

Mündliche Anhörung



**Gemeinsamer
Bundesausschuss**

gemäß 5. Kapitel, § 19 Abs. 2 Verfahrensordnung
des Gemeinsamen Bundesausschusses

hier: Ataluren

Sitzung im Hause des Gemeinsamen Bundesausschusses in Berlin
am 10. Oktober 2016
von 16.20 Uhr bis 17.40 Uhr

– Stenografisches Wortprotokoll –

Angemeldete Teilnehmer der Firma **PTC Therapeutics Limited:**

Herr Boulding
Herr Dr. Hofmann
Herr Dr. Schopen
Herr Dr. Werner

Angemeldete Teilnehmer der Firma **Santhera Pharmaceuticals:**

Herr Dr. Müller-York
Herr Dr. Schupp

Angemeldete Teilnehmer für den **Bundesverband der Pharmazeutischen Industrie e. V. (BPI):**

Herr Dr. Schnorpfeil
Herr Dr. Wilken

Angemeldete Teilnehmer für den **Verband forschender Arzneimittelhersteller e. V. (vfa):**

Herr Dr. Rasch
Herr Dr. Werner

Beginn der Anhörung: 16.20 Uhr

(Die angemeldeten Teilnehmer betreten den Raum)

Herr Prof. Hecken (Vorsitzender): Meine sehr verehrten Damen und Herren, herzlich willkommen im Unterausschuss „Arzneimittel“ des Gemeinsamen Bundesausschusses, erneute Nutzenbewertung nach Fristablauf. Wir beschäftigen uns erneut mit Ataluren, Erstbeschluss vom 21. Mai 2015, befristet wegen noch ausstehender EMA-Daten. Jetzt neues Dossier und eine entsprechende Nutzendossierbewertung des Gemeinsamen Bundesausschusses vom 1. September 2016, die Gegenstand des heutigen Stellungnahmeverfahrens und der heutigen mündlichen Anhörung und Ihnen im Übrigen bekannt ist. Stellung genommen haben PTC Therapeutics Limited, der pharmazeutische Unternehmer, Santhera Pharmaceuticals, der Verband der forschenden Arzneimittelhersteller und der Bundesverband der Pharmazeutischen Industrie. Kliniker haben nicht Stellung genommen, obgleich sie wie in jedem Verfahren zur Stellungnahme aufgefordert oder um Stellungnahme gebeten worden sind.

Ich muss jetzt für das Protokoll die Anwesenheit feststellen. Für PTC Therapeutics Limited sind Herr Boulding, Herr Dr. Hofmann, Herr Dr. Schopen und Herr Dr. Werner da, für Santhera sind Herr Dr. Müller-York und Herr Dr. Schupp da, für den BPI Herr Dr. Schnorpfeil und Herr Dr. Wilken – noch am alten Platz – und für den vfa Herr Dr. Rasch. Herr Werner ist abwesend.

Seien Sie herzlich willkommen. Geschäftsleitender Hinweis: Wir führen Wortprotokoll. Deshalb immer, bevor Sie das Wort ergreifen, den Namen nennen und das Entscheideunternehmen bzw. die Institution. Wir haben hier im vorliegenden Fall, wie gesagt, Neubewertungen nach Fristablauf.

Wir haben die Ergebnisse der verblindeten Placebo-kontrollierten Phase-III-Studie inzwischen vorliegen. Ein neuer EMA-Bewertungsbericht liegt derzeit noch nicht vor. Wir sehen, dass der primäre Endpunkt, die Verbesserung der Gehstrecke, in der Gesamtgruppe nach 48 Wochen nicht signifikant ist. Lebensqualitätsdaten, auch körperliche Funktionen, sind ebenfalls relativ unauffällig. Nur in einer von fünf Subgruppen nach Gehstrecke basal liegen signifikante Ergebnisse vor, über die wir sicher diskutieren müssen. Wie gesagt: Keine Stellungnahme ist von Klinikern eingegangen, und die Fragestellung, die wir heute sicherlich sehr intensiv diskutieren müssen, ist, ob die Placebo-kontrollierte Phase-III-Studie mit den positiven Effekten auf eine von fünf Subgruppen Relevanz hat, was von der Fragestellung abzuleiten ist: Welche Relevanz hat diese eine Subgruppe, in der es dann positive Effekte gibt? Darüber werden wir sicherlich zu diskutieren haben. Selbstverständlich können wir auch alle anderen Punkte, die aus Ihrer Sicht wichtig und wesentlich sind, hier gerne und ausführlich erörtern.

Ich gebe jetzt dem pharmazeutischen Unternehmer die Möglichkeit, kurz aus seiner Sicht einzuführen, was die relevanten Punkte sind. Dann machen wir eine Frage- und Antwort-Runde und schauen einmal, was wir im Verlauf der nächsten Stunde oder eineinhalb Stunden hier noch an Evidenz hervorzulocken vermögen oder an bislang verborgener Evidenz auf den Tisch legen oder an interessanten Fragen beantworten. – Wer beginnt? – Herr Schopen, bitte schön.

Herr Dr. Schopen (PTC Therapeutics): Vielen Dank, Herr Professor Hecken. – Meine Damen und Herren, ich möchte ganz kurz noch zwei Worte mehr zu den PTC-Team-Mitgliedern sagen. Ich selber bin Geschäftsführer bei PTC Deutschland und vertrete heute hier die PTC Therapeutics International. Mit mir im PTC-Team ist Herr Dr. Werner gekommen. Herr Dr. Werner ist Leiter unserer Medizin in Deutschland, und dieses einleitende Statement teilen wir uns. Er wird direkt im Anschluss nach mir noch ein paar vertiefende Kommentare zur aktuellen Nutzenbewertung aus seiner medizinischen Sicht geben wollen.

Mit uns ist auch Herr Mark Boulding. Das ist der Herr, der zu meiner Linken sitzt. Herr Mark Boulding ist Chief Legal Officer in unserer Zentrale in den USA, ist aber vor allen Dingen hier, weil er schon sehr lange beim Unternehmen ist. Er ist 15 Jahre beim Unternehmen, er ist im Vorstand des Unternehmens, und die Zulassungsabteilung berichtet auch zu ihm. Wir haben uns gedacht: Fall es Fragen gibt, die etwas den unmittelbaren Kontext überschreiten – auch von der Zulassungstechnischen Seite, vielleicht auf europäischer Ebene –, wäre er ein Ansprechpartner. Mark Boulding spricht nicht Deutsch, und wir haben dieses Mal davon abgesehen, Übersetzer mitzubringen. Ich glaube, wir haben da letztes Mal zwei Herren hinter uns ziemlich frustriert. Deshalb versuchen wir beide hier – Herr Werner und ich –, dann zu übersetzen und ihn nach Bedarf ins Gespräch zu bringen.

Der Vierte im Bunde von uns ist Herr Hofmann. Er hat uns als Externer bei der Datenanalyse und der Erstellung der Dossiers unterstützt.

Ich glaube, zu PTC brauche ich nicht viel zu sagen: kleines Start-up-Unternehmen, weniger als 300 Mitarbeiter. Wir fokussieren auf die Forschung und Entwicklung von innovativen Therapien, schwerpunktmäßig im Bereich seltener Erkrankungen. Das Produkt, das wir heute besprechen, ist das einzige, aber auch das am weitesten entwickelte Produkt, das wir haben. Es ist gezielt für die Erkrankung von zugrundeliegenden Nonsense-Mutationen entwickelt worden und hat eine EMA-Zulassung. Wir entwickeln basierend auf demselben Wirkstoffmechanismus auch in anderen Indikationen wie zystischer Fibrose, Mukopolysaccharidose 1; all das sei nur am Rande erwähnt.

DMD – das brauche ich auch nicht zu sagen – ist eine seltene Erkrankung. Was noch einmal aus unserer Sicht wichtig ist, ist, dass die Subindikation innerhalb der DMD Duchenne, mit der wir es hier zu tun haben, eine sehr seltene Erkrankung ist. Es sind letztendlich nur die Patienten, die älter als fünf Jahre und gehfähig sind, und wir schätzen, dass es 34 bis 42 Patienten in Deutschland sind. Dazu haben wir die Rationale eingereicht; aber ich kann Ihnen auch hands on sagen: Wir haben in zwei Jahren hier jeden Stein in Deutschland mit allen Muskelzentren umgedreht und gesucht und gesucht. Ich glaube, wir sind jetzt schon am Maximum angekommen und sind insofern sehr sicher, dass wir es hier mit einer sehr kleinen gezielten Patientengruppe zu tun haben.

