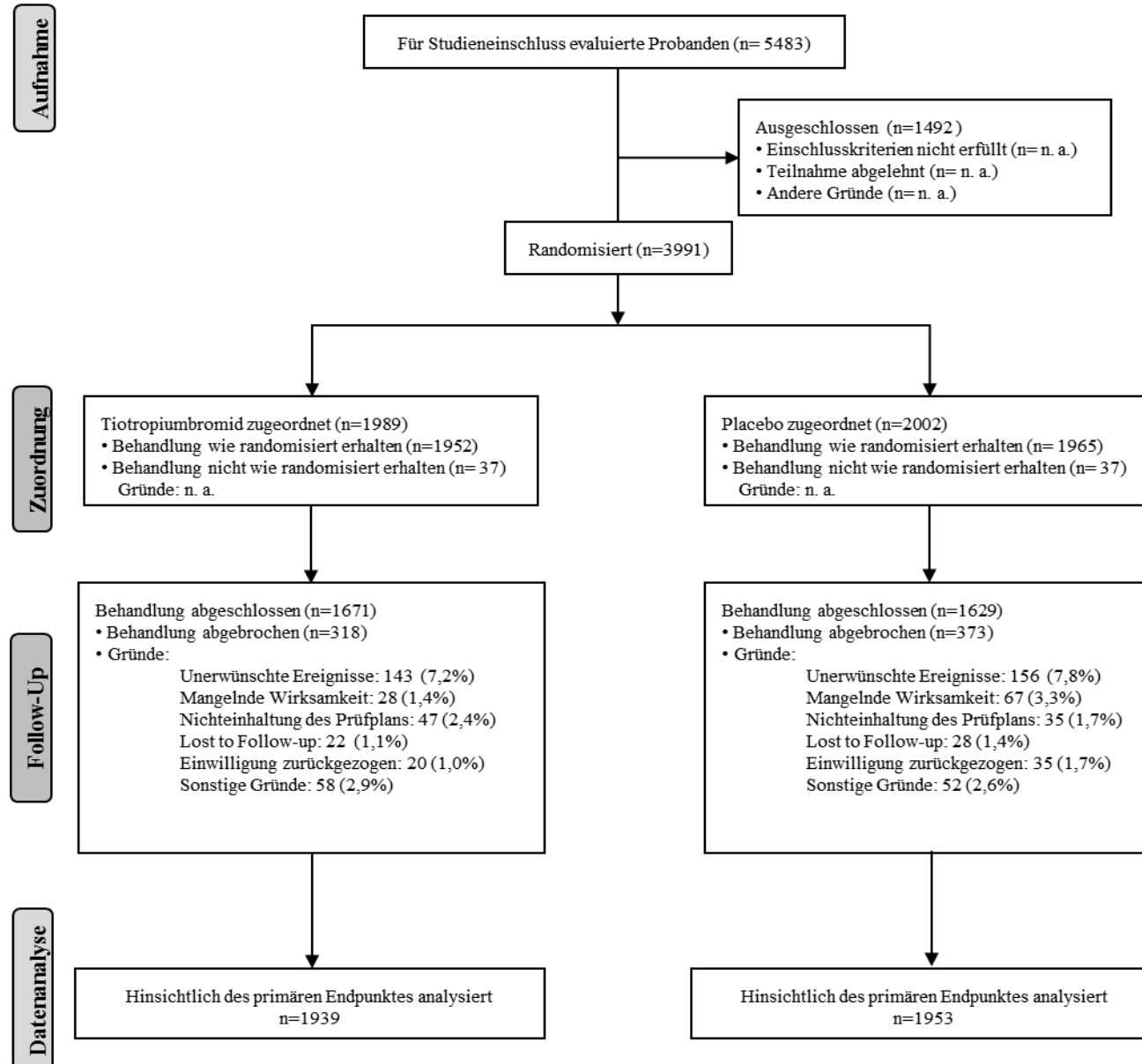


Tabelle 4.7-27 (Anhang): Studiendesign und -methodik für Studie Voshaar 2008

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Vergleich der Wirksamkeit und Sicherheit von zwei Dosen Tiotropiumbromid, verabreicht einmal täglich mit dem RespiMAT Soft-Mist-Inhaler mit Placebo (RCT1) und Ipratropium (RCT2), verabreicht viermal täglich mit einem treibgasbetriebenen Dosieraerosol (pMDI) zur Behandlung von Patienten mit COPD
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Zwei identische randomisierte, doppelblinde, Doppel-Dummy-Parallelgruppenstudien (RCT1 und RCT2). Randomisierung im Verhältnis 1:1:1:1 Insgesamt wurden 361 und 358 Patienten randomisiert und den Behandlungen in den Studien Nr. 205.251 bzw. Nr. 205.252 zugeteilt.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Männer und Frauen im Alter von ≥ 40 Jahren mit diagnostizierter COPD, moderater bis schwerer Atemwegsobstruktion und präbronchodilatatorischem FEV ₁ von ≤ 60 % des Sollwerts, FEV ₁ /forcierte Vitalkapazität (FVC) ≤ 70 % und Raucheranamnese ≥ 10 Packungsjahren. Keine Vorgeschichte von Asthma, allergischer Rhinitis oder anderer bedeutender Atemwegserkrankungen Keine bekannte Hypersensitivität gegenüber Anticholinergika, vorherige Anwendung von Tiotropium, regelmäßige Anwendung einer Sauerstofftherapie tagsüber, kein signifikanter Alkohol- oder Drogenmissbrauch, keine Teilnahme an einer anderen Studie
4b	Studienorganisation und Ort der Studiendurchführung	39 Prüfzentren in Deutschland, Italien, Südafrika und der Schweiz (RCT1) und 25 Prüfzentren in den USA und Kanada (RCT2). Sponsoren: Boehringer Ingelheim, Pfizer Korrespondenzautor: E Rubin. Respiratory, Clinical Operations, Boehringer Ingelheim Pharmaceuticals Inc., 900 Ridgebury Road, Ridgefield, CT 06877, USA. E-Mail-Adresse: erubin@rdg.boehringer-ingelheim.com (E. Rubin).

Item ^a	Charakteristikum	Studieninformation
		E. Rubin und L. Lowe sind Mitarbeiter von Boehringer Ingelheim.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Einmal täglich Tiotropiumbromid (je zwei Inhalationen zu je 5 oder 10 µg, verabreicht mit dem SPIRIVA Respimat Soft-Mist-Inhaler [SMI]) (Boehringer Ingelheim, Ingelheim am Rhein, Deutschland), inhaliertes Ipratropiumbromid (36 mg, verabreicht als treibgasbetriebenes Dosieraerosol [pMDI]) und Placebo (verabreicht mit Respimat SMI und pMDI).</p> <p>Tiotropiumbromid bzw. Tiotropiumbromid-Placebo wurden einmal täglich und Ipratropium bzw. das Ipratropium-Placebo wurden viermal täglich verabreicht. Tiotropiumbromid (5 µg [zwei Hübe mit 2,5 µg] oder 10 µg [zwei Hübe mit 5 µg]) wurden mit dem Respimat SMI einmal täglich morgens mit zwei Inhalationen des Placebos mittels pMDI verabreicht. Weitere zwei Inhalationen des Placebos mittels pMDI wurden mittags, am frühen Abend und vor dem Schlafengehen vorgenommen, um die Verblindung zu erreichen. Ipratropium (36 µg [zwei Hübe mit 18 µg]) wurden mit dem pMDI morgens, zusammen mit zwei Inhalationen des Placebos mittels Respimat SMI verabreicht. Weitere zwei Inhalationen von Ipratropium mittels pMDI wurden mittags, am frühen Abend und vor dem Schlafengehen vorgenommen. Für das Placebo wurden zwei Inhalationen mittels Respimat SMI und zwei Inhalationen mit dem pMDI morgens verabreicht. Weitere zwei Inhalationen des Placebos mittels pMDI wurden mittags, am frühen Abend und vor dem Schlafengehen vorgenommen.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Primärer Wirksamkeitsendpunkt: Trough-FEV₁ nach 12 Behandlungswochen</p> <p>Sekundäre Wirksamkeitsendpunkte: FVC, maximaler expiratorischer Fluss (PEFR) und die Anzahl der Patienten, die eine Zunahme von 15 % beim FEV₁ gegenüber Baseline erzielten</p> <p>Die wöchentliche mittlere Anzahl von Situationen pro Tag, in denen Bedarfsmedikation angewendet wurde; Schwere der COPD-Symptome auf Grundlage der Beurteilung der Patientenverfassung durch den Arzt und der allgemeinen Beurteilung der Erkrankung des Patienten durch den Arzt.</p> <p>Sicherheit: unerwünschte Ereignisse, Vitalzeichen, 12-Kanal-Elektrokardiogramme (EKG), Routine-Labor- und körperliche Untersuchungen</p> <p>Informationen zur Validierung oder Optimierung der Ergebnismessung wurden nicht berichtet. Die Literaturhinweise</p>

Item ^a	Charakteristikum	Studieninformation
		werden nach Ergebnis bereitgestellt und können Informationen über die Validität der Ergebnismessungen enthalten.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	n. a.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	N. z.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	n. a.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung	Sowohl der Prüfarzt als auch der Patient waren beim Randomisierungsprozess verblindet.

Item ^a	Charakteristikum	Studieninformation
	durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Tiotropiumbromid bzw. das Tiotropiumbromid-Placebo wurden einmal täglich und Ipratropium bzw. das Ipratropium-Placebo wurden viermal täglich verabreicht. Tiotropiumbromid (5µg [zwei Hübe mit 2,5 µg] oder 10 µg [zwei Hübe mit 5 µg]) wurden mit dem Respimat SMI einmal täglich morgens mit zwei Inhalationen des Placebos mittels pMDI verabreicht. Weitere zwei Inhalationen des Placebos mittels pMDI wurden mittags, am frühen Abend und vor dem Schlafengehen vorgenommen, um die Verblindung zu erreichen. Ipratropium (36 mg [zwei Hübe mit 18 mg]) wurden mit dem pMDI morgens, zusammen mit zwei Inhalationen des Placebos mittels Respimat SMI verabreicht. Weitere zwei Inhalationen von Ipratropium mittels pMDI wurden mittags, am frühen Abend und vor dem Schlafengehen vorgenommen. Für das Placebo wurden zwei Inhalationen mittels Respimat SMI und zwei Inhalationen mit dem pMDI morgens verabreicht. Weitere zwei Inhalationen des Placebos mittels pMDI wurden mittags, am frühen Abend und vor dem Schlafengehen vorgenommen.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p>Die Primäranalyse war ein schrittweises Verfahren für die sequenziellen Analysen jeder Tiotropium-Dosis, zuerst auf Überlegenheit gegenüber Placebo, dann auf Non-Inferiorität gegenüber Ipratropium und (ausgehend von erfolgreichen ersten Tests) auf Überlegenheit gegenüber Ipratropium (Testung von insgesamt sechs Hypothesen). Eine Kovarianzanalyse (ANCOVA) mit Termen für Raucherstatus, Behandlung und Zentrum und Baseline-Daten als lineare Kovariate wurde für den primären Endpunkt, die sekundären Spirometrie-Endpunkte, die allgemeine Beurteilung durch den Arzt und die COPD-Symptomscores durchgeführt. Separate ANCOVA-Modelle wurden für jeden Zeitpunkt und Testtag angepasst.</p> <p>Alle randomisierten Patienten mit beliebigen verfügbaren Daten wurden in die Sicherheitsbeurteilung aufgenommen und all jene mit Baseline-Daten und mindestens einer adäquaten Messung der Trough-Lungenfunktions-Parameter nach ≥ 5 Tagen der randomisierten Behandlung wurden in die Wirksamkeitsanalyse aufgenommen.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.

Item ^a	Charakteristikum	Studieninformation
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<ul style="list-style-type: none"> • Randomisiert: Tiotropiumbromid 5 µg (180), Tiotropiumbromid 10 µg (180), Ipratropium 36 µg (178), Placebo (181) • Erhielten die vorgesehene Behandlung: n. a. • Analysiert in Bezug auf das Primärergebnis: Tiotropiumbromid 5 µg (175), Tiotropiumbromid 10 µg (173), Ipratropium 36 µg (170), Placebo (171)
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Ein Ablaufdiagramm liegt nicht vor. Das folgende Ablaufdiagramm wurde auf Grundlage des Textes erstellt. Die Gründe für das Ausscheiden aus der Studie wurden bei einem Teil der Patienten berichtet.</p> <p>Tiotropiumbromid 5 µg (16)</p> <p>Unerwünschte Ereignisse: 13, sonstige Gründe: 3 Tiotropiumbromid 10 µg (18)</p> <p>Unerwünschte Ereignisse: 12, sonstige Gründe: 6 Ipratropium 36 µg (31)</p> <p>unerwünschte Ereignisse: 21, sonstige Gründe 10</p> <p>Placebo (22)</p> <p>Unerwünschte Ereignisse: 16, sonstige Gründe: 6</p>
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Die Studien fanden von November 2002 bis Dezember 2003 statt
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
a: nach CONSORT 2010.		
n. a. = nicht angegeben		

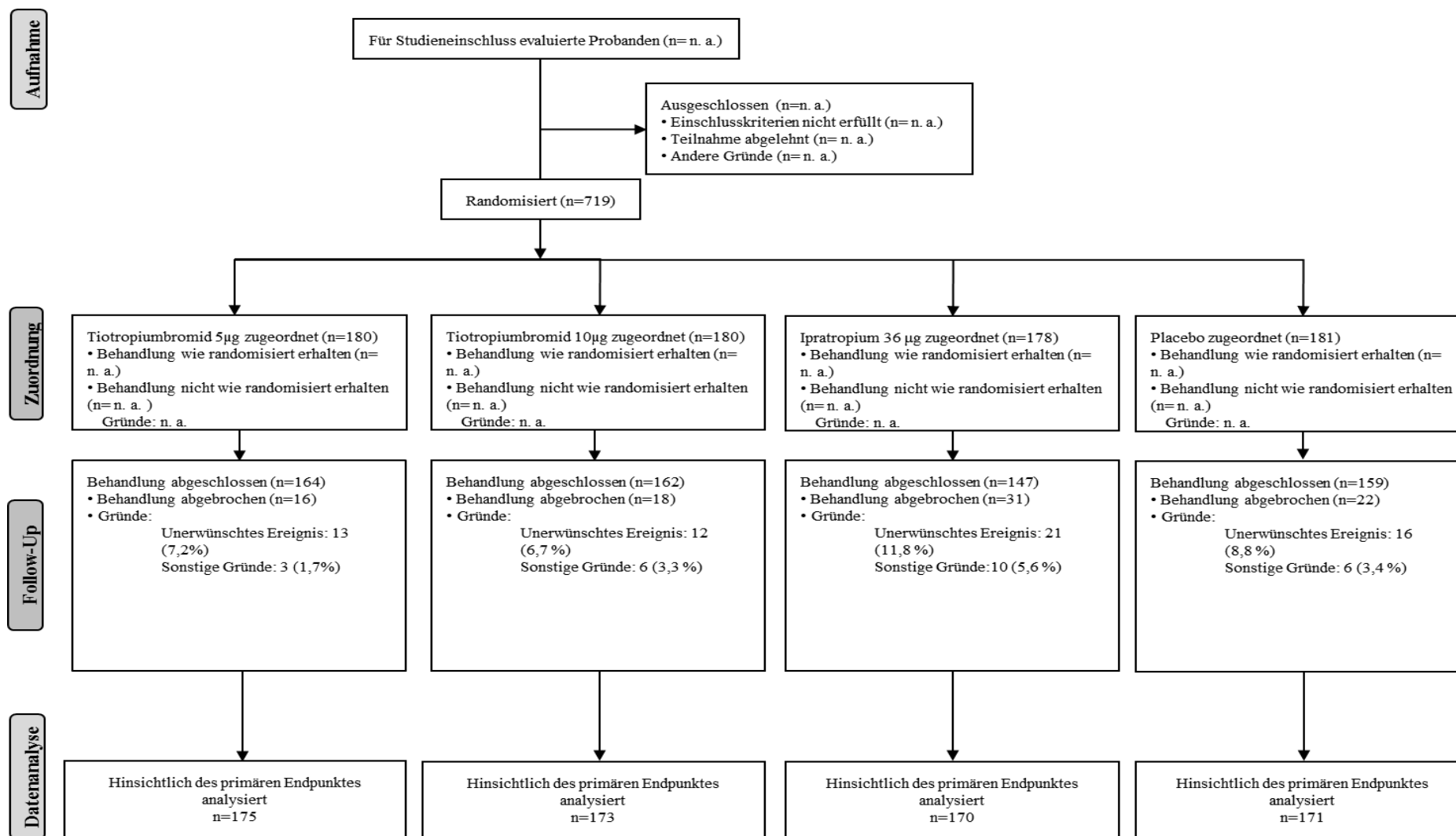


Tabelle 4.7-28 (Anhang): Studiendesign und -methodik für Studie Johansson 2008

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Ziel dieser Studie war die Beurteilung der Wirkung von Tiotropiumbromid 18 µg einmal täglich auf die Lungenfunktion, die klinischen Symptome und die HRQoL bei Patienten mit leichter COPD nach den schwedischen Richtlinien von 2003 (leichte bis moderate COPD nach GOLD-Richtlinien).
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, doppelblinde, multizentrische Parallelgruppen-Studie. Nach einem Screening-Besuch absolvierten die Patienten eine 2-wöchige Anlaufphase. Nach einer weiteren Beurteilung an Tag 1 wurden die geeigneten Patienten im Verhältnis 1:1 randomisiert und erhielten Tiotropiumbromid oder Placebo
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	<p>Einschlusskriterien: ambulante Patienten im Alter von >40 Jahren mit diagnostizierter leichter COPD nach den schwedischen Richtlinien von 2003 (postbronchodilatorische FEV₁/FVC <70 % und FEV₁ ≥60 % des Sollwerts); Raucheranamnese von ≥10 Packungsjahren und ein Dyspnoea-Score nach Medical Research Council (MRC) von ≥2.</p> <p>Ausschlusskriterien: Vorgeschichte von Asthma, allergischer Rhinitis oder Atopie; Eosinophilenzahl >600/mm³; Infektion der unteren Atemwege oder jede Exazerbation innerhalb der letzten sechs Wochen; Myokardinfarkt innerhalb der letzten sechs Monate; instabile kardiale Arrhythmie; regelmäßige Sauerstofftherapie; Anwendung von oralen oder inhalativen Steroiden (innerhalb der letzten drei Monate) und signifikante Erkrankungen außer COPD.</p>
4b	Studienorganisation und Ort der Studiendurchführung	<p>Die Studie wurde von März 2004 bis Juli 2005 in 27 Zentren in Schweden durchgeführt.</p> <p>Sponsoren: Boehringer Ingelheim, Pfizer</p> <p>Korrespondenzautor: Department of Public Health and Caring Sciences, Family Medicine and Epidemiology, Science Park, 751 85 Uppsala, Schweden.</p>