Die Patienten verlieren üblicherweise oder im Durchschnitt mit zehn bis zwölf Jahren die Gehfähigkeit, und insofern ist der Verlust der Gehfähigkeit prognostisch für den weiteren und dann auch unumkehrbaren Krankheitsverlauf; daher ist er der primäre Endpunkt unserer Studien gewesen. Darüber hinaus verspricht der Wirkmechanismus – dazu gibt es sehr gute präklinische Daten –, dass er auch in anderer Muskulatur, die jenseits der Gehfähigkeit zur Wirkung kommt, einen Benefit bringt. Es sei nur am Rande erwähnt, dass letzte Woche bei der World Muscle Society erstmals Daten von uns vorgestellt worden sind, die den Einfluss von Ataluren auf die Lungenfunktion zeigen; wir können das unseren Langzeitstudien entnehmen.

Mir ist sehr wichtig, an dieser Stelle zu sagen: Ataluren haben wir im Dezember 2014 hier in Deutschland im Markt eingeführt. Nach neun Monaten waren alle uns bekannten passenden Patienten auf Therapie. Ich denke, das spricht für den dringenden medizinischen Bedarf. Es spricht vielleicht auch für das Produkt. Es war für uns sicherlich alles andere als einfach, als wir nach Verhandlungen mit dem KGV-SV und nach Schiedsstellenspruch das Produkt außer Vertrieb nehmen mussten. Das war sicherlich eine der schwersten Entscheidungen, die wir treffen mussten. Ich kann dazu, wenn Interesse besteht, etwas mehr sagen. Ich möchte an dieser Stelle nur sagen, dass wir, als wir diesen Beschluss gefasst hatten, aus unserer Sicht sehr sorgfältig mit allen Zentren gesprochen haben und ihnen Zeit gegeben haben, sich darauf vorzubereiten. Es haben dann auch intensive Gespräche der Patienten mit den einzelnen Versicherungen stattgefunden. Ich bin persönlich und, ich glaube, auch im Namen des Unternehmens sehr erleichtert, von den Muskelzentren zu hören, dass alle Patienten

dieses Produkt weiter haben wollten, und es im Augenblick über Importe auch beziehen können. Die einzelnen Versicherungen haben dabei eine sehr konstruktive und pragmatische Rolle gespielt.

Nach vorne blickend kann ich nur sagen: Das ist einer der Hauptgründe, warum wir uns jetzt sehr freuen, wieder hier zu sein und durch diesen Prozess zu gehen, dass wir jede Gelegenheit suchen, dieses Produkt wieder in Deutschland einzuführen. Aber vielleicht kommen wir dazu auch noch einmal separat.

In der aktuellen zweiten Nutzenbewertung können wir jetzt voll umfänglich die gesamte Evidenz präsentieren. In der ersten Nutzenbewertung war es eine Phase-IIb-Studie, jetzt kommt eine Phase-III-Studie hinzu. Wenn man bedenkt, wie klein die Patientenpopulation ist, so denken wir, dass das eine Menge an Evidenz ist, die wir hier zur Verfügung stellen können. Wir hoffen auf Ihre Zustimmung, dass eine Metaanalyse grundsätzlich und vor dem Hintergrund der sehr begrenzten Patientenzahlen die angemessene Methodik ist, zwei unabhängige randomisierte kontrollierte Studien auszuwerten. Wir haben bezugnehmend auf das IQWiG-Methodenpapier die Resultate der Metaanalyse eingereicht. Wir haben signifikante und patientenrelevante Effekte gefunden und demonstriert. In ihrer Gesamtheit und in der kurzen Zusammenfassung hier haben wir signifikante Resultate für den primären Endpunkt präsentiert, den unterstützenden Endpunkt einer dauerhaften 10-prozentigen Verschlechterung der Gehstrecke und drei Zeitfunktionstests.

Wir haben die Nutzenbewertung, die Sie uns gegeben haben, aufmerksam gelesen. Wir haben im Wesentlichen einige Kritikpunkte, und wir haben eine andere Herangehensweise. Unsere Kritikpunkte sind sicherlich, dass die Resultate der Metaanalyse aus unserer Sicht nicht voll umfänglich in der Nutzenbewertung reflektiert worden sind. Der Therapieeffekt in der Gesamtpopulation für den Endpunkt 10-prozentige Verschlechterung wurde nicht berücksichtigt. Analysen zur Patientenrelevanz dieser Endpunkte sind auch nicht dargestellt worden. Des Weiteren sind einige Daten fehlerhaft aus den Studienberichten kopiert worden. Vielleicht können wir heute darüber reden, wenn Sie entsprechende Fragen haben, die Sie uns stellen wollen.

Unsere Herangehensweise ist auch eine andere gewesen. PTC hat die gesamte Studienpopulation einbezogen, der G-BA die sogenannte ADP-Population. Wir sind überzeugt, dass unsere Herangehensweise voll im Einklang mit den Prinzipien der evidenzbasierten Medizin ist und eins zu eins die IQWiG-Methodik reflektiert, und insofern reflektiert es dann auch die von der Zulassung betroffene Patientenpopulation voll umfänglich. Allerdings muss ich, müssen wir, auch sagen: Unabhängig von allen statistischen Überlegungen, ob wir nun die sogenannte G-BA-Metaanalyse oder die PTC-Metaanalyse nehmen, wir denken, dass beide zu fast identischen Aussagen kommen, auch was die vorteilhaften Resultate für Ataluren angeht. Wir interpretieren das so, dass es doch eigentlich nur die Stärke unserer Daten und die Robustheit der Resultate unterstreicht.

Abschließend möchte ich aus meiner Sicht nur noch einmal kurz betonen, dass wir sehr daran interessiert sind, Translarna in Deutschland wieder in den Verkehr bringen zu können. Wir betrachten dieses jetzt angelaufene Verfahren und den sich dann anschließenden Verhandlungsprozess mit dem GKV als Chance. Für den Moment erhoffen wir uns eine Bewertung im Einklang mit den Standards der Arzneimittel-Nutzenverordnung, der Verfahrensordnung des G-BA und der IQWiG-Methodik und sind überzeugt, dass wir in diesem Zusammenhang hier starke Daten eingereicht haben, die einen beträchtlichen Zusatznutzen zeigen werden. – Das ist es einleitend von mir. Herr Dr. Werner würde ganz gerne noch einige ergänzende Anmerkungen zu der Nutzenbewertung hinzufügen. – Vielen Dank zunächst.

Herr Prof. Hecken (Vorsitzender): Vielen Dank, Herr Schopen. – Herr Dr. Werner, bitte.

Herr Dr. Werner (PTC Therapeutica): Vielen Dank. – Sehr geehrter Herr Vorsitzender, sehr geehrte Damen und Herren, ich möchte zunächst ein paar Worte zur Krankheit und zum Wirkmechanismus von Ataluren verlieren und anschließend auf die von Herrn Schopen eingangs erwähnten wesentlichen Aspekte, die wir für die Nutzenbewertung für relevant halten, eingehen. Da ist zum einen das Thema der Metaanalyse, warum wir standardisierte mittlere Unterschiede anstelle der geringsten klinisch relevanten Unterschiede der MCID gewählt haben und abschließend noch einmal auf das durchschnittliche Gewicht der Patienten eingehen, weil das für die Berechnung der Therapiekosten eine wesentliche Rolle spielt.

Wir haben gehört, die Duchenne-Muskeldystrophie, über die wir heute sprechen, ist eine schwerwiegende, lebensbedrohliche genetische Erkrankung, die durch einen genetischen Defekt im Dystrophin-Gen verursacht wird. Bei einer Subgruppe dieser Patienten, und zwar mit einer sogenannten Nonsense-Mutation, kommt es zu einem vorzeitigen Stoppcodon, was zu einer Unterbrechung der Translation der Boten-, der Messenger-RNA, führt, sodass das Protein Dystrophin nicht produziert werden kann. Dystrophin spielt aber im Erhalt der Muskelzellen eine entscheidende Rolle und hat eine stabilisierende Funktion. Beim Fehlen dieses Proteins kommt es zu degenerativen Prozessen, wo die Muskelfasern nach und nach in Binde- bzw. Fettgewebe umgebaut werden. Der Prozess beginnt in der Regel in erster Linie in den unteren Extremitäten und führt nach und nach zu einem Gehverlust im Alter von zehn bis zwölf Jahren. Andere Muskelgruppen des Körpers sind später im Krankheitsverlauf betroffen und führen schließlich dazu, dass die Patienten aufgrund von Atem- oder Herzinsuffizienz versterben.

Die genetische Mutation dieser Erkrankung ist bei Geburt vorhanden. Die ersten auffälligen Symptome zeigen sich in der Regel im Alter von zwei bis drei Jahren, wobei es sich um Gangauffälligkeiten und Probleme, vom Boden aufzustehen, handelt. Weiterhin hat ungefähr ein Drittel dieser Patienten Lernschwierigkeiten und ein vermindertes Intelligenzniveau. Daten aus dem natürlichen, also dem unbehandelten Krankheitsverlauf, zeigen, dass der Verlust der motorischen Funktion und der Gehfähigkeit kein kontinuierlicher, kein linear verlaufender Prozess ist. Innerhalb der ersten Lebensjahre bleibt deshalb die Gehfähigkeit der Patienten nur relativ wenig betroffen. Das hat den Hintergrund, dass in den ersten Jahren das Wachstum und die Entwicklung der Kinder diesen Prozess teilweise kompensieren können. Daher bleibt also die Gehstrecke in den ersten Jahren weitgehend stabil. Wenn jedoch ein gewisses Maß an Muskelgewebe verlorengegangen ist, kommt es zu einer Beeinträchtigung der Gehfähigkeit und wenn – so zeigen es entsprechende Untersuchungen – etwa 80 Prozent des Muskelgewebes durch Binde- oder Fettgewebe ersetzt worden ist, kommt es zu einem raschen Verlust der Gehfähigkeit.

Dieser natürliche Verlauf der Erkrankung, dieser nicht lineare Verlauf, bedingt, dass Studienendpunkte wie der erwähnte Sechs-Minuten-Gehtest nicht in jeder Krankheitsphase gleich sensitiv ist und eine besondere Sensitivität in der Gruppe der Patienten zeigen, die eine abnehmende Gehfähigkeit, aber noch keinen rapiden Verlust der Gehfähigkeit zeigen. Das erklärt auch, warum in den von PTC durchgeführten Studien über den Beobachtungszeitraum von knapp einem Jahr der größte Effekt in der Gruppe der Patienten gezeigt werden konnte, die eine Ausgangsgehstrecke von 300 bis 400 Metern hatten.

Wie funktioniert nun Ataluren? Ataluren wird bei Duchenne-Muskeldystrophie-Patienten mit einer zugrundeliegenden Nonsense-Mutation eingesetzt. Wie gesagt, wird dort durch ein Stoppcodon die Dystrophin-Synthese vorzeitig beendet. In den Studien konnten wir zeigen, dass Ataluren mit dem Ribosom interagiert, die Komponente, die die Translation von der mRNA zum Protein darstellt und ermöglicht, dass das vorzeitige Stoppcodon überlesen werden kann. Dadurch kann ein vollständiges funktionsfähiges Dystrophin-Protein hergestellt werden. Daher sind wir überzeugt, dass Ataluren so

wohl im Rahmen der DMD, aber auch bei anderen Erkrankungen wie der zystischen Fibrose für Patienten mit zugrundeliegender Nonsense-Mutation eine kausale therapeutische Option darstellt.

In der neuen Nutzenbewertung, die wir hier vorgelegt haben, wurden neben der Studie 007 der Phase-IIb-Studie die konfirmatorische Phase-III-Studie mit der Studiennummer 020 vorgestellt. Diese Studie zeigt einen Unterschied zwischen Ataluren und Placebo von 15 Metern im primären Endpunkt und hat damit den primären Endpunkt für den Sechs-Minuten-Gehtest nicht erreicht. Aber es gibt einen Grund, warum dies geschah: Wenn wir uns die Studienpopulation genauer anschauen, sehen wir, dass eine größere als erwartete Zahl von Patienten mit einer langen Ausgangsgehstrecke in die Studie eingeschlossen wurde, bei denen die Gehstrecke über den Beobachtungszeitraum dieser Studie stabil blieb; daher konnte in dem gewählten Beobachtungszeitraum für diese Patienten kein Wirkeffekt dargestellt werden. Aber in der Subpopulation von Patienten, die begonnen haben, die Gehfähigkeit zu verlieren, nämlich der Gruppe der Patienten mit einer Ausgangsgehstrecke von 300 bis 400 Metern, war ein sehr signifikanter Effekt sichtbar, der mit dem in der vorangegangenen Phase-IIb-Studie vergleichbar war. Somit stellt aus unserer Sicht diese Studie eine Bestätigung der Wirksamkeit zur vorangegangenen Studie dar.

In der Phase-III-Studie der Studie 020 zeigte sich für den sekundären Endpunkt des Vier-Treppen-Hinabsteigens – das ist ein sogenannter Time-Function-Test – ein signifikantes Ergebnis. Allerdings sind unserer Auffassung nach die gesamte Wirksamkeit und die gesamte verfügbare Evidenz aus beiden Studien für die Beurteilung der Wirksamkeit zu berücksichtigen. Zu diesem Zweck haben wir eine Metaanalyse durchgeführt, die diese beiden Studien kombiniert und damit die Wirksamkeit auf dem höchsten Evidenzniveau, dem Level 1a, belegt. Diese Analyse haben wir basierend auf den Regeln der evidenzbasierten Medizin und der IQWiG-Methodik durchgeführt und dabei die gesamte Studienpopulation gewählt. Wir wissen, dass im Analyseprotokoll die Metaanalyse für die sogenannte ADP-Population vordefiniert wurde. Jedoch sind wir der Auffassung, dass die gesamte Population die richtige Analysepopulation darstellt. Grund dafür ist, dass die ADP-Population, wenn wir diese wählten, ein methodisches Problem aufwies, das wir durch die Analyse der Gesamtpopulation vermeiden können.

Was sind das für Probleme? Die Randomisierung für die Studie 007, in der wir dann die ADP-Population gewählt hätten, wäre damit verletzt worden. Die Patientenpopulation, die dem Zulassungstext entspricht, wäre in der Metaanalyse nicht vollständig erfasst worden und auch die Subgruppenanalysen wären entsprechend beeinflusst worden. Weiterhin war in der ersten Nutzenbewertung von Ataluren die Wahl der ADP-Population ein Kritikpunkt, da es sich hier um eine Post-hoc-Subgruppe handelt, sodass wir uns hier entschlossen haben, die Gesamtpopulation zu analysieren.

Wenn wir uns diese Metaanalyse wie vorgelegt sowohl im Dossier als auch in der Stellungnahme anschauen, sehen wir, dass in fünf verschiedenen Endpunkten einschließlich des primären Endpunktes signifikante Vorteile im Erhalt der Gehstrecke im Vergleich Ataluren gegenüber Placebo dargelegt werden konnten. Basierend auf dieser Analyse haben wir entsprechend die Subgruppenanalysen durchgeführt und zwar für die, wo ein signifikanter Interaktionseffekt gezeigt werden konnte. Nur für die Untergruppe der Patienten zwischen 300 bis 400 Metern Ausgangsgehstrecke im Sechs-Minuten-Gehtest ist eine Subgruppenanalyse statistisch gerechtfertigt. Wir haben entsprechend die Ergebnisse der Effektmodifikatoren dargestellt, um diese signifikanten Effekte zu erläutern.

Keinen Subgruppeneffekt fanden wir für den Endpunkt der 10-prozentigen Verschlechterung des Sechs-Minuten-Gehtests, sodass für diesen Endpunkt die gesamte Population für die Nutzenbewertung relevant ist. Zusätzliche Analysen, die dies entsprechend belegen, haben wir in der schriftlichen Stellungnahme vorgelegt. Da die untere Grenze des Konfidenzintervalls der Hazard Ratio hier in die-

ser Analyse bei unter 0,95 liegt, kann aus unserer Sicht ein erheblicher Zusatznutzen nach den IQWiG-Methoden beansprucht werden.

Für den primären Endpunkt, den Sechs-Minuten-Gehtest, und die übrigen sekundären Endpunkte, nämlich den Vier-Treppen-Hinauf- und -Herabsteige-Test, sowie den Test eines Zehn-Meter-Rennens bzw. -Laufens, konnten Subgruppeneffekte dargestellt werden. Bewusst haben wir uns hier die vorgeschlagenen MCIDs, also die geringsten klinisch relevanten Unterschiede, nicht für die Interpretation verwendet. Hintergrund war wiederum, dass in der ersten Nutzenbewertung diese MCIDs infrage gestellt worden sind, da es sich hier um verteilungs- und nicht ankerbasierte Werte handelt. Deshalb haben wir für die Analyse der Patientenrelevanz die standardisierten Mittelwertdifferenzen Hedges-g verwendet, um die Patientenrelevanz zu bestimmen. Diese Analyse zeigt, dass in drei der vier Endpunkte ein patientenrelevanter Effekt gezeigt werden kann.