Item ^a	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Tiotropiumbromid 18 µg oder Placebo einmal täglich morgens 12 Wochen lang, verabreicht mit dem HandiHaler®-Inhalator (Boehringer Ingelheim).
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>Der primäre Endpunkt war die Veränderung der Fläche unter der Kurve beim FEV₁ (AUC vor der Verabreichung des Medikaments bis zwei Stunden nach Verabreichung zwischen Baseline und Woche 12). Weitere Endpunkte waren die FEV₁- und FVC-Trough-Reaktion, der Einsatz von Bedarfsmedikation und unerwünschte Ereignisse. Dyspnoe und HRQoL wurden anhand von Fragebögen bewertet.</p> <p>Spirometrien erfolgten an Tag 1, 15 (nach 2 Wochen) und 85 (nach 12 Wochen). FEV₁ und FVC wurden 10 Minuten vor und 0,5, 1 und 2 Stunden nach der Verabreichung der Dosis aufgezeichnet. Die FEV₁- und FVC-Trough-Reaktions-Werte waren die Werte vor Verabreichung der Dosis an Tag 15 und 85. Alle Spirometrien wurden dreifach durchgeführt und der höchste Messwert wurde für die Analyse verwendet. Ausstattung und Methodik basierten auf den Standards der ATS (1994) oder der European Respiratory Society (1993).</p> <p>Fragebögen zur Beurteilung der Dyspnoe und der HRQoL wurden bei Beginn und Ende der Studie ausgegeben. Die Dyspnoe wurde anhand des Dyspnoe-Index bei Baseline (BDI) und der MRC Dyspnoe-Skala beurteilt. Die Patienten definierten mit der MRC Dyspnoe-Skala auch ihre Wahrnehmung der Dyspnoe. Als explorative Messung wurde der BDI auch für die Beurteilung an Tag 85 anstelle des Transitional Dyspnoea Index verwendet. Die gesundheitsbezogene Lebensqualität (HRQoL) wurde mit dem generischen European Quality of Life Questionnaire, dem EuroQol (EQ-5D) ermittelt, der die EQ-5D Visual Analogue Scale (VAS) und den EQ-5D-Index umfasst.</p> <p>Die Patienten zeichneten die tägliche Anzahl der verwendeten Bedarfsmedikations-Dosen auf. Unerwünschte Ereignisse, einschließlich COPD-Exazerbationen, wurden aufgezeichnet.</p> <p>Informationen über die Validierung der Ergebnismessung wurden nicht mitgeteilt. Die Literaturhinweise werden nach Ergebnis bereitgestellt und können Informationen über die Validität der Ergebnismessungen enthalten.</p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	

Item ^a	Charakteristikum	Studieninformation
7a	Wie wurden die Fallzahlen bestimmt?	Einhundert Patienten pro randomisierter Behandlungsgruppe waren erforderlich, um mit einer Aussagekraft von 80 % eine Differenz zwischen den Gruppen von 120 ml beim primären Endpunkt (Veränderung der AUC _{0-2Std.} FEV ₁ versus Baseline-Wert nach 12 Behandlungswochen) mit einem 5 %-igen Signifikanzniveau im zweiseitigen t-Test zu detektieren. Der Stichprobenumfang basierte auf Daten einer vorausgehenden Studie zu Tiotropiumbromid mit ähnlichen Elementen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	n. a.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	n. a.
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht	Alle Beteiligten (Patienten, Ärzte, Ergebnisbewerter) waren verblindet. Es wurde nicht berichtet, wie die Verblindung erzielt wurde.

Item ^a	Charakteristikum	Studieninformation
	verblindet, wie wurde die Verblindung vorgenommen?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Patienten erhielten Tiotropiumbromid 18 mcg oder das Placebo, das mit dem HandiHaler®-Inhalator verabreicht wurde. Der HandiHaler war in beiden Gruppen vom Aussehen her identisch.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Kovarianzanalyse mit den Termenr Behandlungsdauer und gepooltes Zentrum wurde für alle kontinuierlichen Endpunkte verwendet, die Baseline-Daten dienten als Kovariaten. Für fehlende spirometrische Daten wurde der letzte beobachtete Wert verwendet, außer bei Patienten, die aufgrund einer sich verschlechternden COPD aus der Studie ausschieden. Hier wurden die ungünstigsten Werte vor dem Ausscheiden verwendet. Es wurde nicht berichtet, ob der ITT-Ansatz angewendet wurde.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	n. a.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Ein Ablaufdiagramm wurde nicht geliefert. Das im Folgenden präsentierte Ablaufdiagramm wurde auf Grundlage des Textes erstellt, kann jedoch unvollständig sein, da die Berichte über den Patientenfluss begrenzt waren. Randomisierte Patienten: 107 (Tiotropium) und 117 (Placebo). Anzahl der Patienten in der FEV ₁ -Bewertung: 105 (Tiotropium) und 115 (Placebo).
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	n. a. Die Berichte über den Patientenfluss waren begrenzt.
14	Aufnahme/Rekrutierung	

Item ^a	Charakteristikum	Studieninformation
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	Die Studie wurde von März 2004 bis Juli 2005 durchgeführt. Die Behandlungsdauer war 12 Wochen. Es wurden keine weiteren Informationen zur Verfügung gestellt.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		

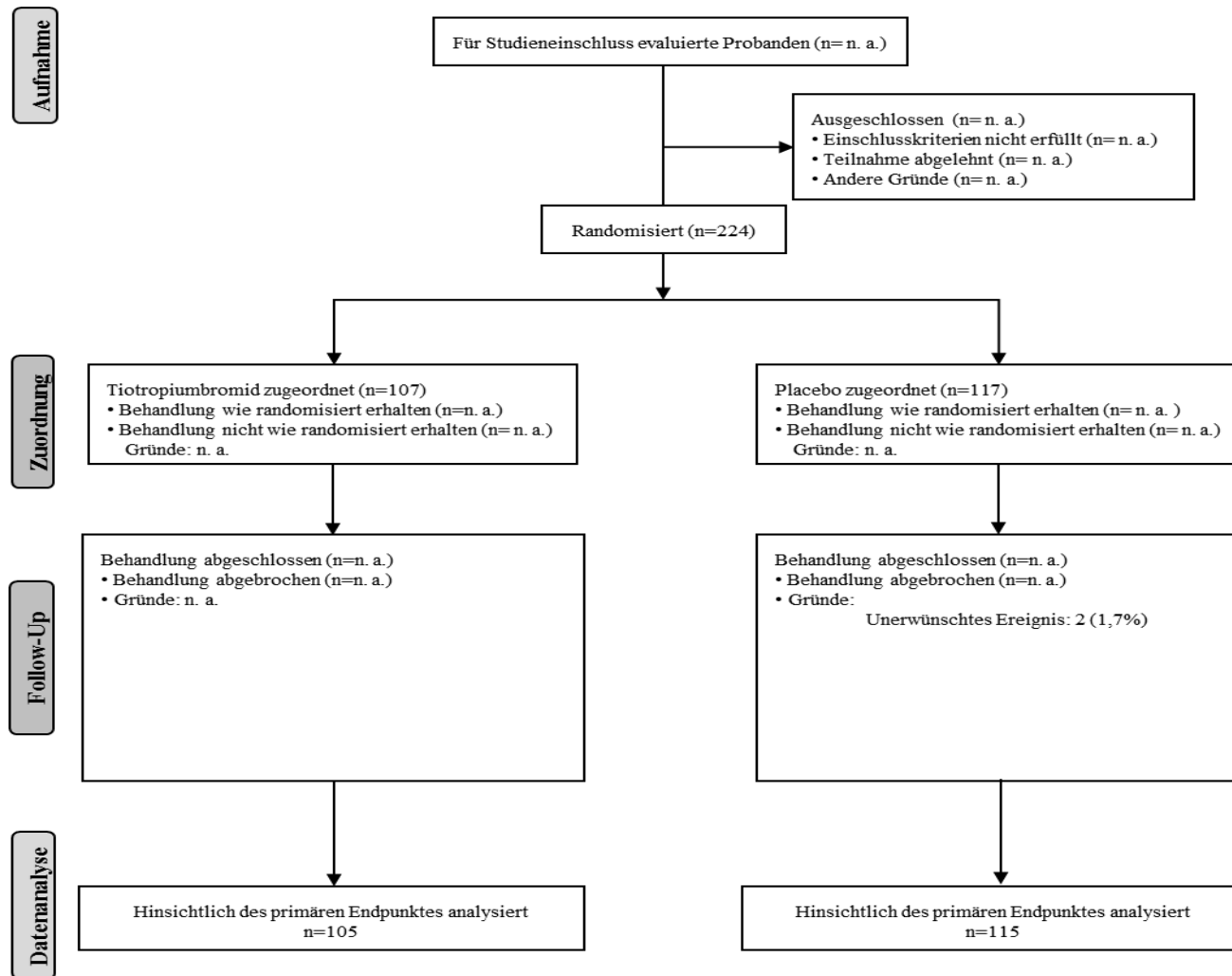


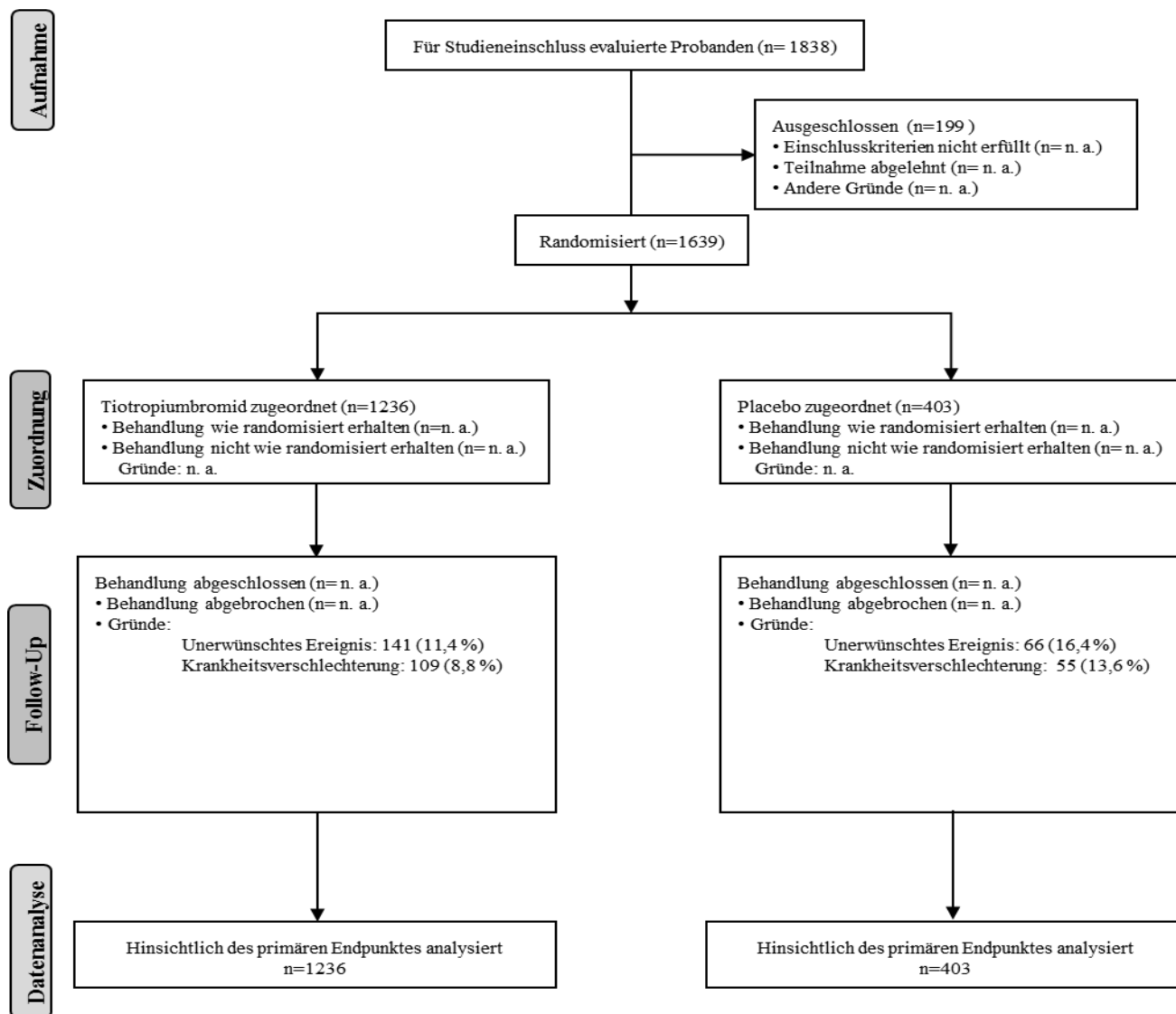
Tabelle 4.7-29 (Anhang): Studiendesign und -methodik für Studie Beeh 2006

Item ^a	Charakteristikum	Studieninformation
-	Studienziel	
2 b	Genauere Ziele, Fragestellung und Hypothesen	Prüfung der Wirksamkeit und Verträglichkeit von Tiotropiumbromid in einem großen Patientenkollektiv verschiedener COPD-Schweregrade unter praxisnahen Bedingungen in deutschen pneumologischen Facharztpraxen.
-	Methoden	
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisiert, doppelblinde Parallelgruppenstudie, Randomisierung im Verhältnis 3:1
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	n. a.
4	Probanden/Patienten	
4a	Ein-/Ausschlusskriterien der Probanden/Patienten	Es wurden COPD-Patienten verschiedener Schweregrade eingeschlossen eine stabile COPD mit einem $FEV_1 \leq 70\%$ vom Sollwert und einem FEV_1/FVC -Verhältnis $< 0,7$ Gefordert war eine Raucheranamnese ≥ 10 Packungsjahren und ein Alter ≥ 40 Jahren Kein Asthma oder Erkrankungen aus dem atopischen Formenkreis Keine Sauerstofflangzeittherapie oder ein respiratorischer Infekt in den sechs Wochen vor Screening
4b	Studienorganisation und Ort der Studiendurchführung	Die multizentrische Studie wurde in 294 pneumologischen Prüfzentren in Deutschland durchgeführt. Prüfzentren waren pneumologische Facharztpraxen oder pneumologische Klinikambulanzen sowie eine geringe Anzahl von internistischen Praxen mit Schwerpunkt Atemwegserkrankungen (< 10). Sponsor: Boehringer Ingelheim, Pfizer Korrespondenzadresse (Die Anschrift des Organizers wurde nicht gefunden): Priv.-Doz. Dr. med. Kai Michael Beeh, insaf Institut für Atemwegsforschung, Biebrichter Allee 34, 65187 Wiesbaden.

Item ^a	Charakteristikum	Studieninformation
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Jeweils morgens eine Kapsel Tiotropiumbromid 18 µg oder Placebo via HandiHaler
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	Primärer Wirksamkeitsendpunkt: • Der morgendliche Ausgangs- (Trough) FEV ₁ -Wert 23 – 24 Stunden nach der vorhergehenden Inhalation der Prüfmedikation und das FEV ₁ zwei Stunden nach Inhalation nach 12-wöchiger Behandlung mit Tiotropiumbromid gegenüber Placebo. Zusätzliche Wirksamkeitsendpunkte: • FVC, IVC, Exazerbationen, unerwünschte Ereignisse Validierung oder Optimierung der Ergebnismessung: nicht angegeben
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	n. a.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Fallzahlplanung ergab, dass mindestens 612 auswertbare Patienten in der Tiotropiumgruppe und 204 in der Placebogruppe benötigt werden, um einen Unterschied von 0,1 Liter in beiden Zielgrößen bei einer angenommenen Standardabweichung von 0,35 l mit 90 %iger Trennschärfe aufzuzeigen.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	n. a.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	n. a.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Tiotropiumbromid oder Placebo. Verhältnis 3:1
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation)	n. a.