Aus unserer Sicht zwei wesentliche Punkte an der G-BA-Sichtweise auf die Nutzenbewertung: Zum einen vermissen wir die Ergebnisse der Metaanalyse und die Hedges-g-Werte, die sich für die Subpopulationen nicht in der Darstellung wiederfinden, und einige Daten, wie eingangs erwähnt, sind aus dem Studienbericht nicht korrekt wiedergegeben worden. Zum anderen sehen wir die Metaanalyse der Gesamtpopulation als die richtige Analysemethode an. Dennoch zeigen, wie auch schon von Herrn Schopen dargestellt, egal welche Population wir wählen, beide Metaanalysen in die gleiche Richtung und demonstrieren einen signifikanten Effekt zugunsten von Ataluren.

Ein paar Worte zur Sicherheit der Substanz: Die Substanz Ataluren ist insgesamt gut verträglich. Entsprechende Datenanalysen zeigen in der Häufigkeit von unerwünschten Ereignissen in den einzelnen Gruppen keinen signifikanten Unterschied zu Placebo. Zu erwähnen sind als die häufigsten unerwünschte Nebenwirkungen bzw. Ereignisse Nasopharyngitis, Erbrechen, Durchfall, Stürze, Rückenschmerzen, Husten, Fieber und Kopfschmerzen.

Zum Schluss noch ein paar Worte zum Gewicht und zur Gewichtsbeurteilung. Wir haben in der schriftlichen Stellungnahme auch noch einige zusätzliche Daten vorgelegt, die belegen, dass aus unserer Sicht das Gewicht der Patienten unter 33 Kilogramm liegt. Wir kommen aufgrund der verfügbaren Daten zu einem Gewichtsband von 27 bis 31 Kilogramm für die Patienten, was, wie eingangs gesagt, für die Berechnung der Therapiekosten eine wesentliche Rolle spielt. Dies wird zum einen belegt durch die Studiendaten der beiden erwähnten Studien, die insgesamt 342 Patienten umfassen. Betrachten wir das durchschnittliche Gewicht, kommen wir in diesen Studien auf 30,9 Kilogramm, betrachten wir den Median, auf 27,1 Kilogramm. Aus unserer Sicht scheint, da die Werte nicht normal verteilt sind, der Median hier die aussagekräftigere Größe. Zum anderen gilt zu berücksichtigen, dass in einer der beiden Studien, nämlich in der Studie 020, Patienten im Alter von sieben bis 16 Jahren eingeschlossen worden sind und damit die zugelassene Altersgruppe von fünf bis sechs Jahren hier unterrepräsentiert ist, was dann dazu führen würde, dass das Gewicht geringer als in der Studie läge.

Einen weiteren Beleg für dieses Gewichtsband bieten die Apothekenbestellungen. Aufgrund der Seltenheit der Erkrankung steht in der Regel hinter einer Apothekenbestellung ein Patient, sodass wir aus den Bestellmustern der Apotheken sowohl durch externe Analyse über INSIGHT Health als durch interne Analysen zu einer durchschnittlichen Tagesdosierung pro Patienten, das entsprechend umgerechnet zu einem Gewichtsband, wiederum von 27 bis 31 Kilogramm, kommen. Von daher sind wir der Auffassung, dass die IQWiG-Berechnung des durchschnittlichen Gewichts basierend auf den gesunden altersgleichen Patienten ein zu hohes Gewicht für Duchenne-Muskeldystrophie-Patienten ergibt.

Abschließend hoffen wir, dass wir die Daten und die Metaanalysen, die wir hier im Dossier präsentiert haben, hinreichend erklären konnten. Die signifikanten Vorteile, die Ataluren über mehrere Eckpunkte

hinweg bietet, belegen die Wirksamkeit von Ataluren und demonstrieren einen patientenrelevanten Zusatznutzen. Die Verlangsamung der Abnahme der Gehstrecke verlängert die Phase der Gehfähigkeit für die Patienten, was den Krankheitsverlauf im erheblichen Maße positiv beeinflusst. – Vielen Dank.

Herr Prof. Hecken (Vorsitzender): Herzlichen Dank, Herr Dr. Werner. – Das Gewicht gucken wir uns dann noch einmal an, damit wir das abgearbeitet haben. – Frau Mostardt, erste Frage.

Frau Dr. Mostardt: Ich möchte dazu direkt etwas sagen. Bis auf die Zulassungsstudien gibt es eigentlich keinen Nachweis, wo Sie wirklich nachweisen können, dass das Gewicht geringer ist, weil die neuere Quelle, die Sie im Gegensatz zu dem alten Verfahren hinzugefügt haben, diese NICE-Guideline ist. Aber in der NICE-Guideline ist es so, dass man ein durchschnittliches Gewicht von Achtjährigen angenommen hat und die dann über fünf Jahre beobachtet hat. Und wenn man sich dann das Gewicht anguckt, was dann zum Beispiel die Neunjährigen mit 39 Kilogramm und für die 13-Jährigen 53 Kilogramm, das entspricht dann genau den Mikrozensusdaten, die wir aufgenommen haben. Deswegen ist es eigentlich auch wieder kein Beleg dafür, dass dort ein geringeres Gewicht vorhanden ist.

Wenn man sich in der Best-Studie die Tabelle 1 anguckt, sieht man auch, dass die Gewichtsangaben für die einzelnen Altersgruppen dem Mikrozensus ungefähr entsprechen. Es gibt eigentlich bis auf die Zulassungsstudie keinen Nachweis für das geringere Gewicht.

Zu den INSIGHT-Health-Daten ist es so – das ist aber noch einmal eine generelle Anmerkung –: Sie haben dazu keine Quelle geliefert. Sie haben eine reine Angabe in dem Dossier gemacht, aber es fehlt sozusagen, wie man die Patienten überhaupt identifiziert hat. Wie ist die Auswertung gelaufen? Dazu finden wir überhaupt keine Angaben im Dossier. Deswegen können wir das auch nicht bewerten.

Herr Prof. Hecken (Vorsitzender): Danke schön. – Herr Schopen.

Herr Dr. Schopen (PTC Therapeutics): Die INSIGHT-Health-Daten sind Daten, die wir natürlich auf vertraulicher Basis zur Verfügung stellen können. Sie sind reale aktuelle Daten, keine theoretischen Überlegungen; das ist das, was wir sehen. Wir haben auch kein Problem, sie zur Verfügung zu stellen, aber dann auf einer vertraulichen Basis. Aber die aktuellen Daten zeigen das ziemlich eindeutig. Wir können das natürlich auch im Tagesgeschäft sehen. Wir wissen, wie viele Patienten wir haben und was an Umsätzen läuft; das können Sie zurückrechnen. Es ist sogar so, dass wir mittlerweile am unteren Band der 27 bis 31 Kilogramm angekommen sind. Es ergibt sich zum Teil einfach dadurch, dass wir zum Ende des Jahres 2015, Anfang des Jahres 2016 fast schon die Maximalzahl der Patienten hatten, weil der Bedarf so hoch war. Daran hat sich im Wesentlichen im Laufe des Jahres nicht viel geändert. Das heißt nicht, dass hier und da ein paar Patienten weggefallen sind und ein paar neue Patienten dazugekommen sind, und dadurch haben wir zunehmenden Shift zu jüngeren Patienten bekommen, die noch leichter im Gewicht sind. Wir können sicherlich noch einmal gucken, ob wir hinsichtlich dieser Datenbasis einen Ausgleich finden; aber es ist Real Life. – Das ist alles, was wir im Augenblick dazu beitragen können.

Herr Prof. Hecken (Vorsitzender): Frau Hornkamp.

Frau Hornkamp: Ich würde das schon bestätigen, dass unsere Kinder leichter sind. Sie sind kleinwüchsig. Abgesehen davon haben sie weniger Muskeln. Die Muskelmasse nimmt ab. Von daher kann ich es bestätigen. Wir fragen auch regelmäßig unsere Eltern ab und lassen uns auch da das Gewicht

zurückmelden. Definitiv ist es halt so, dass die Kinder leichter als gesunde Kinder in dem Alter sind. Dazu können Sie gerne auch noch einmal auf mich zukommen. Wir haben ja auch letztendlich – –

Herr Prof. Hecken (Vorsitzender): Wobei das natürlich die belastbare Evidenz zu ersetzen nicht imstande ist. Aber ich sagte ja: Wir werden uns damit beschäftigen müssen. – Herr Kulig, bitte. Dann Herr Dr. Heckemann.