Item ^a	Charakteristikum	Studieninformation
	concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax/Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	
10	Randomisierung, Durchführung Wer führte die Zuteilung durch, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	n. a.
11	Verblindung	
11a	Waren a) die Probanden/Patienten und/oder b) diejenigen, die die Intervention/Behandlung durchführten, und/oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Doppelblind
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Jeweils morgens eine Kapsel Tiotropiumbromid 18 µg oder Placebo via HandiHaler, derselbe Inhalator wurde in beiden Gruppen verwendet.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Die Zielgrößen wurden mittels Kovarianzanalyse mit Zentrum als zufälligem Einflussfaktor und Baseline als Kovariate auf Behandlungsunterschiede geprüft. Um den Gesamtfehler erster Art unter dem üblichen Signifikanzniveau von 5 % zu halten, wurde für die zwei primären Fragestellungen die Methode der alpha-Adjustierung nach Bonferoni-Holm angewendet. Die Anzahl der Exazerbationen wurde mit dem χ^2 -Test und die Zeit bis zu ihrem ersten Auftreten mit dem log-Rank-Test geprüft“. Die Prüfer definieren „Zentrum“ als das Prüfzentrum, in dem der Patient behandelt wurde. Auch in den anderen Studien war eine Anpassung in Bezug auf die Prüfzentren gängige Praxis. Die Auswertung folgte dem Intention-To-Treat Prinzip (wie randomisiert) mit dem Full-Analysis-Set (FAS) für die Parameter der Lungenfunktion und für die Erfassung der Exazerbationen für alle behandelten Patienten (Safety Set).

Item ^a	Charakteristikum	Studieninformation
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Unerwünschte Ereignisse wurden ebenfalls nach dem ITT-Prinzip für alle behandelten Patientengruppen ausgewertet.
-	Resultate	
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Randomisierung: Tiotropiumbromid 18 µg (1.236) oder Placebo (403) Geplante Behandlung erhalten: n. a. Analysiert für primären Endpunkt: Tiotropium: 1236, Placebo 403
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Tiotropiumbromid 18 µg (n. a.) Unerwünschte Ereignisse: 141; Krankheitsverschlechterung: 109 Placebo (n. a.) Unerwünschte Ereignisse: 66; Krankheitsverschlechterung: 55
14	Aufnahme/Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden/Patienten und der Nachbeobachtung	n. a.
14b	Informationen, warum die Studie endete oder beendet wurde	Geplant
<p>a: nach CONSORT 2010.</p> <p>n. a. = nicht angegeben</p>		



Anhang 4-G: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktspezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Tabelle 4-27 (Anhang 4G)

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie
<Studienbezeichnung>

Grundsätzliche Hinweise zum Ausfüllen des Bewertungsbogens

Dieser Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- **Kriterien zur Bewertung des Verzerrungspotenzials auf Studienebene.** In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- **Kriterien zur Bewertung des endpunktpezifischen Verzerrungspotenzials.** In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf.; auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich werden die Bögen studienbezogen ausgefüllt. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Tabelle 4.7-30 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie M/3473/23

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Fuhr R (2010): Clinical Trial M/3473/23 - A multiple dose, double blind, double-dummy, two-week 3 way cross-over, placebo controlled clinical trial to assess the efficacy and safety of twice daily inhaled aclidinium bromide 400 µg compared to placebo and to an active comparator in patients with stable moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 22/04/2010. Confidential.	M/3473/23
Fuhr R, Magnussen H, Sarem K, Ribera Llovera A, Kirsten AM, Falques M, et al. (2012): Efficacy of Aclidinium Bromide 400 µg Twice Daily Compared With Placebo and Tiotropium in Patients With Moderate to Severe COPD. Chest; 141(3):745-52.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Die Zuteilung zu den Behandlungsgruppen erfolgte randomisiert.

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Patienten erhielten über einen computergenerierten Randomisierungsplan eine Identifikationsnummer und wurden mittels eines ausgeglichenen Randomisierungsverhältnisses (1:1:1:1:1:1) einer der 6 Behandlungssequenzen (vgl. CONSORT-Statement Item 3a) zugeordnet.

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)

- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Therapiezuweisung erfolgte durch einen computergenerierten Randomisierungsplan und war gegenüber Prüfern und Patienten verblindet. Außerdem wurde ein Double-Dummy-Verfahren zur Sicherstellung der doppelblinden Durchführung des Versuchs angewandt. Zusätzlich wiesen die passenden Placebos zu Acclidiniumbromid und Tiotropium, abgesehen vom aktiven Wirkstoff, die gleiche äußere Erscheinungsform und die gleiche inhaltliche Zusammenstellung auf.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

- ja:** Die Patienten waren verblindet.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet. Alle Inhalatoren glichen sich in Bezug auf äußere Erscheinung und Zusammensetzung.

Behandler bzw. Weiterbehandler

- ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet. Randomisierungsdaten wurden streng vertraulich behandelt und waren bis zur Offenlegung der zugeordneten Behandlung aller Versuchspatienten nach der Schließung der Datenbank beim Abschluss der Studie nur berechtigten Personen zugänglich. Erst nach dem Abschluss der Studie und nach der Überprüfung und dem Abschluss der Datensammlung wurden die Randomisierungs-codes zur Datenanalyse zugänglich gemacht.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis

kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer

Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die relevanten Endpunkte wurden vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Transpatenter Patientenfluss (vgl. CONSORT flow chart). Es wurde keine Interimanalyse geplant und durchgeführt.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien,

statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan festgelegt und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Peak-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

 Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan festgelegt und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: FEV 1 AUC

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-

Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungszufriedenheit

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der

Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (serious adverse events, SAE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfer wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung ITT-Prinzip.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Anzahl der Studienabbrüche in den Therapiearmen ist mit der Anzahl der AEs und

SAEs konsistent.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Nächtliche und morgendliche Symptome

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:**Endpunkt: COPD-Symptome****1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-) Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Kardiovaskuläre Morbidität und Mortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung

angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Aclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation) Seite 686 von 1250
Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-31 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie M/3473/39

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Beier J (2011): Clinical Trial Protocol M/3473/39 - A multiple dose, double-blind, double-dummy, placebo controlled, parallel clinical trial to assess the efficacy and safety of twice daily inhaled aclidinium bromide 400 µg compared to placebo and to tiotropium bromide in patients with stable moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 06/05/2011. Confidential.	M/3473/39

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Die Zuteilung zu den Behandlungsgruppen erfolgte randomisiert.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

Aclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation) Seite 690 von 1250
Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Patienten erhielten über ein IVRS eine Identifikationsnummer und wurden im Verhältnis 2:2:1 zu Acclidiniumbromid, Tiotropiumbromid oder Placebo randomisiert.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Therapiezuweisung erfolgte durch ein interaktives Voice-Response-System (IVRS) und war gegenüber Prüfärzten und Patienten verblindet. Die Inhalatoren glichen sich in Bezug auf äußere Erscheinung und Zusammensetzung.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet. Alle Inhalatoren glichen sich in Bezug auf äußere Erscheinung und Zusammensetzung.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Behandler nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Falls es Sicherheit und/oder Wohlbefinden des Patienten erforderten, war der Prüfarzt in Notfallsituationen berechtigt, die Verblindung über das IVRS-System aufzuheben. Im Falle einer Entblindung wurde das Forest Research Institute umgehend informiert und es wurde ein schweres unerwünschtes Ereignis dokumentiert. Nach Offenlegung des Codes am Prüfzentrum wurde der Patient sofort von der Studie ausgeschlossen.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Ableich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Ableich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die relevanten Endpunkte wurden vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Transparenter Patientenfluss. Keine Interimanalyse geplant oder durchgeführt.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr

unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat

berichtet

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: FEV 1 AUC

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfärzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Behandlungszufriedenheit

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfärzte wurden im Hinblick auf das zugewiesene Medikament

Aclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation/704 von 1250 Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)

(Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende

Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse (Adverse events, AEs)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfärzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwere unerwünschte Ereignisse (Serious Adverse Events, SAE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfärzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende

Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Nächtliche und morgendliche Symptome

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und

Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: COPD-Symptome

1. Verblindung der Endpunkterheber

*Aclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation) sowie
Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)*

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der

Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Kardiovaskuläre Morbidität und Mortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfärzte wurden im Hinblick auf das zugewiesene Medikament

Aclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation/720 von 1250 Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)

(Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende

Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid, Tiotropiumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder –wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung

angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Korrekte Umsetzung des ITT-Prinzips.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-32 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie ACCORD I

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Kerwin E (2011): Clinical Trial M/34273/33 - Efficacy and Safety of Acridinium Bromide at Two Dose Levels (200 µg Twice Daily, 400 µg Twice Daily) vs. Placebo When Administered to Patients With Moderate to Severe Chronic Obstructive Pulmonary Disease (COPD). Stand: 21/01/2011. Confidential.	ACCORD I
Kerwin EM, D'Urzo AD, Gelb AF, Lakkis H, Garcia GE, Caracta CF, et al. (2012): Efficacy and safety of a 12-week treatment with twice-daily acridinium bromide in COPD patients (ACCORD COPD I). <i>Copd: Journal of Chronic Obstructive Pulmonary Disease</i> ; 9(2):90-101.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Die Zuteilung zu den Behandlungsgruppen erfolgte randomisiert.

1.

für randomisierte Studien:**Adäquate Erzeugung der Randomisierungssequenz**

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Patienten erhielten über ein IVRS (Interactive-Voice-Response-System von Perceptive Informatics, Inc) eine Identifikationsnummer und wurden im Verhältnis 1:1:1 zu 200 µg Acridiniumbromid BID, 400 µg Acridiniumbromid BID oder Placebo BID randomisiert.

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)

- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Therapiezuweisung erfolgte durch ein interaktives Voice-Response-System (IVRS) und war gegenüber Prüfärzten und Patienten verblindet. Die Inhalatoren glichen sich in Bezug auf äußere Erscheinung und Zusammensetzung.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

- ja:** Die Patienten waren verblindet.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Alle Inhalatoren glichen sich in Bezug auf äußere Erscheinung und Zusammensetzung.

Behandler bzw. Weiterbehandler

- ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Falls es Sicherheit und/oder Wohlbefinden des Patienten erforderten, war der Prüfarzt in Notfallsituationen berechtigt, die Verblindung über das IVRS-System aufzuheben. Im Falle einer Entblindung wurde das Forest Research Institute umgehend informiert und es wurde ein schweres unerwünschtes Ereignis dokumentiert. Nach Offenlegung des Codes am Prüfzentrum wurde der Patient sofort von der Studie ausgeschlossen.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten)

einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die relevanten Endpunkte wurden vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Transparenter Patientenfluss. Bei Patienten, die die Studie aufgrund „unzureichenden therapeutischen Ansprechens“ abbrechen, könnte dieses schlechte Ansprechen auf die mangelnde Wirksamkeit der Prüfpräparate zurückzuführen sein (Seite 249, ACCORD I, mehr Abbrecher mit Placebo oder 200 µg Acridiniumbromid als mit 400 µg Acridiniumbromid).

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Anzahl der Patienten, die die Studie in drei Armen abbrechen: Placebo 19,9 %, Acclidinium 200 µg 17,8 % und Acclidinium 400 µg 12,6 %. Einer der Gründe für den häufigeren Abbruch in Placebo- oder Acclidinium-200 µg-Gruppe war das „unzureichende therapeutische Ansprechen“ (u. U. aufgrund einer mangelnden Wirksamkeit des Prüfpräparats, Seite 249, ACCORD I). Ein Ausschluss dieser Patienten begünstigt die FEV₁-Ergebnisse für Acclidinium 400 µg jedoch nicht. Die analysierte Population umfasste Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan festgelegt und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament

(Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

- nein:** Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Patienten, die die Studie in den drei Armen abbrachen: Placebo 19,9 %, Acclidinium 200 µg 17,8 % und Acclidinium 400 µg 12,6 %. Einer der Gründe für den häufigeren Abbruch in Placebo- und 200 µg-Aclidiniumbromid-Gruppe war das „unzureichende therapeutische Ansprechen“ (u. U. aufgrund einer mangelnden Wirksamkeit des Prüfpräparats, Seite 249, ACCORD I). Die analysierte Population umfasste Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan festgelegt und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: SGRQ**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population umfasste entsprechend des ITT-Prinzips alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: SGRQ Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte waren im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu

beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population umfasste entsprechend des ITT-Prinzips alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: TDI

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu

anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population umfasste alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: TDI Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population umfasste alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung

angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population umfasste alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

Der Therapieeffekt im Hinblick auf Exazerbationen wurde um den Einfluss der Kovariate Alter bereinigt, was für das randomisierte Studiendesign nicht erforderlich ist. Das Ergebnis wurde um den Einfluss von Kovariaten bereinigt, was nach der Randomisierung nicht erforderlich sein sollte.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Krankenhausaufenthalte aufgrund von COPD-Exacerbationen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr

genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Zahl der Krankenhausaufenthalte aufgrund von COPD-Exacerbationen wird sowohl in der Wirksamkeits- als auch in der Sicherheitsanalyse (als schwere unerwünschte Ereignisse) angegeben. In der Wirksamkeitsanalyse wurde bei 4 Patienten eine COPD, die zu einer Hospitalisierung führte, dokumentiert (je einmal in Placebo und Acridiniumbromid 200 µg-Gruppe und zweimal in der Acridiniumbromid 400 µg-Gruppe). In der Sicherheitsanalyse wurden 5 Patienten mit dem schwerem unerwünschten Ereignis „COPD-Exacerbation, die zur Krankenhauseinweisung führte“ angegeben (je einmal in Placebo und Acridiniumbromid 200 µg-Gruppe und dreimal in der Acridiniumbromid 400 µg-Gruppe).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es ist unklar, warum hinsichtlich der Analyse der COPD-bedingten Krankenhausaufenthalte unterschiedliche Zahlen für die 400 µg Acclidiniumbromid-Gruppe berichtet wurden.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt eine Inkongruenz zwischen Wirksamkeits- und Sicherheitsanalyse, die sich nicht durch die unterschiedliche Zusammensetzung von Sicherheits- und ITT-Population erklären lässt. Die beiden Populationen unterscheiden sich lediglich durch einen zusätzlichen Patienten in der Placebo-Gruppe:

- Studienbericht Seite 3:
Sicherheitspopulation (N): Placebo 186; Acclidinium 200 µg 184; Acclidinium 400 µg 190
ITT-Population (N): Placebo 185; Acclidinium 200 µg 184; Acclidinium 400 µg 190

Bei Wirksamkeits- und Sicherheitsanalyse gibt es jedoch eine Differenz in der Acclidiniumbromid 400 µg-Gruppe:

- Studienbericht Seite 166 (Wirksamkeitsanalyse):
„Krankenhausaufenthalte aufgrund von COPD Exazerbationen
Von allen ITT-Patienten wurden 4 Patienten stationär aufgenommen, 1 (0,5 %) in der Placebo-Gruppe, 1 (0,5 %) in der 200-µg-Aclidiniumbromid-Gruppe und 2 (1,1 %) in der 400-µg-Aclidiniumbromid-Gruppe.“
- Studienbericht Seite 198 (Sicherheitsanalyse):
„Das am häufigsten angegebene SUE war eine Exazerbation der COPD (die einen Krankenhausaufenthalt erforderte). Bei 5 Patienten traten schwere COPD Exazerbationen auf: 1 Patient in der Placebo-Gruppe, 1 Patient in der 200-µg-Aclidiniumbromid-Gruppe und 3 Patienten in der 400-µg-Aclidiniumbromid-Gruppe.“

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Der Grund für den Unterschied bei den in der 400-µg-Aclidiniumbromid-Gruppe stationär aufgenommenen Patienten, der in der Sicherheitspopulation verglichen mit der ITT-Population auftrat, ist unklar.

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitsanalysen basierten auf der Sicherheitspopulation (alle randomisierten Patienten, die mindestens eine Dosis der doppelblinden Medikation einnahmen).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab definiert, verblindet erhoben und adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SAE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüferärzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitspopulation umfasste alle Patienten, die wenigsten eine Dosis der Studienmedikation erhielten. Das ITT-Prinzip wurde adäquat umgesetzt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab definiert, verblindet erhoben und adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei

einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde streng nach dem ITT-Prinzip in der Sicherheitspopulation erhoben

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab definiert, verblindet erhoben und adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Nächtliche Symptome

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Der Endpunkt wurde anhand der Notwendigkeit des Gebrauchs der Bedarfs-Medikation erhoben.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population entsprach der ITT-Population (alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab definiert, verblindet erhoben und adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Kardiovaskuläre Morbidität und Mortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüferärzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Safetyanalysen basierten auf der Safetypopulation (alle Patienten der randomisierten Population eingeschlossen, die mindestens eine Dosis der doppelblinden Medikation einnahmen).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende

Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung

angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Safetyanalysen basierten auf der Safetypopulation (alle Patienten der randomisierten Population eingeschlossen, die mindestens eine Dosis der doppelblinden Medikation einnahmen).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Während der Dauer der Studie wurde nur ein Todesfall (aufgrund von Lungenkrebs, nicht therapiebedingt) verzeichnet. Der Patientenbericht ist im Studienbericht enthalten.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-33 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie ACCORD II

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Taglietti M (2011): Clinical Trial M/34273/38 Part A - A Randomized, Double-blind, Placebo-Controlled Study Evaluating the Efficacy, Safety, and Tolerability of 2 Doses of Acclidinium Bromide Compared With Placebo for 12 Weeks in Patients With Moderate to Severe, Stable Chronic Obstructive Pulmonary Disease Followed by a 40-Week Evaluation of the Higher Acclidinium Bromide Dose. Stand: 16/03/2011. Confidential.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Die Zuteilung zu den Behandlungsgruppen erfolgte randomisiert.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

Acclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation) Seite 7/72 von 1250
Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Patienten erhielten eine Identifikationsnummer über das IVRS (Interactive-Voice-Response-System der Premier Research Group). Eine Liste der Patienten-Randomisierungs-codes wurde von der Abteilung für statistische Programmierung im Forest Research Institute Inc. generiert. Die Patienten wurden in einem Randomisierungsverhältnis von 1:1:1 zu 200 µg Acridiniumbromid BID, 400 µg Acridiniumbromid BID oder Placebo BID randomisiert.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern

- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Zuteilung zu den Behandlungsgruppen erfolgte durch ein interaktives Voice Response-System (IVRS), das den Patienten eine Identifikationsnummer zuteilte. Alle Inhalatoren waren äußerlich identisch und hatten die gleiche Zusammensetzung. Die Placebo-Inhalatoren enthielten allerdings keinen Wirkstoff. Die aktive Substanz hatte keinen wahrnehmbaren Geschmack, Erscheinung, Geruch oder Farbe, so dass die Verblindung nicht durchbrochen werden konnte.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

- ja:** Die Patienten waren verblindet.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Alle Inhalatoren glichen sich in Bezug auf äußere Erscheinung und Zusammensetzung. Die aktive Substanz hatte keinen wahrnehmbaren Geschmack, Erscheinung, Geruch oder Farbe. Die Verblindung wurde nur in medizinischen Notfällen aufgehoben. Nachdem der Code am Prüfzentrum offengelegt wurde, wurde der Patient sofort von der Studie ausgeschlossen.

Behandler bzw. Weiterbehandler

- ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Für die Sicherheit und das Wohlbefinden des Patienten war der Prüfarzt in Notfallsituationen berechtigt, die Verblindung über das IVRS-System aufzuheben. Im Falle einer Entblindung wurde das Forest Research Institute umgehend informiert und es wurde eine umfassende schriftliche Erklärung vorgelegt. Nachdem der Code am Prüfzentrum offengelegt wurde, wurde der Patient sofort von der Studie ausgeschlossen.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h.

vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.
- Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.
- Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.
- Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- Subgruppen
- Zeitpunkte / -räume
- Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)
- Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)
- Cut-off-points bei Dichotomisierung
- statistischer Verfahren

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).
- Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.

Zulässige Gründe sind:

- erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung
- Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe
- geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben
- Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.
- Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive

Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen /Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Zwischen den einzelnen Behandlungsgruppen traten im Hinblick auf die FEV₁ zu Studienbeginn signifikante Unterschiede auf (1.249 l in der 400-µg-Aclidiniumbromid-Gruppe, 1.397 l in der 200-µg-Aclidiniumbromid-Gruppe und 1.459 l in der Placebo-Gruppe), genauso wie im Hinblick auf den Anteil der Patienten mit schwerer COPD Stadium III (55 % in der 400-µg-Aclidiniumbromid-Gruppe, 48 % in der 200-µg-Aclidiniumbromid-Gruppe und 37 % in der Placebo-Gruppe für Stadium III).

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Studiendesign und die Implementierung dürften keine Verzerrungen zulassen. Dieses Ungleichgewicht bei der zugelassenen Population könnte sich jedoch möglicherweise auf die Studienergebnisse auswirken.

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der

Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Das ITT-Prinzip wurde adäquat umgesetzt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab definiert und adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Zwischen den einzelnen Behandlungsgruppen traten im Hinblick auf die FEV₁ zu Studienbeginn signifikante Unterschiede auf (1.249 l in der 400-µg-Aclidiniumbromid-Gruppe, 1.397 l in der 200-µg-Aclidiniumbromid-Gruppe und 1.459 l in der Placebo-Gruppe), genauso wie im Hinblick auf den Anteil der Patienten mit schwerer COPD Stadium III (55 % in der 400-µg-Aclidiniumbromid-Gruppe, 48 % in der 200-µg-Aclidiniumbromid-Gruppe und 37 % in der Placebo-Gruppe für Stadium III). Dies könnte sich auf die Ergebnisse für den primären Endpunkt (Veränderung gegenüber der Baseline beim Trough-FEV₁ in Woche 12) auswirken.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Ungleichgewicht zwischen den Behandlungsgruppen bei Baseline (wie in A5 beschrieben) hat die Ergebnisse im Hinblick auf den primären Endpunkt unter Umständen beeinflusst.

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfer wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population (ITT) umfasste Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Zwischen den einzelnen Behandlungsgruppen traten im Hinblick auf die FEV₁ signifikante Unterschiede auf (1.249 l in der 400-µg-Aclidiniumbromid-Gruppe, 1.397 l in der 200-µg-Aclidiniumbromid-Gruppe und 1.459 l in der Placebo-Gruppe), genauso wie im Hinblick auf den Anteil der Patienten mit schwerer COPD Stadium III (55 % in der 400-µg-Aclidiniumbromid-Gruppe, 48 % in der 200-µg-Aclidiniumbromid-Gruppe und 37 % in der Placebo-Gruppe für Stadium III). Dadurch könnten die Ergebnisse im Hinblick auf die Veränderung des Peak-FEV₁ zur Baseline in Woche 12 beeinflusst werden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Behandlungsgruppen konnten zur Baseline nicht verglichen werden. Dadurch könnten die Ergebnisse im Hinblick auf die Veränderungen des Peak-FEV₁ zur Baseline beeinflusst worden sein.

Endpunkt: SGRQ

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population (ITT) umfasste alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde im Prüfplan vorab definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Zwischen den einzelnen Behandlungsgruppen traten im Hinblick auf die FEV₁ signifikante Unterschiede auf (1.249 l in der 400-µg-Aclidiniumbromid-Gruppe, 1.397 l in der 200-µg-Aclidiniumbromid-Gruppe und 1.459 l in der Placebo-Gruppe), genauso wie im Hinblick auf den Anteil der Patienten mit schwerer COPD Stadium III (55 % in der 400-µg-Aclidiniumbromid-Gruppe, 48 % in der 200-µg-Aclidiniumbromid-Gruppe und 37 % in der Placebo-Gruppe für Stadium III). Die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline können sich u. U. auf die Ergebnisse des SGRQ-Endpunkts auswirken.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Ungleichgewicht zwischen den Behandlungsgruppen zur Baseline (wie in B.3 beschrieben) hat die Ergebnisse im Hinblick auf den SGRQ-Endpunkt unter Umständen beeinflusst.

Endpunkt: SGRQ Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population (ITT) umfasste alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde im Prüfplan vorab definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Zwischen den einzelnen Behandlungsgruppen traten im Hinblick auf die FEV₁ signifikante Unterschiede auf (1.249 l in der 400-µg-Aclidiniumbromid-Gruppe, 1.397 l in der 200-µg-Aclidiniumbromid-Gruppe und 1.459 l in der Placebo-Gruppe), genauso wie im Hinblick auf den Anteil der Patienten mit schwerer COPD

Stadium III (55 % in der 400-µg-Aclidiniumbromid-Gruppe, 48 % in der 200-µg-Aclidiniumbromid-Gruppe und 37 % in der Placebo-Gruppe für Stadium III). Die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline können sich u. U. auf die Ergebnisse des SGRQ-Endpunkts auswirken.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Ungleichgewicht zwischen den Behandlungsgruppen zur Baseline (wie in B.3 beschrieben) hat die Ergebnisse im Hinblick auf den SGRQ-Endpunkt unter Umständen beeinflusst.

Endpunkt: TDI

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

Die analysierte Population (ITT) umfasste alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Zwischen den einzelnen Behandlungsgruppen traten im Hinblick auf die FEV₁ signifikante Unterschiede auf (1.249 l in der 400-µg-Aclidiniumbromid-Gruppe, 1.397 l in der 200-µg-Aclidiniumbromid-Gruppe und 1.459 l in der Placebo-Gruppe), genauso wie im Hinblick auf den Anteil der Patienten mit schwerer COPD Stadium III (55 % in der 400-µg-Aclidiniumbromid-Gruppe, 48 % in der 200-µg-Aclidiniumbromid-Gruppe und 37 % in der Placebo-Gruppe für Stadium III). Die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline können sich u. U. auf die Ergebnisse der TDI-Bewertungen auswirken.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Ungleichgewicht zwischen den Behandlungsgruppen zur Baseline (wie in B.3 beschrieben) hat die Ergebnisse im Hinblick auf den TDI-Endpunkt unter Umständen beeinflusst.

Endpunkt: TDI Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüferärzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die analysierte Population (ITT) umfasste alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und im Studienbericht adäquat berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Zwischen den einzelnen Behandlungsgruppen traten im Hinblick auf die FEV₁ signifikante Unterschiede auf (1.249 l in der 400-µg-Aclidiniumbromid-Gruppe, 1.397 l in der 200-µg-Aclidiniumbromid-Gruppe und 1.459 l in der Placebo-Gruppe), genauso wie im Hinblick auf den Anteil der Patienten mit schwerer COPD Stadium III (55 % in der 400-µg-Aclidiniumbromid-Gruppe, 48 % in der 200-µg-Aclidiniumbromid-Gruppe und 37 % in der Placebo-Gruppe für Stadium III). Die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline können sich u. U. auf die Ergebnisse der TDI-Bewertungen auswirken.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Ungleichgewicht zwischen den Behandlungsgruppen zur Baseline (wie in B.3 beschrieben) hat die Ergebnisse im Hinblick auf den TDI-Endpunkt unter Umständen beeinflusst.

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüferärzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die in Bezug auf Exazerbationen analysierten Populationen waren die ITT-Population (beschrieben als Patienten, die mindestens eine Dosis des Prüfpräparats erhielten und deren FEV₁ Post-Baseline mindestens einmal bewertet wurde) und die Sicherheitspopulation (Patienten, die mindestens eine Dosis des Prüfpräparats erhielten).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde im Prüfplan vorab definiert und im Studienbericht adäquat beschrieben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Die Behandlungsgruppen unterschieden sich zu Baseline erheblich (Einzelheiten s. A5).

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Behandlungsgruppen unterschieden sich zu Baseline erheblich (Einzelheiten s. A5).

Endpunkt: Krankenhausaufenthalte aufgrund von COPD-Exacerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Aclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation) Seite 7/99 von 1250
Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die in Bezug auf Exazerbationen analysierten Populationen waren die ITT-Population (beschrieben als Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen und deren FEV₁ Post-Baseline mindestens einmal bewertet wurde) und die Sicherheitspopulation (beschrieben als Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Die Behandlungsgruppen konnten zur Baseline in Bezug auf den Schweregrad der COPD (s. A5) nicht verglichen werden. Dies könnte die Anzahl der Krankenhausaufenthalte während der verbleibenden Dauer der Studie beeinflusst haben.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende

Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Behandlungsgruppen konnten zur Baseline in Bezug auf den Schweregrad der COPD (s. A5) nicht verglichen werden. Dies könnte die Anzahl der Krankenhausaufenthalte während der verbleibenden Dauer der Studie beeinflusst haben.

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu

definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitsanalyse (unerwünschte Ereignisse einschließlich COPD Exazerbationen) wurde in der Sicherheitspopulation (alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen) durchgeführt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Starke Konzentration auf therapiebedingte Nebenwirkungen (TEAEs). Insgesamt wurden unerwünschte Ereignisse nur im Haupttext für UE angegeben, wenn sie bei mindestens 2 % der Patienten auftraten. Für detailliertere Informationen müssten zusätzliche Berechnungen auf Basis der Rohdaten in den Anhängen durchgeführt werden. Therapiebedingte Nebenwirkungen werden im Haupttext sehr detailliert beschrieben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Zwischen den einzelnen Behandlungsgruppen traten zur Baseline im Hinblick auf das FEV₁ signifikante Unterschiede auf, genauso wie im Hinblick auf den Anteil der Patienten mit schwerer COPD Stadium III (55 % in der 400-µg-Aclidiniumbromid-Gruppe, 48 % in der 200-µg-Aclidiniumbromid-Gruppe und 37 % in der Placebo-Gruppe für Stadium III). Die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline könnten sich auf die Anzahl der angegebenen unerwünschten Ereignisse auswirken.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Behandlungsgruppen konnten zur Baseline nicht verglichen werden. Dies könnte Auswirkungen auf die Berichterstattung haben. Bei der Erfassung der UE's fand keine selektive Berichterstattung statt.

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitsanalyse (unerwünschte Ereignisse einschließlich COPD Exazerbationen) wurde in der Sicherheitspopulation (alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen) durchgeführt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Die Behandlungsgruppen konnten zur Baseline nicht verglichen werden und die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline könnten sich auf die Anzahl der angegebenen unerwünschten und schwerwiegenden unerwünschten Ereignisse auswirken.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Behandlungsgruppen waren nicht vergleichbar zur Baseline.

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Acidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr

genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitsanalyse (unerwünschte Ereignisse einschließlich COPD Exazerbationen) wurde in der Sicherheitspopulation (alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen) durchgeführt.
Die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline könnten sich auf die Anzahl der Studienabbrüche aufgrund unerwünschter Ereignisse in den einzelnen Therapiearmen ausgewirkt haben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Die Behandlungsgruppen konnten zur Baseline nicht verglichen werden und die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline könnten sich auf die Anzahl der Studienabbrüche aufgrund unerwünschter Ereignisse in den einzelnen Therapiearmen auswirken.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Behandlungsgruppen waren nicht vergleichbar zur Baseline.

Endpunkt: Kardiovaskuläre Morbidität und Mortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfer wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitsanalyse (unerwünschte Ereignisse einschließlich COPD Exazerbationen) wurde in der Sicherheitspopulation (alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen) durchgeführt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Die Behandlungsgruppen konnten zur Baseline nicht verglichen werden und die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline könnten die Anzahl der Patienten beeinflussen, bei denen kardiovaskuläre Morbiditäten auftraten oder die in Folge von kardiovaskulären Ereignissen verstarben. Im Durchschnitt enthielt die Placebo-Gruppe mehr Patienten mit einer weniger schweren COPD. Aus diesem Grund können die Ergebnisse für diesen kombinierten Endpunkt nicht unbedingt zugunsten von Acridiniumbromid gedeutet werden. Es ist jedoch kein unverzerrter Vergleich zwischen Placebo und Acridiniumbromid möglich.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Behandlungsgruppen waren nicht vergleichbar zur Baseline.

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien,

statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Safetyanalyse (unerwünschte Ereignisse einschließlich COPD Exazerbationen) wurde in der Safetypopulation (alle Patienten, die mindestens eine Dosis des Prüfpräparats einnahmen) durchgeführt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Die Behandlungsgruppen konnten zur Baseline nicht verglichen werden und die Unterschiede beim COPD-Schweregrad zwischen den Therapiegruppen zur Baseline

könnten sich auf die Anzahl der während der Studie verstorbenen Patienten auswirken. Im Durchschnitt enthielt die Placebo-Gruppe mehr Patienten mit weniger schwerer COPD. Aus diesem Grund können die Ergebnisse für diesen Endpunkt nicht unbedingt zugunsten von Acclidiniumbromid gedeutet werden. Es ist jedoch kein unverzerrter Vergleich zwischen Placebo und Acclidiniumbromid möglich.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Behandlungsgruppen waren nicht vergleichbar zur Baseline.

Tabelle 4.7-34 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie ATTAIN

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Jones PW (2011): Clinical Trial M/34273/34 - Efficacy and safety of acclidinium bromide at two dose levels vs placebo when administered to patients with moderate to severe chronic obstructive pulmonary disease (COPD). Stand: 12/04/2011. Confidential.	ATTAIN
Jones PW, Singh D, Bateman ED, Agusti A, Lamarca R, de Miquel G, et al. (2012): Efficacy and safety of twice-daily acclidinium bromide in COPD patients: The ATTAIN study. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology;	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Die Therapie wurde mittels Randomisierung zugewiesen.

1. für randomisierte Studien:

Aclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation) Seite 817 von 1250
Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Von einer Arbeitsgruppe Statistik- / Programmierung bei Almirall S.A. wurde gemäß SOP ein computergenerierter Randomisierungsplan erstellt, um einer Behandlungsgruppe eine Randomisierungsnummer zuzuweisen. Die Blockgröße war den Prüfarzten nicht bekannt. In dem Moment, in dem der Patient die Einwilligungserklärung (ICF) unterzeichnete, wiesen die Studienmitarbeiter jedem Patient eine Patientenidentifikationsnummer zu, indem Sie das IVRS (Interactive-Voice Response System von PAREXEL International) anriefen. Die Patienten wurden mithilfe des IVRS randomisiert.

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)

- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Art der Behandlung wurde gegenüber Prüfärzten und Patienten verblindet.
Alle Inhalatoren waren äußerlich identisch.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

- ja:** Die Patienten waren verblindet.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Alle Inhalatoren glichen sich im Hinblick auf ihre äußere Erscheinung.