Herr Kulig: Ich habe zwei Fragen, die sich auf die Definition der Subgruppen, für die unterschiedliche Effekte auftreten, beziehen. Das ist eine Frage, die ich zuerst gar nicht hatte, die sich aber auf Ihr eben abgegebenes Statement referenziert. Sie sagten, die Sensitivitätstests zur Messung der Gehstrecke sei für die Population, die eine mittlere Beeinträchtigung hat, also die ADP-Population, – am sensitivsten; für die anderen Gruppen sei das nicht so sensitiv. Was heißt denn das, dass dieser Test, diese Endpunktmessung, für die anderen Gruppen gar nicht geeignet ist, weil er nicht sensitiv genug ist? Wäre das dann gar kein geeignetes Messinstrument für diese Gruppen? Oder heißt es, dass dann dort gar kein Therapieeffekt vorhanden ist? Ich war jetzt ein bisschen über die Aussage irritiert, die Sensitivität des Tests für die Gruppe ist dort am besten.

Herr Dr. Werner (PTC Therapeutica): Ja, es liegt hier tatsächlich an der Sensitivität des Tests, gerade auch für die Betrachtung des gewählten Zeitraums. Ich hatte versucht, dazustellen, dass der natürliche Verlauf der Erkrankung zunächst mit einer über einen längeren Beobachtungszeitraum hinweg relativ stabilen Gehstrecke einhergeht, die dann in einer späteren Krankheitsphase abnimmt und dann kurz vor Verlust der Gehfähigkeit doch relativ rapide abgeht, sodass die Gehstrecke abnimmt. Das hatten wir auch in den Unterlagen in einer Studie entsprechend dargestellt, wie sich der natürliche Verlauf der Erkrankung zeigt. Das heißt, in einer Patientenpopulation mit einer Ausgangsgehstrecke von 300 bis 400 Metern haben wir die Möglichkeit, über einen Behandlungszeitraum von, wie in den Studien gewählt, 48 Wochen einen Effekt zu zeigen. Patienten, die eine Ausgangsgehstrecke von 400 Metern und mehr haben, verlieren auch im natürlichen Verlauf über diese Zeit nur sehr wenig an Gehstrecke, sodass hier kein Effekt eines therapeutischen Wirkstoffs dargestellt werden kann. Würden wir eine längere Studiendauer wählen, würde man auch in dieser Patientengruppe möglicherweise einen Effekt sehen. – Haben Sie ergänzend noch etwas?

Herr Dr. Hofmann (PTC Therapeutics): Vielleicht ergänzend, Herr Kulig. Wenn man sich Patienten ansieht, die sehr jung sind – im Alter zwischen fünf bis sieben Jahren – und diagnostische Methodik verwendet, in der man die Zerstörung des Muskelgewebes messen kann, also Magnetresonanztomografie, dann kann man erkennen, dass das Muskelgewebe schon sehr früh geschädigt wird, wie erwähnt im Alter zwischen fünf und sieben Jahren. Studien von [Lynn *nachträglich ergänzt*], die wir auch im Nutzendossier präsentiert haben, zeigen Folgendes: Wenn man die Kinder in diesem Alter in einen Sechs-Minuten-Gehtest mit diesen relativ geringen Veränderungen – das können bis zu 30, 40 Prozent des Gewebes von gewissen Muskeln sein – schickt, dann wird man im Sechs-Minuten-Gehtest keinen Unterschied messen können. Mit anderen Worten: Wenn Patienten in dieser Phase behandelt werden, wird man auch über den Sechs-Minuten-Gehtest keine Wirksamkeit nachweisen können. Das geht genau zu dem Punkt, den Sie in Ihrer Frage zu Patienten angesprochen haben, die noch sehr früh in der Krankheitsprogression sind – um die 400 Meter Baseline-Gehstrecke –, bei denen wird man mit dem Sechs-Minuten-Gehtest weder eine Krankheitsprogression noch eine Wirksamkeit von einem neuen Arzneimittel messen können. Unsere Voraussagen gehen in die Richtung, dass man wahrscheinlich in zukünftigen Studien keine Patientenpopulationen mehr einschließt, die in diesem frühen Bereich der Krankheitsprogression zu finden sind. Da sind auch andere Unternehmen, die dieselben Ergebnisse finden; das ist nicht nur PTC.

Herr Prof. Hecken (Vorsitzender): Ich habe jetzt konkret dazu Fragen von Herrn Wilken. Anschließend Herr Kulig, dann Herr Heckemann.

Herr Dr. Wilken (BPI): Die Frage bezieht sich auf die von Ihnen angesprochenen Effekte, die eigentlich in der Gruppe der Ambulatory Decline Phase nur darzustellen sind und das neben der nicht eingeschränkten Indikation des Arzneimittels. Können Sie dies belegen?

Herr Dr. Hofmann (PTC Therapeutics): Bevor ich die Frage beantworte, Herr Wilken: Ich glaube, wir sollten vorab einmal definieren, über welche Subgruppen wir jetzt eigentlich sprechen. Wir sprechen im Nutzendossier und auch in anderen Kommunikationen nur über eine Subgruppe oder ein Subgruppengebilde, nämlich die, die sich durch die 300- bis 400-Meter-Subgruppenmerkmale bestimmen, und die Patienten, die darüber liegen oder darunter. Das ist die einzige Subgruppe, für die eine statistisch berechnete Möglichkeit einer Analyse vorliegt. Die ADP-Subgruppe ist unseres Wissens keine demonstrierte Subgruppe. Wir haben hier keinen signifikanten Interaktionstest mit der komplementären Subgruppe, und das stimmt auch so für alle anderen Subgruppen, die wir im Rahmen der Phase-III-Studie analysiert hatten. Wir sprechen in der Regel immer nur von der Subgruppe, die sich durch diese drei Subgruppen 300 bis 400, über 400 und unter 300 charakterisiert.

Herr Prof. Hecken (Vorsitzender): Herr Kulig.

Herr Kulig: Das war auch meine zweite Frage. Wir fanden es jetzt in der Bewertung schwierig, wie sich diese unterschiedlichen Subgruppen im Laufe des Studienprogramms gebildet haben, nach welchen Kriterien und auch, wann das definiert wurde. Die Einschlusskriterien dieser 020-er-Studie haben sich, wenn wir das richtig verstanden haben, aus den Ergebnissen der vorherigen Studie 007 gebildet. Dort hat man ja gesagt: Okay, die Patienten müssen mindestens eine Gehstrecke von 150 Metern oder mindestens 80 Prozent der vorhergesagten Gehfähigkeit oder Gehstrecke haben. Die Frage ist, wie weit es da schon Unterschiede, was wir auch in der Nutzenbewertung adressiert haben, zwischen den Studien gibt, aber vor allen Dingen geht es um die Bildung der Subgruppen. Wenn Sie sagen, Sie heben hauptsächlich oder viel auf die Kleinphase-Patienten zwischen 300 und 400 Metern ab, dann stelle ich mir die Frage, warum Sie aber eine stratifizierte Randomisierung aufgrund von plus/minus 350 Meter gemacht haben. Man führt eine stratifizierte Randomisierung eigentlich durch, wenn man schon das Gefühl hat: Okay, da gibt es gewisse Subgruppen, wo man unterschiedliche Effekte hat, und deswegen möchte man auch innerhalb dieser Subgruppen randomisierte balancierte Patienten haben. Warum haben Sie dann nicht in der Studie schon diese Gruppe als Stratifizierungsfaktor gewählt? Wann und wie das im Studienprogramm ablief, ist eigentlich nicht so recht nachvollziehbar gewesen. Wollen Sie den Effekt für die Metaanalyse eigentlich für die Gesamtgruppe nachweisen oder wirklich nur für die ADP-Gruppe? Das ist jetzt auch noch nicht so ganz klar, Das wäre auch ein bisschen widersprüchlich: einmal für die Gesamtgruppe und schon allein aufgrund der stratifizierten Randomisierung aufgrund dieser Einschlusskriterienwahl dann doch nur für diese Subgruppe. Können Sie mir einfach noch einmal ein bisschen ausführen, wie das zustande kam?

Herr Prof. Hecken (Vorsitzender): Wer macht das? – Herr Hofmann.

Herr Dr. Hofmann (PTC Therapeutics): Das klinische Entwicklungsprogramm war von Beginn an getrieben von der Fragestellung: Können in der klinischen Studie, in der eine relativ breite Patientengruppe eingeschlossen wird – – Der Grund hierfür sind unter anderem die Anforderungen von der FDA. Kann man an allen Patienten einen Therapieeffekt messen? Aufgrund der heterogenen Studienlage war das von vornherein schon infrage gestellt. In der ersten Studie war das Verständnis der his-

torischen Entwicklung noch relativ früh und, wie Sie richtig feststellen, gab es dort eine Stratifizierung mit einem Cut-off bei der Baseline-Gehstrecke. Das ist ein wichtiger prognostischer Faktor, nicht nur prognostisch, auch prädiktiv. Hier hat man versucht, in relativ frühe Populationen oder späte Populationen einzuteilen, also ein relativ simpler Ansatz.