Behandler bzw. Weiterbehandler

- ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet. Die Prüfarzte erhielten keine versiegelten Notfall-Umschläge, in denen die Behandlungsgruppen identifiziert werden konnten. Wenn es zur Wahrung der Sicherheit und des Wohlergehens des Patienten notwendig war, konnte der Prüfarzt nach eigenem Ermessen Informationen zum tatsächlich verabreichten Studienmedikament erhalten. Dazu musste er das IVRS anrufen und den entsprechenden Patienten nennen, der daraufhin entblindet wurde. Der Prüfarzt erhielt dann ein Bestätigungsformular, auf dem die jeweilige Therapie aufgeführt wurde. Almirall und die klinische Forschungsorganisation wurden umgehend informiert, wenn ein Patient entblindet wurde.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten)

einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die relevanten Endpunkte wurden im Prüfplan prospektiv definiert und im klinischen Studienbericht entsprechend berichtet.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament

(Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

- nein:** Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts erfolgte streng nach dem ITT-Prinzip. Die Analyse aller Wirksamkeitsvariablen wurde in der ITT-Population durchgeführt. Diese umfasste alle randomisierten Patienten, die mindestens eine Dosis der doppelblinden Prüfmedikation einnahmen und deren FEV₁ zur Baseline und mindestens einmal Post-Baseline bewertet wurde. Der Patientenfluss wurde eindeutig beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und in Studienbericht adäquat berichtet.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse aller Wirksamkeitsvariablen wurde in der ITT-Population durchgeführt. Diese umfasste alle randomisierten Patienten, die mindestens eine Dosis der doppelblinden Prüfmedikation einnahmen und deren FEV₁ zur Baseline und mindestens einmal Post-Baseline bewertet wurde. Der Patientenfluss wurde eindeutig beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und adäquat im Studienbericht berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: SGRQ

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung

angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die SGRQ-Population umfasste in der Placebo-Gruppe 2 Patienten weniger als in der ITT-Population. Der Grund für diese Diskrepanz ist unklar. Eine relevante Verzerrung ist unwahrscheinlich. Die 400 µg-Aclidiniumbromid-Population entsprach der ITT-Population.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Obwohl unklar ist, warum die SGRQ-Beurteilung nicht für die gesamte ITT-Population durchgeführt wurde, ist nicht anzunehmen, dass dies zu Verzerrungen im Hinblick auf diesen Endpunkt führen wird. Die Anzahl der nicht analysierten Patienten ist äußerst gering und würde die Ergebnisse des SGRQ-Endpunkts logischerweise nicht wesentlich beeinflussen.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: SGRQ Responder**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die SGRQ-Population umfasste in der Placebo-Gruppe 2 Patienten weniger als in der ITT-Population. Der Grund für diese Diskrepanz ist unklar. Eine relevante Verzerrung ist unwahrscheinlich. Die 400 µg-Aclidiniumbromid-Population entsprach der ITT-Population.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Obwohl unklar ist, warum die SGRQ-Beurteilung nicht für die gesamte ITT-Population durchgeführt wurde, ist nicht anzunehmen, dass dies zu Verzerrungen im Hinblick auf diesen Endpunkt führen wird. Die Anzahl der nicht analysierten Patienten ist äußerst gering und würde die Ergebnisse des SGRQ-Endpunkts logischerweise nicht wesentlich beeinflussen.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: TDI

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse erfolgte in der ITT-Population nach dem LOCF-Prinzip. Folgende Patienten brachen die Studie in den verschiedenen Armen ab: Placebo 14,9 %, Aclidinium 200 µg 8,6 % und Aclidinium 400 µg 6,3 %. Ein Grund für den Abbruch war die „mangelnde Wirksamkeit“. Dies könnte sich auf die Ergebnisse in Bezug auf den TDI-Endpunkt auswirken. Es ist nicht klar, bei wie vielen Patienten letztendlich der TDI ermittelt wurde. Die Anzahl der im Hinblick auf den BDI bewerteten Patienten war geringer als die Anzahl der Patienten in der Sicherheitspopulation. Der BDI wird als Ausgangswert für die Bestimmung des TDI verwendet. In allen Armen fehlten Patienten (verglichen mit der Sicherheitspopulation): Acht Patienten in der Placebo-Gruppe, 11 Patienten in der Aclidiniumbromid 200 µg-Gruppe und 3 Patienten in der Aclidiniumbromid 400 µg-Gruppe. Es wurde nicht eindeutig vermittelt, warum diese Patienten in der Analyse nicht erfasst sind.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Obwohl unklar ist, warum die BDI-Beurteilung nicht für die gesamte ITT-Population durchgeführt wurde und obwohl die Populationsgröße für den TDI nicht eindeutig ist, ist nicht anzunehmen, dass dies zu Verzerrungen im Hinblick auf diesen Endpunkt führen wird. Die Anzahl der nicht analysierten Patienten ist relativ klein und würde die Ergebnisse des TDI-Endpunkts logischerweise nicht wesentlich

beeinflussen.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: TDI Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse erfolgte in der ITT-Population nach dem LOCF-Prinzip. Folgende Patienten brachen die Studie in den verschiedenen Armen ab: Placebo 14,9 %, Aclidinium 200 µg 8,6 % und Aclidinium 400 µg 6,3 %. Ein Grund für den Abbruch war die „mangelnde Wirksamkeit“. Dies könnte sich auf die Ergebnisse in Bezug auf den TDI-Endpunkt auswirken. Es ist nicht klar, bei wie vielen Patienten letztendlich der TDI ermittelt wurde. Die Anzahl der im Hinblick auf den BDI bewerteten Patienten war geringer als die Anzahl der Patienten in der Sicherheitspopulation. Der BDI wird als Ausgangswert für die Bestimmung des TDI verwendet. In allen Armen fehlten Patienten (verglichen mit der Sicherheitspopulation): Acht Patienten in der Placebo-Gruppe, 11 Patienten in der Aclidiniumbromid 200 µg-Gruppe und 3 Patienten in der Aclidiniumbromid 400 µg-Gruppe. Es wurde nicht eindeutig vermittelt, warum diese Patienten in der Analyse nicht erfasst sind.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Obwohl unklar ist, warum die BDI-Beurteilung nicht für die gesamte ITT-Population durchgeführt wurde und obwohl die Populationsgröße für den TDI nicht eindeutig ist, ist nicht anzunehmen, dass dies zu Verzerrungen im Hinblick auf diesen Endpunkt führen wird. Die Anzahl der nicht analysierten Patienten ist relativ klein und würde die Ergebnisse des TDI-Endpunkts logischerweise nicht wesentlich

beeinflussen.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Endpunkt wurde vorab im Prüfplan definiert und entsprechend berichtet. Der Therapieeffekt im Hinblick auf Exazerbationen wurde um den Einfluss der Kovariate Alter bereinigt, was für das randomisierte Studiendesign nicht erforderlich ist. Das Ergebnis wurde um den Einfluss von Kovariaten bereinigt, was nach der Randomisierung nicht erforderlich sein sollte.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Krankenhausaufenthalte

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament

(Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse aller Wirksamkeitsvariablen wurde in der ITT-Population durchgeführt. Diese umfasste alle randomisierten Patienten, die mindestens eine Dosis der doppelblinden Prüfmedikation einnahmen und deren FEV₁ zur Baseline und mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitsergebnisse wurden anhand der Sicherheitspopulation (alle Patienten der randomisierten Population, die mindestens eine Dosis der doppelblinden Medikation einnahmen) analysiert. Der Patientenfluss ist eindeutig.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitsergebnisse wurden anhand der Sicherheitspopulation (alle Patienten der randomisierten Population, die mindestens eine Dosis der doppelblinden Medikation einnahmen) analysiert.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Insgesamt wurden vier Patienten (3 in der Placebogruppe und 1 in der 400-µg-Aclidinium-Gruppe) in den Bewertungen nicht berücksichtigt (Abbruch aufgrund „anderer“ Gründe angegeben); Grund waren fehlende Aufzeichnungen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Nächtliche und frühmorgendliche Symptome**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte

Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse aller Wirksamkeitsvariablen wurde in der ITT-Population durchgeführt. Diese umfasste alle randomisierten Patienten, die mindestens eine Dosis der doppelblinden Prüfmedikation einnahmen und deren FEV₁ zur Baseline und

mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Im Artikel von Jones („A“) fehlen Angaben zu nächtlichen und frühmorgendlichen Symptomen an den Stellen, wo sie im klinischen Studienbericht („ATTAIN“) aufgeführt werden.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: EXACT-PRO-Fragebogen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-

Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse aller Wirksamkeitsvariablen wurde in der ITT-Population durchgeführt. Diese umfasste alle randomisierten Patienten, die mindestens eine Dosis der doppelblinden Prüfmedikation einnahmen und deren FEV₁ zur Baseline und mindestens einmal Post-Baseline bewertet wurde.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Artikel von Jones („A“) enthält keine Angaben zu den EXACT-PRO-Punktwerten, die im klinischen Studienbericht („ATTAIN“) aufgeführt und beschrieben werden. „ATTAIN“ macht Angaben zu sowohl den Ergebnissen im Hinblick auf die deskriptive Statistik als auch zu den Ergebnissen aus dem statistischen Modell.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Kardiovaskuläre Morbidität und Mortalität**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Sicherheitsergebnisse wurden anhand der Sicherheitspopulation (alle Patienten der randomisierten Population, die mindestens eine Dosis der doppelblinden Medikation einnahmen) analysiert.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten und Prüfarzte wurden im Hinblick auf das zugewiesene Medikament (Aclidiniumbromid oder Placebo) verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Ausschlusskriterien für die Aufnahme verhinderten Verzerrungen, da die Aufnahme von Patienten mit bestehenden Komorbiditäten beschränkt war.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Todesfälle werden für alle Therapiearme objektiv angegeben und umfassen ausführliche Informationen über die Todesursache und die Verbindung mit der Studienmedikation.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-35 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Bateman 2010a

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Bateman E, Singh D, Smith D, Disse B, Towse L, Massey D, et al. (2010): Efficacy and safety of tiotropium Respimat SMI in COPD in two 1-year randomized studies. International Journal of COPD; 5:197-208.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurden zwei identische einjährige, multizentrische, randomisierte, doppelblinde Parallelgruppenstudien analysiert. (Seite 198).

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Randomisierungsmethode wurde nicht beschrieben.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Nicht angegeben.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde nur angegeben, dass die Studie doppelt verblindet war. Wir gehen daher davon aus, dass Patient und Arzt verblindet wurden.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde nur angegeben, dass die Studie doppelt verblindet war. Wir gehen daher davon aus, dass Patient und Arzt verblindet wurden.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Therapieeffekte im Hinblick auf das FEV₁ wurden anhand der Behandlung mit dem Baseline-ICS stratifiziert. In Tabllle 3 wurden alle Daten ohne die Stratifizierung dargestellt.

Therapieeffekte im Hinblick auf FEV₁, TDI und SGRQ sowie sekundäre Ergebnisse wurden hinsichtlich des Einflusses von Kovariaten angepasst, Effekte im Hinblick

auf Exazerbationen wurden jedoch nicht adjustiert. (Seite 199)

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Der Patientenfluss war unklar. Patienten wurden ausgeschlossen, weil Daten nicht ausgewertet werden konnten (Seite 200). Der Hauptgrund für den Ausschluss aus der Analyse war, dass die Patienten weniger als 5 Tage die Studienbehandlung erhalten hatten. Ob dieses Ausschlusskriterium vorab festgelegt wurde, bleibt unklar.

Des Weiteren wurden zwei Studien in einer Analyse kombiniert. Die Randomisierung gilt nur für eine Studie. Es ist nicht klar, ob die Adjustierung hierfür erfolgt ist (z. B. Betrachtung der Therapieeffekte pro Studie).

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die oben genannte Bewertung der Risikoverzerrung (Punkte 1 – 5) weist darauf hin, dass die Berichterstattung in dieser Publikation nicht ausreichend war. Wir sahen uns daher außerstande, die Qualität der Studie zu bescheinigen. Darüber hinaus wurden Patienten ausgeschlossen, weil Daten nicht ausgewertet werden konnten. Die Bereinigung um Therapieeffekte ist unter Umständen nicht angemessen.

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr

unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Dieses Dokument stellt die kombinierten Analysen (Per-Protocol) vor (Seite 198). Des Weiteren kann die Anzahl der Patienten, die Lost-to-Follow-Up waren, nicht von der Anzahl der Patienten unterschieden werden, die gegen den Prüfplan verstießen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Therapieeffekte im Hinblick auf den FEV₁ wurden, wie in den Methoden beschrieben, angegeben. Infolge der Randomisierung werden die Behandlungsgruppen in Bezug auf die Prognose ausgeglichen sein. Die Therapieeffekte wurden jedoch um die Baseline-Merkmale der Patienten bereinigt. Dies könnte möglicherweise zu einem Ungleichgewicht zwischen den Behandlungsgruppen geführt haben. Diese Anpassungen könnten durchgeführt worden sein, um die beiden Studien besser kombinieren zu können. Dies wurde jedoch nicht angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Patienten wurden ausgeschlossen, weil Daten nicht ausgewertet werden konnten (Seite 200).

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die oben genannte Bewertung der Risikoverzerrung weist darauf hin, dass die Berichterstattung in dieser Publikation nicht ausreichend war. Wir sahen uns daher außerstande, die Qualität der Studie zu bescheinigen. Darüber hinaus wurden Patienten ausgeschlossen, weil Daten nicht ausgewertet werden konnten. Die Bereinigung um Therapieeffekte ist unter Umständen nicht angemessen.

Endpunkt: SGRQ

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und

Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Dieses Dokument stellt die kombinierten Analysen (Per-Protocol) vor (Seite 198). Des Weiteren kann die Anzahl der Patienten, die Lost-to-Follow-Up waren, nicht von der Anzahl der Patienten unterschieden werden, die gegen den Prüfplan verstießen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Therapieeffekte im Hinblick auf den SGRQ wurden wie in den Methoden beschrieben vorgestellt. Infolge der Randomisierung werden die Behandlungsgruppen in Bezug auf die Prognose ausgeglichen sein. Die Therapieeffekte wurden jedoch um die Baseline-Merkmale der Patienten bereinigt. Dies könnte möglicherweise zu einem Ungleichgewicht zwischen den Behandlungsgruppen geführt haben. Diese Anpassungen könnten durchgeführt worden sein, um die beiden Studien besser kombinieren zu können. Dies wurde jedoch nicht angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Patienten wurden ausgeschlossen, weil Daten nicht ausgewertet werden konnten (Seite 200).

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die oben genannte Bewertung der Risikoverzerrung weist darauf hin, dass die Berichterstattung in dieser Publikation nicht ausreichend war. Wir sahen uns daher außerstande, die Qualität der Studie zu bescheinigen. Darüber hinaus wurden Patienten ausgeschlossen, weil Daten nicht ausgewertet werden konnten. Die Bereinigung um Therapieeffekte ist unter Umständen nicht angemessen.

Endpunkt: TDI

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu

beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Dieses Dokument stellt die kombinierten Analysen (Per-Protocol) vor (Seite 198). Des Weiteren kann die Anzahl der Patienten, die Lost-to-Follow-Up waren, nicht von der Anzahl der Patienten unterschieden werden, die gegen den Prüfplan verstießen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Therapieeffekte im Hinblick auf den TDI wurden, wie in den Methoden beschrieben, vorgestellt. Infolge der Randomisierung werden die Behandlungsgruppen in Bezug auf die Prognose ausgeglichen sein. Die Therapieeffekte wurden jedoch um die Baseline-Merkmale der Patienten bereinigt. Dies könnte möglicherweise zu einem Ungleichgewicht zwischen den Behandlungsgruppen geführt haben. Diese Anpassungen könnten durchgeführt worden sein, um die beiden Studien besser kombinieren zu können. Dies wurde jedoch nicht angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Patienten wurden ausgeschlossen, weil Daten nicht ausgewertet werden konnten (Seite 200).

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die oben genannte Bewertung der Risikoverzerrung weist darauf hin, dass die Berichterstattung in dieser Publikation nicht ausreichend war. Wir sahen uns daher außerstande, die Qualität der Studie zu bescheinigen. Darüber hinaus wurden Patienten ausgeschlossen, weil Daten nicht ausgewertet werden konnten. Die Bereinigung um Therapieeffekte ist unter Umständen nicht angemessen.

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Dieses Dokument stellt die kombinierten Analysen (Per-Protocol) vor (Seite 198). Des Weiteren kann die Anzahl der Patienten, die Lost-to-Follow-Up waren, nicht von der Anzahl der Patienten unterschieden werden, die gegen den Prüfplan verstießen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Therapieeffekte im Hinblick auf Exazerbationen wurden, wie in den Methoden beschrieben, vorgestellt. Im Gegensatz zu den anderen primären Ergebnissen wurden jedoch keine Einflüsse von Kovariaten bereinigt.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die oben genannte Bewertung der Risikoverzerrung weist darauf hin, dass die Berichterstattung in dieser Publikation nicht ausreichend war. Wir sahen uns daher außerstande, die Qualität der Studie zu bescheinigen. Außerdem wurden die Therapieeffekte im Gegensatz zu den anderen beobachteten Ergebnissen nicht um den Einfluss der Kovariaten bereinigt.

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Patienten mit einer verfälschenden Erkrankung, einschließlich anderer signifikanter respiratorischer Erkrankungen, wurden ausgeschlossen, genauso wie Patienten mit Erkrankungen, die durch die Studienteilnahme möglicherweise zu einem erhöhten Risiko führen konnten. Diese Einschlusskriterien führten bei Baseline zu einem verringerten Mortalitätsrisiko. Informationen zur Gesamtmortalität und zur behandlungsbedingten Mortalität standen zur Verfügung.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-36 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Bateman 2010b

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Bateman ED, Tashkin D, Siafakas N, Dahl R, Towse L, Massey D, et al. (2010): A one-year trial of tiotropium Respimat plus usual therapy in COPD patients. <i>Respiratory Medicine</i> ; 104(10):1460-72.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die zugewiesene Behandlung wurde mithilfe eines computergenerierten Randomisierungscode bestimmt, der von Boehringer Ingelheim zur Verfügung gestellt wurde. Die Randomisierung in einem Verhältnis von 1:1 wurde vom Prüfzentrum und innerhalb der Zentren stratifiziert und blockweise durchgeführt, damit eine ausgewogene Verteilung der Behandlungsgruppen jederzeit gewährleistet war.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Art der Behandlung wurde gegenüber den Prüffärzten, den auswertenden Wissenschaftlern und den Patienten verblindet.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die direkt mit der Durchführung und Analyse der Studie beauftragten Personen erhielten erst nach Beendigung der Studie Zugriff auf die Informationen über die zugewiesene Behandlung.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die direkt mit der Durchführung und Analyse der Studie beauftragten Personen erhielten erst nach Beendigung der Studie Zugriff auf die Informationen über die zugewiesene Behandlung.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*

- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Ergebnisse wurden um den Einfluss von Kovariaten bereinigt, was im Rahmen randomisierter Studien unnötig sein sollte (z. B. Behandlung mit LABA). Eine Subgruppenanalyse wurde nur für mit LABA behandelte Patienten durchgeführt (ja / nein). Andere Subgruppenanalysen könnten unter Umständen ebenfalls informativ sein.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Patientenfluss: Es ist nicht völlig klar, wie das „behandelte Set“ für die Analyse kalkuliert wurde.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Prüfärzte, Bewerter und Patienten waren in Bezug auf die Identität der Medikamente verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs,

Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Patienten wurden wegen „anderer“ Gründe vom Full-Analysis-Set ausgeschlossen (Abb. 1). Die Anzahl der Patienten, bei denen die Bewertung des FEV₁ erfolgte, entsprach nicht der Anzahl der Patienten, die in der Wirksamkeitsanalyse berücksichtigt werden sollten. Die Anzahl der Patienten, die bei der Analyse nicht berücksichtigt wurden, lag bei knapp über 5%..

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Das Ergebnis wurde um den Einfluss von Kovariaten bereinigt, was nach der Randomisierung nicht erforderlich sein sollte. Zudem wurde eine Subgruppenanalyse nur für mit LABA behandelte Patienten durchgeführt (ja / nein). Andere Subgruppenanalysen könnten unter Umständen ebenfalls informativ sein.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Es ist nicht klar, welche Patienten bei der FEV₁-Bewertung berücksichtigt wurden.

Endpunkt: SGRQ

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Prüfärzte, Bewerter und Patienten waren in Bezug auf die Identität der Medikamente verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Patienten wurden wegen „anderer“ Gründe vom Full-Analysis-Set ausgeschlossen (Abb. 1). Die Anzahl der im Rahmen des SGRQ bewerteten Patienten wurde nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der SGRQ wurde mithilfe eines ANCOVA-Modells analysiert. Im Gegensatz zu anderen Ergebnissen wurde jedoch nicht beschrieben, ob die Auswirkungen der Behandlung um den Einfluss von Kovariaten bereinigt wurden. Zudem wurde eine Subgruppenanalyse nur für mit LABA behandelte Patienten durchgeführt (ja / nein). Andere Subgruppenanalysen könnten unter Umständen ebenfalls informativ sein.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Im Gegensatz zu anderen Ergebnissen wurde die Anzahl der Patienten, die bei der Analyse des SGRQ berücksichtigt wurden, nicht explizit angegeben.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Verzerrungsrisiko ist aufgrund des nicht eindeutigen Patientenflusses schwierig zu bestimmen.

Endpunkt: Exazerbationen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Prüfärzte, Bewerter und Patienten waren in Bezug auf die Identität der Medikamente verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Patientenfluss wurde eindeutig beschrieben, außer dass Patienten wegen „anderer“ Gründe vom Full-Analysis-Set ausgeschlossen wurden (Abb. 1).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden die Auswirkungen der Behandlung auf zwei definierte Exazerbationen angegeben (Abb. 2), was für eine größere Transparenz sorgt. Der Therapieeffekt im Hinblick auf Exazerbationen wurde jedoch um den Einfluss von Kovariaten bereinigt, was für das randomisierte Studiendesign nicht erforderlich ist. Das Ergebnis wurde um den Einfluss von Kovariaten bereinigt, was nach der Randomisierung nicht erforderlich sein sollte. Zudem wurde eine Subgruppenanalyse nur für mit LABA behandelte Patienten durchgeführt (ja / nein). Andere Subgruppenanalysen könnten unter Umständen ebenfalls informativ sein.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Kardiovaskuläre Morbidität und Mortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Prüfärzte, Bewerter und Patienten waren in Bezug auf die Identität der Medikamente verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Patientenfluss wurde eindeutig beschrieben, außer dass Patienten wegen „anderer“ Gründe vom Full-Analysis-Set ausgeschlossen wurden (Abb. 1). Patienten mit bekannten kardiovaskulären Erkrankungen wurden in die Studie aufgenommen, es sei denn, sie erfüllten folgende Kriterien: frühere instabile Arrhythmien, Myokardinfarkt in den letzten 6 Monaten oder kardiales Versagen, das in den letzten 12 Monaten eine stationäre Behandlung erforderte. Informationen über die kardiovaskuläre Anamnese und die Randomisierung wurden nicht zur Verfügung gestellt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Informationen über die kardiovaskuläre Morbidität und Mortalität werden im angemessenen Umfang gegeben und können nicht zu Gunsten der Studienmedikation gedeutet werden.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Prüfärzte, Bewerter und Patienten waren in Bezug auf die Identität der Medikamente verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Patientenfluss wurde eindeutig beschrieben, außer dass Patienten wegen „anderer“ Gründe vom Full-Analysis-Set ausgeschlossen wurden (Abb. 1). Strenge Ausschlusskriterien für Patienten im Hinblick auf Komorbiditäten, ausgenommen kardiovaskuläre Erkrankungen. Angaben zu einer adäquaten Randomisierung der Patienten mit kardiovaskulären Erkrankungen liegen nicht vor.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Angaben zur Mortalität nach Organklasse pro Behandlungsarm.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-37 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Brusasco 2003

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Brusasco V, Hodder R, Miravittles M, Korducki L, Towse L, Kesten S (2003): Health outcomes following treatment for six months with once daily tiotropium compared with twice daily salmeterol in patients with COPD.[Erratum appears in Thorax. 2005 Feb;60(2):105]. Thorax; 58(5):399-404.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie war doppelblind.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studie war doppelblind. Wir sind davon ausgegangen, dass Patient und Arzt verblindet waren, der auswertende Wissenschaftler jedoch nicht notwendigerweise.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).
- Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:
 - erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung
 - Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe
 - geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben
- Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.
- Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

FEV₁: Die Therapieeffekte wurden im Hinblick auf ihre Signifikanz beschrieben.

Die Therapieeffekte wurden nicht immer angegeben.

TDI und SGRQ wurden in Woche 8, 16 und 24 untersucht. Die Ergebnisse wurden

jedoch nur für die Beurteilung nach 24 Wochen angegeben.

Exazerbationen: Es wurden Ergebnisse für alle definierten Exazerbationen angegeben.

Informationen über die Fallzahlplanung wurden nicht gegeben.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es wurden zwei Studien kombiniert, die Randomisierung gilt nur innerhalb einer Studie. Die Effekte der Kombination dieser Studien wurden nicht bereinigt.

Unvollständige Berichterstattung aufgrund von Loss-to-Follow-Up.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die kombinierten Studien waren vergleichbar. Es ist deshalb davon auszugehen, dass die Kombination die Ergebnisse nicht verzerrt hat.

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Begründung für die Studienabbrüche war unzureichend. Es wurde keine ITT-Analyse durchgeführt und der Anteil der Patienten, die die Studie abbrachen, lag bei über 5 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

FEV₁: Die Therapieeffekte wurden im Hinblick auf ihre Signifikanz beschrieben.
Die Therapieeffekte wurden nicht immer angegeben.

Therapieeffekte für alle durchgeführten Messungen wurden nur für das mittlere FEV₁ angegeben, nicht aber für das Peak- und Trough-FEV₁.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Therapieeffekte für alle durchgeführten Messungen wurden nur für das mittlere FEV₁ angegeben, nicht aber für das Peak- und Trough-FEV₁.

Endpunkt: SGRQ**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Begründung für die Studienabbrüche war unzureichend. Die Anzahl der Patienten, die anhand des SGRQ bewertet wurden, wurde nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

TDI und SGRQ wurden in Woche 8, 16 und 24 untersucht. Die Ergebnisse wurden jedoch nur für die Beurteilung nach 24 Wochen angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: SGRQ Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Begründung für die Studienabbrüche war unzureichend. Die Anzahl der Patienten, die anhand des SGRQ bewertet wurden, wurde nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

TDI und SGRQ wurden in Woche 8, 16 und 24 untersucht. Die Ergebnisse wurden jedoch nur für die Beurteilung nach 24 Wochen angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: TDI

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei

einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Anzahl der Patienten, die hinsichtlich des TDI untersucht wurden, wurde nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

TDI und SGRQ wurden in Woche 8, 16 und 24 untersucht. Die Ergebnisse wurden jedoch nur für die Beurteilung nach 24 Wochen angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: TDI Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Anzahl der Patienten, die hinsichtlich des TDI untersucht wurden, wurde nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

TDI und SGRQ wurden in Woche 8, 16 und 24 untersucht. Die Ergebnisse wurden jedoch nur für die Beurteilung nach 24 Wochen angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Exazerbationen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Exazerbationen: Es wurden Ergebnisse für verschiedene definierte Exazerbationen angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Exazerbationen: Es wurden Ergebnisse für verschiedene definierte Exazerbationen angegeben.

Endpunkt: Krankenhausaufenthalte

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es handelte sich um eine randomisierte doppelblinde Studie.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Im Artikeltext wurden Angaben zu Krankenhausaufenthalten aus beliebigen Ursachen gemacht. In Tabelle 2 fehlen die Standardfehler-Werte für

Krankenhausaufenthalte infolge von Exazerbationen, für Krankenhausaufenthalte aus beliebiger Ursache werden diese Werte angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Vollständige Angaben liegen nur zu den Krankenhausaufenthalten aufgrund beliebiger Ursachen vor.

Tabelle 4.7-38 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Beeh 2006

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Beeh KM, Beier J, Buhl R, Stark-Lorenzen P, Gerken F, Metzendorf N, et al. (2006): [Efficacy of tiotropium bromide (Spiriva) in patients with chronic-obstructive pulmonary disease (COPD) of different severities]. [German]. Pneumologie (Stuttgart, Germany); 60(6):341-6.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Identische Behälter (HandiHaler) wurden in beiden Gruppen verwendet.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Nur das Trough-FEV₁ und Exazerbationen wurden als Ergebnisse vorgezählt.
Weitere Ergebnisse erschienen im Abschnitt Ergebnisse.

Das FEV₁ wurde hinsichtlich der Prüfzentren angepasst, die Exazerbationen hingegen nicht.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Patientenfluss ist unklar im Hinblick auf Studienabbrüche und das Full-Analysis-Set.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Das FEV₁ wurde hinsichtlich der Prüfzentren angepasst, die Exazerbationen hingegen nicht.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Bereinigung um Prüfzentren, jedoch keine Adjustierung der Exazerbationen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Patientenfluss ist unklar im Hinblick auf Studienabbrüche und das Full-Analysis-Set.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Das FEV₁ wurde hinsichtlich der Prüfzentren angepasst, die Exazerbationen hingegen nicht.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Der Patientenfluss ist unklar und die statistischen Methoden weichen von anderen Endpunkten ab.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-39 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Casaburi 2000

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Casaburi R, Briggs DD, Jr., Donohue JF, Serby CW, Menjoge SS, Witek TJ, Jr. (2000): The spirometric efficacy of once-daily dosing with tiotropium in stable COPD: a 13-week multicenter trial. The US Tiotropium Study Group. Chest; 118(5):1294-302.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde erwähnt, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Die Beschreibung der statistischen Analyse war unzureichend.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Obwohl Loss-to-Follow-Up geschildert wurde (>10 %), beschrieben die Autoren die verwendeten statistischen Analysemethoden ausgiebig.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Methoden der statistischen Analyse werden nicht detailliert beschrieben. Aus diesem Grund ist es schwierig, die Ergebnisse auf Verzerrungen zu untersuchen.

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Obwohl Loss-to-Follow-Up geschildert wurde (>10 %), beschrieben die Autoren die verwendeten statistischen Analysemethoden ausgiebig.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Methoden der statistischen Analyse werden nicht detailliert beschrieben. Aus diesem Grund ist es schwierig, die Ergebnisse auf Verzerrungen zu untersuchen.

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

Der Artikel gibt an, dass im Placeboarm unterschiedliche Prozentsätze von Patienten die Studie aufgrund von unerwünschten Ereignissen abbrachen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**
-

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Der Artikel gibt an, dass im Placeboarm unterschiedliche Prozentsätze von Patienten die Studie aufgrund von unerwünschten Ereignissen abbrachen.

Tabelle 4.7-40 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Casaburi 2002

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Casaburi R., Mahler D.A., Jones P.W., Wanner A., San P.G., ZuWallack R.L. et al. Langfristige Bewertung von einmal täglich inhaliertem Tiotropium bei chronisch-obstruktiver Lungenerkrankung. European Respiratory Journal Februar 2002;19(2):217-24.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde erwähnt, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

Aclidiniumbromid (Eklira® Genuair® 322 Mikrogramm Pulver zur Inhalation) Seite 960 von 1250
Bretaris® Genuair® 322 Mikrogramm Pulver zur Inhalation)

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine detaillierten Angaben zur Erzeugung der Randomisierungssequenz in der Publikation

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Beide Gruppen bekamen Wirkstoff bzw. Placebo mit dem Handyhaler appliziert.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Loss-to-Follow-Up war erheblich (>10%). Zwischen den Behandlungsgruppen gab es einen Unterschied beim Loss-to-Follow-Up (>5 %). Die ITT wurde zur Analyse herangezogen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: TDI

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Loss-to-Follow-Up war erheblich (>10%). Zwischen den Behandlungsgruppen gab es einen Unterschied beim Loss-to-Follow-Up (>5 %). Obwohl angegeben wurde, dass eine ITT-Analyse erfolgte, war die Anzahl der Patienten, für die eine TDI-Bewertung durchgeführt wurde, geringer als die Anzahl der randomisierten Patienten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Exazerbationen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Loss-to-Follow-Up war erheblich (>10%). Zwischen den Behandlungsgruppen gab es einen Unterschied beim Loss-to-Follow-Up (>5 %).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**
-

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-41 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SAFE (Chan 2007)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Chan CK, Maltais F, Sigouin C, Haddon JM, Ford GT, Group SS (2007): A randomized controlled trial to assess the efficacy of tiotropium in Canadian patients with chronic obstructive pulmonary disease. Canadian Respiratory Journal; 14(8):465-72.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine detaillierten Angaben zur Erzeugung der Randomisierungssequenz publiziert.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet.

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Beide Gruppen bekamen Wirkstoff bzw. Placebo mit dem Handyhaler appliziert.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*

- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, die die Studie abbrachen, lag bei <5 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-42 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Covelli 2005

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Covelli H, Bhattacharya S, Cassino C, Conoscenti C, Kesten S (2005): Absence of electrocardiographic findings and improved function with once-daily tiotropium in patients with chronic obstructive pulmonary disease. <i>Pharmacotherapy: The Journal of Human Pharmacology & Drug Therapy</i> ; 25(12):1708-18.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunkübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Genauere Beschreibung fehlt.

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlages, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Genauere Beschreibung fehlt.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Der Patientenfluss war unklar. Bei der Anzahl der randomisierten Patienten und der Anzahl der Patienten, deren FEV₁ analysiert wurde, lag eine Differenz vor.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu

beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde nicht beschrieben, ob das ITT-Prinzip für die Analyse herangezogen wurde. Der Patientenfluss war unklar.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde nicht beschrieben, ob das ITT-Prinzip für die Analyse herangezogen wurde. Der Patientenfluss war unklar.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Die Veränderungen des Peak-FEV₁ gegenüber Baseline wurden mithilfe eines ANCOVA-Modells analysiert; dieses umfasste als Faktoren die Therapie und die Zentren und als Kovariaten die Baseline-FEV₁-Werte.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde nicht beschrieben, ob das ITT-Prinzip für die Analyse herangezogen wurde. Der Patientenfluss war unklar.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Es wurden nicht für alle Patienten Angaben zur Art der SUE gemacht, nur das häufigste SUE für Placebo (kardiologisches Ereignis) wurde detaillierter beschrieben.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Patienten mit einer signifikanten Erkrankung außer COPD wurden von der Studie ausgeschlossen. Patienten mit einer vorbestehenden kardiovaskulären Erkrankung wurde die Teilnahme an der Studie gestattet, sofern sie keines der folgenden Kriterien erfüllten: Myokardinfarkt in den letzten sechs Monaten, Krankenhausaufenthalt aufgrund von kardialem Versagen im letzten Jahr oder lebensbedrohliche Arrhythmien, die eine Intervention oder eine Änderung der medikamentösen Therapie im letzten Jahr erforderten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-43 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie INHANCE (Donohue 2010)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Donohue JF, Fogarty C, Lotvall J, Mahler DA, Worth H, Yorgancioglu A, et al. (2010): Once-daily bronchodilators for chronic obstructive pulmonary disease: indacaterol versus tiotropium. American Journal of Respiratory & Critical Care Medicine; 182(2):155-62.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurde ein automatisches interaktives Voice-Response-System verwendet (Anhang).