Als die Phase-III-Studie geplant wurde, wurde aus dem Grund der Reproduzierbarkeit auch dasselbe Stratifizierungsmerkmal gewählt. Die Erfahrungen haben sich aber relativ schnell weiterentwickelt, unter anderem durch die Interaktion mit dem Herrn Liswini (?), den ich schon erwähnt hatte. Aus diesem Grund wurde dann auch ein anderes Paar von Cut-offs im Studienprotokoll definiert, um eine Patientenpopulation zu definieren, für die die Tests, die eingesetzt werden, auch sensitiv sind. Die 300- bis 400-Subgruppe findet sich auch im SAP; er war auch prädefiniert. Die Studiendaten von zwei unabhängigen Studien, die wir vorlegen, zeigen mehr als deutlich, dass das Grenzwerte sind, die eine Population relativ gut beschreiben, in der man auch einen Therapieeffekt messen kann.

Wir führen Metaanalysen, Heterogenitätstests zwischen diesen Subgruppen durch. Die Homogenität zwischen diesen Subgruppen liegt bei null, wenn man den i -Quadrat-Wert annimmt, und zwar nicht nur für einen Endpunkt, sondern für fünf. Das Bild ist also extrem homogen, ich glaube, der Studienpool, den wir vorlegen, zeigt extrem überzeugend, dass wir hier bei einer Definition der Studienpopulation sind, in der man tatsächlich und wirklich einen Therapieeffekt nachweisen kann.

Zum zweiten Teil Ihrer Frage: Können wir einen Therapieeffekt in der Gesamtpopulation nachweisen? Die Antwort ist ganz klar Ja. Wir haben das auch so im Nutzendossier dargestellt, und zwar für den Endpunkt der 10-prozentigen dauerhaften Verschlechterung im Sechs-Minuten-Gehtest. Hier finden wir keine Subgruppeneffekte, aber wir finden über die beiden pivotalen Studien in der Metaanalyse einen signifikanten patientenrelevanten Endpunkt.

Herr Prof. Hecken (Vorsitzender): Herr Kulig.

Herr Kulig: Nachfrage. Wann genau wurde dann das geändert bzw. wann genau wurde diese ADP-Population definiert, dass das die wichtige Zielpopulation ist?

Herr Dr. Hofmann (PTC Therapeutics): Für die Nutzenbewertung?

Herr Kulig: Ja.

Herr Dr. Hofmann (PTC Therapeutics): Ich glaube, wir haben da noch immer ein Missverständnis. Die ADP-Population kommt aus einer Definition von der ersten klinischen Studie. Man hat aber gelernt, dass die ADP-Population wahrscheinlich nicht die Population ist, in der man Therapieeffekte am besten messen kann. Aus dem Grund haben wir uns auf die 300 bis 400 und die entsprechenden Komplementären fokussiert.

Herr Kulig: Gut, dann war das ein Missverständnis. Ich meinte wirklich diese Gruppe 300 bis 400 Meter Gehstrecke zu Baseline. Wann wurde sie als Zielpopulation der Auswertung definiert? Sie haben gesagt: prädefiniert im Analyseplan.

Herr Dr. Hofmann (PTC Therapeutics): Sie wurde vor Datalog definiert. Wann der genaue Zeitpunkt war – –

Herr Dr. Schopen (PTC Therapeutics): Entschuldigung, die 300- bis 400-Meter-Gruppe wurde aus den National History-Studien identifiziert, während die Studie lief, und bevor die Studie entblindet

wurde, wurde das statistische Protokoll entsprechend adaptiert und das Auswertkriterium hineingenommen. Das ist das, was Herr Boulding gerade dazu beiträgt.

Herr Kulig: Wenn Sie darauf hingewiesen haben, dass es da auch Interaktionen gibt oder dass das eine Subgruppe ist, wieso sollte dann für die Metaanalyse die Gesamtpopulation für einen therapeutischen Effekt für den gleichen Endpunkt herangezogen werden? Einmal sind es Absolut-Werte, da nimmt man dann die Subgruppe 300 bis 400 Meter zu Baseline; aber bei der 10-prozentigen Verschlechterung – das ist eine andere Operationalisierung des gleichen Endpunktes der Gehstrecke – wird dann die Gesamtpopulation herangezogen. Vielleicht würde bei einer 20-prozentigen oder 5-prozentigen Verschlechterung – wir wissen auch gar nicht, ob diese 10-prozentige Verschlechterung der patientenrelevante Cut-off ist – auch ein anderes Ergebnis herauskommen. Einmal Gesamtpopulation, einmal diese spezielle Population, das ist mir von der Argumentation her noch nicht einleuchtend.

Herr Prof. Hecken (Vorsitzender): Herr Hofmann.

Herr Dr. Hofmann (PTC Therapeutics Limited): Okay. Die Frage ist: Wie wertet man Responderanalysen aus? Es gibt da zwei Effekte, die man berücksichtigen muss. Ich gehe da noch einmal ganz kurz auf die Sechs-Minuten-Gehstrecke zurück. Hier haben wir einen Endpunktparameter, der eine absolute Veränderung misst, also eine Veränderung in der Gehstrecke oder bei den Time-Function-Tests eine absolute Veränderung in der Zeit, die benötigt wird, einen Test durchzuführen. Bei dem Endpunkt der 10-prozentigen [Verschlechterung der 6-Minuten-Gehstrecke *nachträglich ergänzt*] – Ich bleibe noch einmal ganz kurz bei dem Sechs-Minuten-Gehtest. Der Therapieeffekt ist von einem Baseline-Parameter abhängig, nämlich der Baseline Sechs-Minuten-Gehstrecke. Wenn eine absolute Veränderung von dem Baseline-Wert abhängig ist, dann würde man in dieser Population Subgruppen erwarten. Die sehen wir auch, die haben wir auch beschrieben. Das sind die 300 bis 400 und die entsprechenden anderen Subgruppen. Bei der 10-prozentigen Veränderung haben wir einen Endpunktparameter, wo wir eine relative Veränderung messen. Auch diese ist formell abhängig von dem wichtigen Baseline-Parameter der initialen Sechs-Minuten-Gehstrecke. Es ist aber sehr gut bekannt, sehr weit beschrieben und auch im IQWiG-Methodenpapier explizit hervorgehoben, dass die Responder-Analysen, die auf relativen Messungen beruhen, robust sind. Mit anderen Worten: dass die Subgruppeneffekte, die bei Absolut-Messungen da sind, bei relativen Messungen verschwinden. Wir sehen also genau, was man in der Theorie erwarten würde: Die Subgruppeneffekte verschwinden, und wir sehen für die Gesamtpopulation einen signifikanten Effekt für die 10-prozentige Veränderung.

Herr Prof. Hecken (Vorsitzender): Nachfrage, Herr Kulig.

Herr Kulig: Ganz schlüssig ist mir das noch nicht, auch aufgrund Ihrer eingangs gemachten Ausführungen. Der Sechs-Minuten-Gehtest ist für die Gruppe derer, denen es schon sehr schlecht geht oder die am Anfang noch eine relativ gute Gehstrecke zurücklegen können, sehr wenig sensitiv. Wenn ich jetzt, was weiß ich, Baseline 450 Meter laufen kann, dann muss ich bei einer 10-prozentigen Veränderung laut Ihren vorherigen Ausführungen schon mindestens 45 Meter Veränderung zeigen. Dann ist der Test kaum sensitiv und kann das gar nicht erfassen, auch keine Therapieeffekte. Mir ist die Argumentation noch nicht klar, warum ich jetzt plötzlich alle Gruppen in einen Pott schmeißen darf, einmal so salopp gesagt, warum ich da plötzlich den Therapieeffekt zeigen kann, während eingangs gesagt worden ist, dass man es gar nicht messen kann.

Herr Dr. Hofmann (PTC Therapeutics): Ich glaube, ich verstehe Ihre Fragestellung, Herr Kulig. Die Frage ist, davon gehe ich aus, wenn man auf die absoluten Veränderungen sieht: Man sieht in den Populationen keine statistisch signifikanten Effekte in der unter 300- und über 400-Gruppe. Warum sieht man dann Effekte in der 10-prozentigen Veränderung. Richtig? War das Ihre Frage?

Herr Kulig: In den Subgruppen, über 400 Meter einzeln, sehe ich die auch nicht. Und jetzt sehe ich alle zusammen und damit ergibt sich eben ein Effekt. Aber wie eingangs ausgeführt habe ich starke Zweifel, die ganze Population betrachten zu können.

Herr Dr. Hofmann (PTC Therapeutics): Genau. Nehmen wir das Beispiel der über 400-Meter-Subgruppe und gucken auf die gesamte Population der über 400-Meter-Population, dann liegt der Punktschätzer sehr knapp bei der Null, er zeigt einen nicht signifikanten Effekt. Es sind aber nicht alle Patienten in dieser Population, die genau auf dem Punktschätzer liegen. Wir gucken ja auf eine Verteilung, die durch das Konfidenzintervall repräsentiert ist, und in dem Konfidenzintervall gibt es auch Patienten, die noch einen Respons zeigen. Das sind natürlich weniger als die Gesamtpopulation, aber es sind in der Population noch Responder vorhanden, genauso bei der unter 300-Subgruppe. Bei einer Responder-Analyse – das machen Sie jeden Tag – pickt man sich die Responder explizit heraus und lässt alle anderen links liegen und weiß gar nicht, was mit den anderen passiert. Aber es ist bei Responder-Analysen so. Wenn wir Responder-Analysen durchführen, entweder für die 300 bis 400 oder für die unter 300 oder für die über 400, sind in jeder dieser Populationen Responder, und zwar signifikant mehr als in der Placebo-Gruppe. Aus diesem Grund sehen wir auch in der Gesamtpopulation einen statistisch signifikanten Effekt und einen patientenrelevanten Effekt. Das ist eigentlich recht einfach. Im IQWiG-Methodenpapier ist das explizit auch so beschrieben.

Herr Prof. Hecken (Vorsitzender): Frau Grell, Sie haben dazu eine konkrete Frage. Anschließend Herr Heckemann und Herr van Poppel.

Frau Dr. Grell: Ich möchte zu dem, was Herr Kulig gesagt hat, gerne noch eine Verständnisfrage nachschieben. Sie haben als Einschlusskriterium, dass Sie an zwei Tagen den Sechs-Minuten-Gehtest zu Baseline bestimmen und er nicht mehr als 20 Prozent auseinanderliegen kann. Das heißt, Sie gehen bei diesem Test von einer Varianz von 20 Prozent aus. Jetzt sagen Sie: Die Verschlechterung darf aber nur 10 Prozent betragen. Wenn ich aber bei diesem Test bereits von einer Varianz von 20 Prozent ausgehe, dann verstehe ich die 10 Prozent als Marker einfach nicht.

Herr Dr. Hofmann (PTC Therapeutics): Die 10 Prozent sind für jede Gruppe unterschiedlich. Wir haben bei der über 400-Meter-Subgruppe 40 Meter Veränderung, die wir als Response-Kriterium erfragen.

Frau Dr. Grell: Sie hatten eben ausgeführt, dass Sie insbesondere die 10-prozentige Verschlechterung betrachtet haben. Die Frage ist: Woher kommen 10 Prozent? Warum haben Sie nicht 30 Prozent genommen? Dann liegen Sie deutlich über den von Ihnen als Abweichung sowieso bereits akzeptierten 20 Prozent. In 10 Prozent liegen Sie deutlich unter dem, was Sie als Varianz per Einschlusskriterium schon vorab definiert hatten.

Herr Prof. Hecken (Vorsitzender): Herr Werner.

Herr Dr. Werner (PTC Therapeutics): Die 20 Prozent versuchen auszuschließen, dass man eine zu große tagesformabhängige Schwankung an dieser Stelle zeigt, wohingegen das andere, die 10-prozentige Verschlechterung, ein Effekt über einen längeren Zeitraum ist, der dargestellt wird.

Herr Prof. Hecken (Vorsitzender): Frau Grell.

Frau Dr. Grell: Nun wissen wir ja alle, die wir diese Tests durchgeführt haben, dass das sehr tagesformabhängig ist.

Herr Dr. Werner (PTC Therapeutics): Richtig.

Frau Dr. Grell: Sie haben bei der Ergebniserhebung, wenn ich das richtig verstanden habe, auch nicht an zwei unterschiedlichen Tagen gemessen, sondern Sie haben nur an einem Tag gemessen. Da passt es für mich einfach nicht zusammen, wenn Sie davon ausgehen, dass Sie sowieso tagesformabhängig eine 20-prozentige Abweichung haben, dass Sie jetzt sagen: Aber nur die 10-prozentige Verschlechterung ist patientenrelevant. Ich verstehe das einfach nicht.

Herr Prof. Hecken (Vorsitzender): Herr Werner.

Herr Dr. Werner (PTC Therapeutics): Entschuldigung! Die 20-prozentige Variabilität zwischen zwei Messpunkten bezog sich insbesondere auf die Baseline-Erhebung, um sicherzustellen, dass wir hier einen verlässlichen Baseline-Wert erheben können, wohingegen die 10-prozentige Verschlechterung dann ja auch über einen Zeitraum von 48 Wochen durchaus ein in anderen Studien akzeptierter Endpunkt ist.

Herr Prof. Hecken (Vorsitzender): Okay. – Herr Dr. Heckemann.

Herr Dr. Heckemann: Es ging die ganze Zeit um das Thema, zu dem ich auch fragen wollte.

Herr Prof. Hecken (Vorsitzender): Entschuldigung. Dann fragen Sie jetzt.

Herr Dr. Heckemann: Macht nichts. Man ist ja hier in der Zeit nicht dümmer geworden. Ich hatte vorhin gehört: 48 Wochen und die Zeit ist zu kurz, weil es dort in der Gehstrecke zwischen 620 und 400 langsamer geht; dazu reichen 48 Wochen nicht. Das kann ich nachvollziehen und sagen: Okay, wenn man drei Jahre nimmt, gibt es den Effekt. Es gibt doch keine Rationale dafür, warum das Medikament nur dann wirkt, wenn schon Dreiviertel der Muskulatur kaputt oder weg ist, sondern es müsste doch eigentlich von Anfang an wirken. Zumindest kann die Begründung dafür, dass es über 400 Meter nicht nachweisbar ist, nicht die gleiche sein, wie wenn es unter 300 Meter nicht nachweisbar ist. Das müssten dann ja verschiedene Effekte sein. Bei über 400 Metern – das kann ich mir vorstellen – ist die Zeit, in der beobachtet wurde, einfach zu kurz. Bei unter 300 Metern kann man vielleicht sagen, ist die Krankheit nicht mehr aufzuhalten, egal, was ich tue. Dann geht das von alleine weiter. Aber die eigentliche Frage wäre doch: Müsste man nicht im Prinzip am Anfang, wo absolut die meiste Muskulatur kaputt geht, eigentlich schon therapieren, und müsste man nicht insofern eine längere Studie für die leichteren oder früheren Stadien haben?

Die andere Frage ist: Würde beim Erreichen einer Untertherapie die Gehstrecke trotzdem abnehmen, egal wie stark, dann müsste man die Entscheidung treffen: Bei 300 Metern höre ich auf, dann gebe ich das nicht mehr. Würde man diese Entscheidung so treffen? Diese beiden Fragen habe ich.

Herr Dr. Werner (PTC Therapeutics): Wenn wir uns die präklinischen Daten anschauen, die eine entsprechende Förderung der Dystrophin-Produktion zeigen, so ist das auch von den verschiedenen Krankheitsstadien unabhängig und genau, wie Sie sagen, würde man davon ausgehen, dass ein solcher Therapieansatz möglichst frühzeitig begonnen werden sollte, um einen Verlust des Muskelgewebes aufzuhalten und zu verlangsamen. Wie Sie völlig richtig sagen, würde man das erwarten. Da werden uns hoffentlich die entsprechenden offenen Landzeitstudien, die in Kürze abgeschlossen und

ausgewertet werden, einen Hinweis geben, inwieweit das hier bestätigt werden kann. Aber für diesen Beobachtungszeitraum ist dieser Messparameter in dieser Patientenpopulation von 300 bis 400 Metern besonders sensitiv.

Herr Prof. Hecken (Vorsitzender): Herr Heckemann, befriedigt Sie das?

Herr Dr. Heckemann: Ob der Arzt das Medikament unter 300 Metern absetzen sollte, das wäre schon noch die zweite Frage gewesen.

Herr Dr. Werner (PTC Therapeutics): Danke für die Erinnerung, dass ich darauf nicht eingegangen bin. – Es ist so, dass dieser Effekt der Dystrophin-Produktion durchaus auch in anderen Muskelgewebsgruppen – ich nehme als Beispiel die Atemmuskulatur – gezeigt werden kann. Die von Herrn Schopen angesprochenen Ergebnisse zum pulmonalen Funktionserhalt zeigen das also auch, wenn wir jenseits der Gehstrecke sogar bei Patienten, die die Gehfähigkeit vollständig verloren haben, in dieser Betrachtung einen entsprechenden Effekt sehen, dass der pulmonale Funktionserhalt in der Studie, von der ich spreche, als FEV1 gemessen, der FEV1-Verlust um über 10 Prozent gegenüber nicht mit Ataluren behandelten Patienten reduziert werden kann, wobei es sich hier auch um eine offene Studie und eine historische Vergleichskohorte handelt. Das sind keine unter doppelblinden Bedingungen erhobenen Daten, die aber sehr wohl einen Hinweis darauf liefern, dass auch bei geringeren Gehstrecken hier entsprechende Effekte, die für den Patienten relevant sind, da sind.

Herr Prof. Hecken (Vorsitzender): Dann frage ich einmal: Sind die Daten zu diesen pulmonalen Effekten im Dossier enthalten oder haben Sie sie nachgeliefert, oder wo haben Sie sie? Wir haben uns ja bislang auf die Gehstrecke fokussiert. Darauf bezog sich auch das, was Herr Dr. Heckemann gesagt hat. Die Frage ist: Was haben Sie wahrgenommen und was haben Sie an Evidenz, was wir hier zur Bewertung haben? Sie haben das eben schon vorgetragen und ich habe jetzt gerade einmal gefragt. Es kann ja auch sein, dass man nicht alles liest. Aber ich habe diese Daten nicht gefunden, die hier jetzt bei dieser Gruppe der unterhalb von 300 Meter liegenden Gehstrecke dann möglicherweise aufgrund anderer Effekte positiv bewertet werden können.

Herr Dr. Werner (PTC Therapeutics): Diese Daten stammen nicht aus der Studie 020, sondern aus einer offenen Extensionsstudie und sind in der vergangenen Woche erst publiziert worden. Deshalb sind sie noch nicht im Dossier enthalten.

Herr Prof. Hecken (Vorsitzender): Okay. – Herr van Poppel und Frau Grell.

Herr van Poppel: Ich habe noch eine Frage bezüglich des Zulassungsverfahrens: Die EMA hat ihre Bewertung bezüglich Studie 020 noch nicht abgeschlossen. Anscheinend erwartet die CHMP auch Schwierigkeiten bei der Interpretation dieser Studie. Aus den öffentlich verfügbaren Unterlagen des CHMP lese ich heraus, dass das CHMP eine Scientific Advisory Group zu Fragen im Zusammenhang mit dieser Studie einberufen hat. Können Sie uns sagen, über welche Fragen Sie aktuell mit dem CHMP diskutieren oder sprechen? Wäre es möglich, die List of Questions der SAG und weitere Unterlagen für den G-BA zur Verfügung zu stellen?

Herr Dr. Schopen (PTC Therapeutics Limited): Lassen Sie mich gerade übersetzen. Das ist eine Frage für Mark Boulding.

In der Tat sind wir in einem intensiven Dialog mit den Behörden, der EMA und dem CHMP, das ist ein Grund, warum unser CEO heute nicht hier ist; er ist gerade in England. Das ist ein voranschreitender

Prozess. Eine der großen Fragen, die im Augenblick primär diskutiert werden – und auch das CHMP hat gelernt, was wir in der Zwischenzeit gelernt haben, und bringt sich da sehr aktiv ein –, dass vieles über den natürlichen Krankheitsverlauf der Erkrankung nicht bekannt war, bevor Studien gemacht worden sind, und wir lernen vieles jetzt im Nachhinein. Sie bringt sich im Augenblick sehr aktiv in die Diskussion ein, und das ist der Kernpunkt der Fragen, die im Augenblick sind. Soweit kann Mark Boulding dazu etwas öffentlich sagen. Es geht um die Frage: Wie können zukünftige Studien aussehen? Diese Diskussion läuft im Augenblick auf Hochtouren. Wir wollen nicht ausschließen, dass es dazu relativ zeitnah eine Pressemitteilung gibt. Das ist schwerpunktmäßig die Rückmeldung, die er mir jetzt kurzfristig geben konnte.

Herr Prof. Hecken (Vorsitzender): Danke schön. Wir können ihn daher nicht mehr abringen. Wenn über zukünftige Studiendesigns diskutiert wird, dann leite ich daraus ab, dass man hier jedenfalls sicher nicht das optimale Studiendesign vorgefunden hat; denn ansonsten bräuchte es ja nicht über künftige Studien noch solche Diskussionsprozesse. – Bevor ich Frau Grell das Wort gebe, möchte ich gerne die Frage stellen: Was ist denn mit der FDA? Da hatten wir bei der letzten Bewertung die Situation, dass sie immer zugelassen hatte, dass bei der FDA noch keine Entscheidung getroffen war. Gibt es dazu irgendwelche Fortschritte? Wie ist da der aktuelle Stand des Verfahrens? – Aber jetzt als nächstes Frau Dr. Grell mit ihrer Frage. Dann kann man das in einem Aufwasch machen.

Frau Dr. Grell: Ich bleibe transatlantisch. Sie hatten ein Conditional Marketing Authorisation, das immer noch in der deutschen Fachinformation steht, und in der deutschen oder europäischen Fachinformation steht ja auch die Studie 020 noch nicht und eigentlich war die Time Line, dass Sie im dritten Quartal 2015 nach EPAR Ihre Daten vorlegen sollten. Jetzt sind wir schon im vierten Quartal 2016, das heißt, gehe ich richtig in der Annahme, dass es jetzt seit einem Jahr eine Diskussion um die Conditional Approvals von Ataluren gibt?

Herr Prof. Hecken (Vorsitzender): Herr Schopen, bitte.

Herr Dr. Schopen (PTC Therapeutics Limited): Ich versuche mein Bestes im meinem Zweitjob. Ich möchte eines ergänzen, was Herr Boulding nicht gesagt hat. Ein bisschen kenne ich mich auch aus. Seit einem Jahr wird da nicht diskutiert. Wir haben die Daten Anfang des Jahres eingereicht. Sie haben richtig beobachtet, dass wir zum jetzigen Zeitpunkt eigentlich schon etwas Konkretes an Rückmeldungen erwartet hätten. Die Diskussion geht im Augenblick weiter. Wir erwarten aber ganz zeitnah jetzt eine konkrete Rückmeldung von der EMA. Wir gehen davon aus, dass das Conditional Approval bestätigt wird. Die Zulassung wird auf keinen Fall rückgängig gemacht. Das steht überhaupt nicht zur Diskussion. Aber darüber können wir im Augenblick auch nur spekulieren. Aber das, was wir, glaube ich, mit hinreichender Sicherheit andeuten können: Da kommt jetzt etwas ganz kurzfristig.

Herr Prof. Hecken (Vorsitzender): FDA?

Herr Dr. Schopen (PTC Therapeutics): Die Zusammenfassung der Rückmeldung von Herrn Boulding ist Folgende: Die FDA hat uns ein sogenanntes Refuse to File gegeben. Für mich als kommerziellen Menschen: Die haben sich das Ding gar nicht angesehen. Dagegen haben wir Einspruch erhoben. Dazu gibt es einen Prozess bei der FDA. Das ist entsprechend auf die höhere Etage weitergeleitet worden. Da sind wir in einem konkreten Prozess und der geht auch soweit gut voran. Wir denken, dass sich das bis zum Ende des Jahres gelöst haben sollte. Letztendlich – Sie haben sicherlich auch den Zarepter Case (?) beobachtet – stimmt es uns recht optimistisch, dass dafür im Augenblick auch ganz gute Rahmenbedingungen sind.

Herr Prof. Hecken (Vorsitzender): Okay. Also konkret: keine Zulassung bislang und das Rechtsverfahren dagegen läuft. – Herr Kulig, noch eine Frage.

Herr Kulig: Das war genau das mit dem Refuse to File.

Herr Prof. Hecken (Vorsitzender): Okay. Alles klar. Wunderbar. – Dann sehe ich keine Frage mehr. Dann für Sie noch die Möglichkeit, kurz zusammenzufassen, was Sie aus der letzten guten Stunde, eine Stunde fünfzehn, als relevant ansehen, die wir hier eben hatten. Wer möchte das machen? Herr Schopen? – Bitte schön, Herr Schopen.

Herr Dr. Schopen (PTC Therapeutics Limited): Ich möchte sagen, dass wir uns zum einen für die Fragen sehr herzlich bedanken. Ich glaube, wir haben hier Fragen diskutiert und angesprochen, die wir als extrem relevant ansehen. Die Metaanalyse bezogen auf die Gesamtpopulation betrachten wir als geeignete Methodik. Wir haben auch versucht, die signifikanten und den patientenrelevanten Nutzen in mehreren Endpunkten hier zu beschreiben. Worüber wir nicht gesprochen haben – aber nur rein von der Beobachtung her –: So ziemlich alle Effekte, die in diesen Studien gezeigt worden sind, zeigen in die gleiche Richtung. Das Sicherheitsprofil der Studie ist, glaube ich, besprochen worden.

Die besondere Bedeutung