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Behandlung mit Tiotropium war Open-Label.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung

findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, der aufgrund von Loss-to-Follow-Up nicht evaluiert wurde, lag bei >10 %. Die Anzahl der Patienten, die in die Wirksamkeitsanalyse (N=415 und 418, Abb. 1) aufgenommen werden sollten, entsprach nicht der Anzahl der Patienten in der FEV₁-Bewertung.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Loss-to-Follow-Up war erheblich und die Tiotropium-Gruppe war nicht verblindet..

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B.

diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, der aufgrund von Loss-to-Follow-Up nicht evaluiert wurde, lag bei >10 %. Die Anzahl der Patienten, die in die Wirksamkeitsanalyse (N=415 und 418, Abb. 1) aufgenommen werden sollten, entsprach nicht der Anzahl der Patienten in der FEV1-Bewertung.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Loss-to-Follow-Up war erheblich und die Tiotropium-Gruppe war nicht verblindet..

Endpunkt: SGRQ

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

- nein:** Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, der aufgrund von Loss-to-Follow-Up nicht evaluiert wurde, lag bei >10 %. Die Anzahl der Patienten, die in die Wirksamkeitsanalyse (N=415 und 418, Abb. 1) aufgenommen werden sollten, entsprach nicht der Anzahl der Patienten in der SGRQ-Bewertung.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen

Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Loss-to-Follow-Up war erheblich und die Tiotropimu-Gruppe war nicht verblindet.

Endpunkt: SGRQ Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass

diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zu Gunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, der aufgrund von Loss-to-Follow-Up nicht evaluiert wurde, lag bei >10 %. Die Anzahl der Patienten, die in die Wirksamkeitsanalyse (N=415 und 418, Abb. 1) aufgenommen werden sollten, entsprach nicht der Anzahl der Patienten in der SGRQ-Bewertung.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**
-
-

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Loss-to-Follow-Up war erheblich und die Tiotropimu-Gruppe war nicht verblindet.

Endpunkt: TDI**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, der aufgrund von Loss-to-Follow-Up nicht evaluiert wurde, lag bei >10 %. Die Anzahl der Patienten, die in die Wirksamkeitsanalyse (N=415 und 418, Abb. 1) aufgenommen werden sollten, entsprach nicht der Anzahl der Patienten in der TDI-Bewertung.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die nicht adäquate Verblindung und die unzureichende Anwendung des ITT-Prinzips wird die geschätzte Wirksamkeit beeinträchtigen.

Endpunkt: TDI-Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien,

statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, der aufgrund von Loss-to-Follow-Up nicht evaluiert wurde, lag bei >10 %. Die Anzahl der Patienten, die in die Wirksamkeitsanalyse (N=415 und 418, Abb. 1) aufgenommen werden sollten, entsprach nicht der Anzahl der Patienten in der TDI-Bewertung.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die nicht adäquate Verblindung und die unzureichende Umsetzung des ITT-Prinzips wird die beobachteten Effekte beeinflussen und die Ergebnisse verzerren..

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Die Patienten verzeichneten unerwünschte Ereignisse auf Tagebuchkarten. Die Sicherheitsdaten wurden durch ein gemischtes Modell mit Behandlung, Raucherstatus und Land als festen Effekten und dem Baseline-FEV₁-Wert und der Reversibilität als Kovariaten analysiert. Fehlende Werte in Woche 12 wurden durch den vorausgehenden Wert (aus Woche 2 oder später) ersetzt. Die Aufzeichnung durch die Patienten und das Ersetzen fehlender Werte in Woche 12 durch Werte aus Woche 2 könnte die Qualität der Ergebnisse beeinträchtigen. Dieses Vorgehen war jedoch für alle Arme identisch, weshalb aufgrund dieser Datensammlungsmethode und der Methode, fehlende Werte zu ersetzen, keine vermehrten Verzerrungen zwischen den Armen erwartet werden.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die nicht adäquate Verblindung wird die geschätzte Wirksamkeit beeinträchtigen.

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT(„intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien,

statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die nicht adäquate Verblindung wird die geschätzte Wirksamkeit und Sicherheit beeinträchtigen.

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die nicht adäquate Verblindung wird die geschätzte Wirksamkeit und Sicherheit beeinträchtigen.

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der Tiotropium-Arm wurde nicht verblindet. Es wird jedoch nicht erwartet, dass diese Tatsache die Letalitäts- und die Gesamtmortalitätsrate beeinflusst.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Mortalitätsanalyse basierte höchstwahrscheinlich auf der Sicherheitspopulation. Eindeutige Angaben hierzu fehlen jedoch. Die Sicherheitspopulation umfasste alle Patienten, die mindestens eine Dosis des Studienmedikaments erhalten hatten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die inadäquate Verblindung beeinflusste die Ergebnisse im Hinblick auf Letalität und Mortalität während der Studie wahrscheinlich nicht.

Tabelle 4.7-44 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Donohue 2002

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Donohue JF, van Noord JA, Bateman ED, Langley SJ, Lee A, Witek TJ, Jr., et al. (2002): A 6-month, placebo-controlled study comparing lung function and health status changes in COPD patients treated with tiotropium or salmeterol. Chest; 122(1):47-55.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung

findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der Bewertung des FEV₁ berücksichtigt wurden, lag bei >10 %. Es war unklar, ob das ITT-Prinzip angewendet wurde, um die Patienten einzuschließen (Full-Analysis-Set?).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Loss-to-Follow-Up in der Placebo-Gruppe war erheblich. Keine ITT-Analyse.

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien

erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der

Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der Bewertung des FEV₁ berücksichtigt wurden, lag bei >10 %. Es war unklar, ob das ITT-Prinzip angewendet wurde, um die Patienten einzuschließen (Full-Analysis-Set?).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts
(ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Der Loss-to-Follow-Up in der Placebo-Gruppe war erheblich; die Gründe für die Studienabbrüche wurden nicht im Detail beschrieben.

Endpunkt: SGRQ

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der SGRQ-Bewertung berücksichtigt wurden, lag bei >10 %. Es war unklar, ob das ITT-Prinzip angewendet wurde, um die Patienten einzuschließen (Full-Analysis-Set?).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Loss-to-Follow-Up in der Placebo-Gruppe war erheblich. Keine ITT-Analyse.

Endpunkt: SGRQ Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei

einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der SGRQ-Bewertung berücksichtigt wurden, lag bei >10 %. Es war unklar, ob das ITT-Prinzip angewendet wurde, um die Patienten einzuschließen (Full-Analysis-Set?).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Loss-to-Follow-Up in der Placebo-Gruppe war erheblich. Keine ITT-Analyse.

Endpunkt: TDI

1. Verblindung der Endpunkterheber

*Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.
In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.*

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante

Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der TDI-Bewertung berücksichtigt wurden, lag bei >10 %. Es war unklar, ob das ITT-Prinzip angewendet wurde, um die Patienten einzuschließen (Full-Analysis-Set?).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*

- Anwendung inadäquater statistischer Verfahren

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Das Loss-to-Follow-Up in der Placebo-Gruppe war erheblich. Keine ITT-Analyse.

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, die bei der Bewertung von Exazerbationen berücksichtigt wurden, war nicht klar.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**
-

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-45 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MISTRAL (Dusser 2006)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Dusser D, Bravo ML, Iacono P (2006): The effect of tiotropium on exacerbations and airflow in patients with COPD.[Erratum appears in Eur Respir J. 2006 May;27(5):1076]. European Respiratory Journal; 27(3):547-55.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde erwähnt, dass die Therapie mittels Randomisierung zugewiesen wurde.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die

Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Anzahl der Patienten, die lost-to-follow-up waren, war gering.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-46 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SPRUCE (Freeman 2007)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Freeman D, Lee A, Price D (2007): Efficacy and safety of tiotropium in COPD patients in primary care--the SPiRiva Usual Care (SPRUCE) study. Respiratory Research; 8:45.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genauen Angaben zur Generierung der Randomisierungssequenz.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden verschiedene Subgruppenanalysen für verschiedene Ergebnisse präsentiert. Die Subgruppenanalysen wurden vorab nicht spezifiziert. (Tabelle 3-5)

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Angaben zum Randomisierungsverfahren war detailliert beschrieben. Die Subgruppenanalysen wurden vorab nicht spezifiziert.

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die

Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die für die FEV₁-Bewertung verwendete Zahl entspricht nicht den Zahlen im Ablaufdiagramm. Das LOCF-Verfahren wurde angewendet.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die beschriebenen Subgruppenanalysen wurden vorab nicht spezifiziert. Der Therapieeffekt in den Subgruppen war jedoch nicht viel größer. (Tabelle 4)

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Anzahl der Patienten, die für die Bewertung des FEV1 berücksichtigt wurden, entspricht nicht den Zahlen im Ablaufdiagramm der Studie. Dies könnte auf Verzerrungen in Bezug auf die Ergebnisse der FEV1-Bewertung hinweisen.

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser

Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die für die klinischen Bewertungen verwendete Zahl entspricht nicht den Zahlen im Ablaufdiagramm. Dies könnte sich auch auf die angegebenen Sicherheitsergebnisse auswirken. Die UEs wurden jedoch für die Dauer der Studie erhoben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine Angaben zu den angewendeten statistischen Methoden gemacht und nur sehr begrenzte Informationen zur Verfügung gestellt.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial

beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein****Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die für die klinischen Bewertungen verwendete Zahl entspricht nicht den Zahlen im Ablaufdiagramm. Dies könnte sich auch auf die angegebenen Sicherheitsergebnisse auswirken. Des Weiteren wird die Anzahl der Studienabbrüche aufgrund von UE nicht eindeutig angegeben.

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
 - Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die für die klinischen Bewertungen verwendete Zahl entspricht nicht den Zahlen im Ablaufdiagramm. Dies könnte sich auch auf die angegebenen Sicherheitsergebnisse auswirken.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Anzahl der Studienabbrüche aufgrund von UE wird nicht eindeutig angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen

Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die für die klinischen Bewertungen verwendete Zahl entspricht nicht den Zahlen im Ablaufdiagramm. Dies könnte sich auch auf die angegebenen Sicherheitsergebnisse auswirken. Des Weiteren wird die Anzahl der Studienabbrüche aufgrund von UE nicht eindeutig angegeben.

Tabelle 4.7-47 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Johansson 2008

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Johansson G, Lindberg A, Romberg K, Nordstrom L, Gerken F, Roquet A (2008): Bronchodilator efficacy of tiotropium in patients with mild to moderate COPD. Primary Care Respiratory Journal; 17(3):169-75.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde erwähnt, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Genaue Angaben zur Methodik der Erzeugung der Randomisierungssequenz fehlen in der Publikation.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Beide Gruppen erhielten Wirkstoff bzw. Placebo im Handihaler appliziert.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*

- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
Zulässige Gründe sind:
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es lagen keine Angaben zum Patientenfluss vor.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Es liegen zwar keine Informationen zum Patientenfluss vor, dennoch kann das Verzerrungspotential auf Studienebene als niedrig eingestuft werden.

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-) Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung

ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gesamtanzahl der Patienten, die die Studie abbrachen oder lost-to-follow-up waren, wird nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung

vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante

Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gesamtanzahl der Patienten, die die Studie abbrachen oder lost-to-follow-up waren, wird nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Die Kovarianzanalyse mit der Behandlungsdauer und gepooltem Zentrum wurde für als kontinuierliche Endpunkte verwendet, die Baseline-Daten dienten als Kovariaten.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gesamtanzahl der Patienten, die die Studie abbrachen oder lost-to-follow-up waren, wird nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gesamtanzahl der unerwünschten Ereignisse wird angegeben, genauso wie eine Auswahl der unerwünschten Ereignisse, die bei >3 % der Patienten in den Gruppen auftraten.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- unplausible Angaben*
- Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**
-
-

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-

Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gesamtanzahl der Patienten, die die Studie abbrachen oder lost-to-follow-up waren, wird nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gesamtanzahl der Patienten, die die Studie abbrachen oder lost-to-follow-up waren, wird nicht angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R.

auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-48 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SAFE Portugal (Moita 2008)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Moita J, Barbara C, Cardoso J, Costa R, Sousa M, Ruiz J, et al. (2008): Tiotropium improves FEV1 in patients with COPD irrespective of smoking status. Pulmonary Pharmacology & Therapeutics; 21(1):146-51.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunkübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde erwähnt, dass die Therapie mittels Randomisierung zugewiesen wurde.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Genauere Angaben zur Generierung der Randomisierungssequenz fehlen in der Publikation.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

Genauere Angaben fehlen. Beide Gruppen erhielten Wirkstoff bzw. Placebo mit dem Handihaler appliziert.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. A. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. D. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und

für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. A. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der Analyse / Bewertung berücksichtigt wurden, lag bei >5 %. Die Analysestrategie wirkte sich jedoch nicht positiv auf Tiotropium aus (Therapie von Interesse).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**

Im Tiotropium-Arm trat eine Differenz beim Baseline-FEV₁ auf sowie bei der Rauchervergangenheit für Nichtraucher verglichen mit Rauchern. Der Einfluss dieser Differenz auf die Ergebnisse ist voraussichtlich minimal.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Der Anteil der Patienten, die in der FEV₁-Ergebnisanalyse nicht berücksichtigt wurden, lag bei >5 %.

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die

Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. A. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. D. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. A. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Anzahl der Patienten in der Placebo-Gruppe, die nicht in der Analyse / Bewertung berücksichtigt wurden, lag bei >5 %. Die Analysestrategie kann jedoch nicht zugunsten von Tiotropium (Therapie von Interesse) gedeutet werden.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. A. Patienten, die die zugeteilte Therapie nicht

protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. D. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. A. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der Analyse / Bewertung berücksichtigt wurden, lag bei >5 %. Die Analysestrategie wirkte sich jedoch nicht positiv auf Tiotropium aus (Therapie von Interesse).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. A. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. D. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in

relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. A. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der Analyse / Bewertung berücksichtigt wurden, lag bei >5 %. Die Analysestrategie wirkte sich jedoch nicht positiv auf Tiotropium aus (Therapie von Interesse).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein****Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Klare Angaben, Beschreibung der auftretenden unerwünschten Ereignisse für beide Arme separat

Endpunkt: Kardiovaskuläre Morbidität und Mortalität**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. A. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. D. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. A. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Patienten wurden nicht aufgenommen, wenn bei ihnen in der Vergangenheit ein Myokardinfarkt, eine instabile Arrhythmie oder eine klinisch signifikante Erkrankung aufgetreten war, die für sie aufgrund der Teilnahme an der Studie zum Risiko werden konnte. Darüber hinaus wurden Patienten ausgeschlossen, bei denen ≥ 3 COPD Exazerbationen im vergangenen Jahr oder eine Exazerbation der unteren Atemwege innerhalb der sechs Wochen vor der Randomisierung auftraten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Für diesen Endpunkt wurden keine eindeutigen Definitionen oder Ergebnisse berichtet (es wurden nur Ergebnisse für kardiale Erkrankungen als Teil der insgesamt dokumentierten unerwünschten Ereignisse berichtet, ohne genauere Angabe, ob Todesfälle aus kardiovaskulären Ereignissen resultierten). Es ist daher nicht klar, ob diese Angaben verzerrt sind oder nicht.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Für diesen Endpunkt wurden keine eindeutigen Definitionen oder Ergebnisse angegeben. Es ist daher nicht klar, ob diese Angaben verzerrt sind oder nicht.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Für diesen Endpunkt wurden keine eindeutigen Definitionen oder Ergebnisse berichtet (es wurden nur Ergebnisse für kardiale Erkrankungen als Teil der insgesamt dokumentierten unerwünschten Ereignisse berichtet, ohne genauere Angabe, ob Todesfälle aus kardiovaskulären Ereignissen resultierten). Es ist daher nicht klar, ob diese Angaben verzerrt sind oder nicht.

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. A. Patienten, die die zugeteilte Therapie nicht

protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. D. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. A. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Patienten wurden nicht aufgenommen, wenn bei ihnen in der Vergangenheit ein Myokardinfarkt, eine instabile Arrhythmie oder eine klinisch signifikante Erkrankung aufgetreten war, die für sie aufgrund der Teilnahme an der Studie zum Risiko werden konnte. Darüber hinaus wurden Patienten mit ≥ 3 COPD Exazerbationen im vergangenen Jahr oder einer Exazerbation der unteren

Atemwege innerhalb der sechs Wochen vor der Randomisierung auftraten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**
-
-

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Einschlusskriterien für Patienten sind voraussichtlich wirksam in Bezug auf die Verringerung des Mortalitätsrisikos zur Baseline.

Tabelle 4.7-49 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Niewoehner 2005

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Niewoehner DE, Rice K, Cote C, Paulson D, Cooper JA, Jr., Korducki L, et al. (2005): Prevention of exacerbations of chronic obstructive pulmonary disease with tiotropium, a once-daily inhaled anticholinergic bronchodilator: a randomized trial.[Summary for patients in Ann Intern Med. 2005 Sep 6;143(5):I20; PMID: 16144886]. Annals of Internal Medicine; 143(5):317-26.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunkübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Eine geblockte Randomisierungsliste wurde zentral generiert.

für nicht randomisierte Studien:**Zeitliche Parallelität der Gruppen**

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)**

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genauen Angaben.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4 Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. H. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*

- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
 - *Die Methodik muss exakt beschrieben sein (z. B. Alpha Spending Approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).*
 - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Studienabbruch war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).*
 - *Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.*
 - *Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):**

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. D. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-) Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung

ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten aus der Placebo-Gruppe, die nicht bei der Analyse / Bewertung berücksichtigt wurden, lag bei >5 %. Die Analysestrategie wirkte sich jedoch nicht positiv auf Tiotropium aus (Therapie von Interesse).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Der Anteil der Patienten, die in der Ergebnisanalyse nicht berücksichtigt wurden, war beachtlich.

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Anzahl der Subgruppenanalysen war hoch (10). Es wird somit keine selektive Angabe von Ergebnissen für ein bestimmtes Variablen-Set erwartet. Das Gesamtergebnis wurde ebenfalls angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Krankenhausaufenthalte

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Die Endpunkte in Bezug auf die Krankenhausaufenthalte aufgrund von COPD wurden mit einem Cochran-Mantel-Haenszel-Test mit dem Zentrum als Stratum analysiert. Für die Analyse wurde ein zweiseitiges α -Niveau von 0,05 eingesetzt. Diese Methode weicht leicht von dem Verfahren von Böhning und Sarol ab, mit dem die Therapieunterschiede und die 95 % KI geschätzt wurden. Der geschätzte P-Wert liegt somit leicht über 0,050, während die Obergrenze für das 95 % KI leicht niedriger ist als 0.0.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Analyse mithilfe eines stratifizierten Wilcoxon-Mann-Whitney-Test unter Verwendung beobachteter (nicht ergänzter) Daten.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der

Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-50 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Powrie 2007

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Powrie DJ, Wilkinson TM, Donaldson GC, Jones P, Scrine K, Viel K, et al. (2007): Effect of tiotropium on sputum and serum inflammatory markers and exacerbations in COPD. European Respiratory Journal; 30(3):472-8.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Generierung der Randomisierungssequenz nicht im Einzelnen beschrieben.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Beide Gruppen applizierten die Studienmedikation mit dem Handihaler.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).
- Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.

Zulässige Gründe sind:

- erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung
- Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe
- geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben
- Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.
- Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“-)Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-51 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie UPLIFT (Tashkin 2008)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Tashkin DP, Celli B, Senn S, Burkhart D, Kesten S, Menjoge S, et al. (2008): A 4-year trial of tiotropium in chronic obstructive pulmonary disease. <i>New England Journal of Medicine</i> ; 359(15):1543-54.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern**Patient**

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*

- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, die in der FEV1-Analyse nicht berücksichtigt wurden, lag bei >10 %. (Abb. 1, 2) Die Begründung für den Drop-out nach 24 Wochen war nicht umfassend. Alle Studienabbrecher wurden im Detail aufgeführt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Bei der FEV₁-Bewertung wurde eine erhebliche Anzahl an Patienten nicht berücksichtigt.

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden

relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: SGRQ

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr

genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der prozentuale Anteil der Patienten, die bei der SGRQ-Bewertung nicht berücksichtigt wurden, lag bei >10 %. Insgesamt wurden über 2.500 Patienten pro Gruppe ausgewertet. Daher kann von validen Ergebnissen ausgegangen werden.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Anzahl der Patienten, die nicht im Rahmen des SGRQ bewertet wurden, war beträchtlich. Jedoch wurden insgesamt über 5.000 Patienten ausgewertet, daher kann von validen Ergebnissen ausgegangen werden.

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-52 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie TIPHON (Tonnel 2008)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Tonnel AB, Perez T, Grosbois JM, Verkindre C, Bravo ML, Brun M, et al. (2008): Effect of tiotropium on health-related quality of life as a primary efficacy endpoint in COPD. International Journal of COPD; 3(2):301-10.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Den Patienten wurde mithilfe eines computergenerierten Randomisierungsplans ohne Stratifizierung (Blockgröße = 4) eine Behandlung zugewiesen.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Beide Gruppen erhielten Wirkstoff bzw. Placebo mit dem Handihaler appliziert.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*

- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: SGRQ**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Bei der SGRQ-Bewertung wurden >5 % der randomisierten Patienten nicht berücksichtigt; Gründe für die Ausschlüsse wurden nicht umfassend beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: SGRQ Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Bei der SGRQ-Bewertung wurden >5 % der randomisierten Patienten nicht berücksichtigt; Gründe für die Ausschlüsse wurden nicht umfassend beschrieben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Exazerbationen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten

entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der Anteil der Patienten, die bei der Bewertung von Exazerbationen berücksichtigt wurden, wurde nicht eindeutig angegeben.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante

Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-53 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Verkindre 2006

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Verkindre C, Bart F, Aguilaniu B, Fortin F, Guerin JC, Le MC, et al. (2006): The effect of tiotropium on hyperinflation and exercise capacity in chronic obstructive pulmonary disease. <i>Respiration; international review of thoracic diseases</i> ; 73(4):420-7.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde erwähnt, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genau Beschreibung.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Beide Gruppen applizierten Placebo bzw. Wirkstoff mit dem Handihaler.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*

- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein

Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

In der Placebo-Gruppe wurde bei >5 % der Patienten das Trough-FEV₁ nicht bewertet, weil die Baseline-Messung fehlte.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

In der Placebo-Gruppe wurde bei >5 % der Patienten das Trough-FEV1 nicht bewertet, weil die Baseline-Messung fehlte.

Endpunkt: Peak-FEV₁

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Analyse des Peak-FEV₁ wurde in der ITT-Population durchgeführt, ausgenommen ein Patient in der Tiotropium-Gruppe aufgrund fehlender Wirksamkeitsdaten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja **nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: SGRQ

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die

Einstufung:

Die Zahl der Patienten, bei denen die SGRQ-Analyse durchgeführt wurde, unterschied sich für die einzelnen Items; eine Erklärung erfolgt nicht.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Der SGRQ wurde außerdem auf einer Unterskala des Fragebogens angegeben.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Zahl der Patienten, bei denen die SGRQ-Analyse durchgeführt wurde, unterscheiden sich für die einzelnen Items; eine Erklärung erfolgt nicht.

Endpunkt: SGRQ Responder

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten

entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Zahl der Patienten, bei denen die SGRQ-Analyse durchgeführt wurde, unterschied sich für die einzelnen Items; eine Erklärung erfolgt nicht.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Der SGRQ wurde außerdem auf einer Unterskala des Fragebogens angegeben.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Die Zahl der Patienten, bei denen die SGRQ-Analyse durchgeführt wurde, unterscheiden sich für die einzelnen Items; eine Erklärung erfolgt nicht.

Endpunkt: TDI**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es ist nicht klar, warum die Anzahl der Probanden, bei denen der BDI / TDI-Fokal-Score analysiert wurde, sich von allen anderen Analysen unterscheidet. Eine Erklärung erfolgt nicht.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch

mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tabelle 4.7-54 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Vogelmeier 2008

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Vogelmeier C, Kardos P, Harari S, Gans SJ, Stenglein S, Thirlwell J (2008): Formoterol mono- and combination therapy with tiotropium in patients with COPD: a 6-month study. <i>Respiratory Medicine</i> ; 102(11):1511-20.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1. für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2. für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierte Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Tiotropium wurde offen gegeben.

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Behandlung mit Tiotropium war Open-Label.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Behandlung mit Tiotropium war Open-Label.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*

- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Ergebnisse zu Exazerbationen wurden nicht wie im Abschnitt Statistische Analysen beschrieben angegeben, sie wurden anders als in anderen Studien angeführt. Es wurden keine statistischen Analysen in Bezug auf den SGRQ beschrieben.

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Bei der Therapie von Interesse (Tiotropium) handelte es sich um eine Open-Label-Behandlung. Die Ergebnisse können unter Umständen in einer für Tiotropium positiven Art und Weise präsentiert werden.

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: SGRQ**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Behandlung mit Tiotropium war Open-Label.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen

Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

>5 % der Patienten brachen die Studie ab. Alle randomisierten Patienten wurden jedoch bei der SGRQ-Bewertung berücksichtigt. Für die Analyse des SGRQ wurde die ITT-Methode angewandt. Der Einfluss der ITT-Analyse auf die SGRQ-Bewertung wirkte sich für Tiotropium nicht positiv aus.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es wurden keine statistischen Analysen in Bezug auf den SGRQ beschrieben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Open-Label-Behandlung mit Tiotropium.

Endpunkt: Krankenhausaufenthalte

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Behandlung mit Tiotropium war Open-Label.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante

Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

>5 % der Patienten brachen die Studie ab. Alle randomisierten Patienten wurden jedoch bei der Bewertung der Exazerbationen berücksichtigt (ITT-Methode).

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Anzahl der Tage mit COPD Exazerbationen basiert auf einer anderen Anzahl von Probanden als die Anzahl der COPD Exazerbationen, die einen Krankenhausaufenthalt erforderten. Warum diese Zahlen voneinander abweichen ist nicht klar. Die Anzahl der COPD Exazerbationen, die einen Krankenhausaufenthalt erforderten, wird jedoch für die ITT-Population angegeben.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*

- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Open-Label-Behandlung mit Tiotropium.

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Behandlung mit Tiotropium war Open-Label.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugewiesene Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Unerwünschte Ereignisse wurden für die Gesamtpopulation und die Population im Alter von ≥ 65 Jahren berichtet. Der Anteil der Patienten ≥ 65 Jahre variierte in den unterschiedlichen Behandlungsarmen (von 39,5 % für Formoterol bis 48,4 % für Tiotropium) und im Hinblick auf das Auftreten von UE wurden zwischen dem Gesamtpatientenpool und den Patienten ≥ 65 Jahre in verschiedenen Behandlungsarmen Unterschiede beobachtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Open-Label-Behandlung mit Tiotropium und möglicher Einfluss der Zusammensetzung der Population in den verschiedenen Behandlungsarmen.

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Behandlung mit Tiotropium war Open-Label.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- unplausible Angaben*
- Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Open-Label-Behandlung mit Tiotropium.

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Tiotropium wurde als Open-Label-Behandlung verabreicht. Diese Tatsache wird die in der Studie beobachteten Letalitäts- und Gesamtmortalitätsraten jedoch voraussichtlich nicht beeinflussen.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten

ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Patienten mit Atemwegsinfektionen oder Krankenhausaufenthalten aufgrund akuter

COPD Exazerbationen in dem Monat vor dem Screening wurden von der Studie ausgeschlossen. Patienten mit einer klinisch signifikanten Erkrankung wie einer ischämischen Herzkrankheit, die die Sicherheit oder Compliance des Patienten beeinträchtigen konnte, wurden ebenfalls ausgeschlossen. Es ist unklar, ob Patienten, die zuvor beispielsweise einen Myokardinfarkt erlitten, aufgenommen wurden.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- un plausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Tiotropium wurde als Open-Label-Behandlung verabreicht. Diese Tatsache wird die in der Studie beobachteten Letalitäts- und Gesamtmortalitätsraten jedoch voraussichtlich nicht beeinflussen.

Tabelle 4.7-55 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie Voshaar 2008

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
Voshaar T, Lapidus R, Maleki-Yazdi R, Timmer W, Rubin E, Lowe L, et al. (2008): A randomized study of tiotropium Respimat Soft Mist inhaler vs. ipratropium pMDI in COPD. <i>Respiratory Medicine</i> ; 102(1):32-41.	A

Kriterien zur Bewertung des Verzerrungspotenzials der Ergebnisse:

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A: Kriterien des Verzerrungspotenzials der Ergebnisse auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anhaltspunkte vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es wurde angegeben, dass die Therapie mittels Randomisierung zugewiesen wurde.

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genauen Angaben in der Publikation.

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

ja: Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln / Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

nein: Die Gruppenzuteilung erfolgte nicht verdeckt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Keine genauen Angaben in der Publikation

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja: Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anhaltspunkte dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

nein: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und Behandlern

Patient

ja: Die Patienten waren verblindet.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

Behandler bzw. Weiterbehandler

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, den primären Behandler (Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte / -räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*

- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.
Zulässige Gründe sind:*
 - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
 - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
 - *geplante Interimsanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anhaltspunkte für eine ergebnisgesteuerte Darstellung eines Endpunkts das Verzerrungspotenzial für die Ergebnisse der übrigen Endpunkte erhöhen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anhaltspunkten dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, ist das Verzerrungspotenzial für die anderen Endpunkte u. U. erhöht. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Erhöhung des Verzerrungspotenzials der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen.

Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und das Verzerrungspotenzial anderer Endpunkte nicht berührt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Fehlen sonstiger (endpunktübergreifender) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimsanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimsanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, der Grund für den Abbruch der Studie war, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, der Grund für den Abbruch der Studie war, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Das Ansprechen auf die Behandlung in Bezug auf das FEV₁ wurde um den Einfluss des Rauchens bereinigt. Dies ist im Rahmen eines randomisierten Behandlungsmodells möglicherweise nicht angebracht.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

B: Kriterien des Verzerrungspotenzials der Ergebnisse pro Endpunkt

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: Trough-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu

beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es lagen keine Angaben zum Patientenfluss vor.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

 ja **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Peak-FEV₁**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden,

wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren

möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es lagen keine Angaben zum Patientenfluss vor.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Das Peak-FEV₁ wurde als Endpunkt oder Teil der Beurteilungen im Abschnitt Methoden des Artikels nicht erwähnt.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Unerwünschte Ereignisse (adverse events, AE)**1. Verblindung der Endpunkterheber**

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine

Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es lagen keine Angaben zum Patientenfluss vor. Die Bewertung der Sicherheit basierte auf allen randomisierten Patienten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Unerwünschte Ereignisse, die bei ≥ 3 % der Patienten in den Behandlungsgruppen auftraten, wurden berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Schwerwiegende unerwünschte Ereignisse (SUE)

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in

begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es lagen keine Angaben zum Patientenfluss vor. Die Bewertung der Sicherheit basierte auf allen randomisierten Patienten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
 - Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).

- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es lagen keine Angaben zum Patientenfluss vor.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Die Anzahl der Studienabbrüche aufgrund von unerwünschten Ereignissen ist für Ipratropium höher als in den anderen Gruppen; dies entspricht dem Auftreten schwerwiegender unerwünschter Ereignisse. Die Anzahl beliebiger unerwünschter Ereignisse ist in der Ipratropium-Gruppe jedoch geringer als in den anderen Gruppen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:

Endpunkt: Letalität und Gesamtmortalität

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

ja: Der Endpunkt wurde verblindet erhoben.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Es wurde angegeben, dass diese Studie doppelblind war.

2. Adäquate Umsetzung des ITT-Prinzips

Lost-to-follow-up-Patienten sind Patienten, bei denen die Zielkriterien nicht vollständig bis zum Studienende erhoben werden können (z. B. weil ein Patient seine Einwilligung zurückzieht). Protokollverletzer sind u. a. Patienten, die die zugeteilte Therapie nicht protokollgemäß durchlaufen (z. B. Therapieabbrecher oder -wechsler, Einnahme unerlaubter Begleitmedikationen). Zu beachten ist, dass Begriffe wie Lost-to-follow-up und Protokollverletzer in Publikationen jedoch zum Teil sehr unterschiedlich definiert sind. Des Weiteren finden sich auch Begriffe wie Drop-outs, Withdrawals etc., die in diesem Bogen möglichst zu vermeiden bzw. genau zu definieren sind. Kommen in einer Studie solche Patienten vor, ist es notwendig, dass diese ausreichend genau beschrieben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen berücksichtigt (i. d. R. ITT-Analyse, siehe Äquivalenzstudien) wurden. Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen bei Lost-to-follow-up-Patienten fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen richtlinienkonform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

ja: Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung <5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenanzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.
- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Neben anderen Kriterien wurden Patienten mit einer Erkrankung, die ihre Fähigkeit beeinflussen konnte, an der Studie teilzunehmen, von der Studie ausgeschlossen. Dies liefert keine ausreichenden Angaben darüber, ob Patienten mit einem erhöhten Mortalitätsrisiko zur Baseline von der Teilnahme ausgeschlossen wurden.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anhaltspunkte für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Todesfälle (einschließlich Ursache) wurden für Tiotropium 10 µg und Ipratropium berichtet.

4. Fehlen sonstiger (endpunktspezifischer) Aspekte, die das Verzerrungspotenzial beeinflussen

z. B.

- relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen
- unplausible Angaben
- Anwendung inadäquater statistischer Verfahren

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Falls hoch, obligate Begründung für die Einstufung:
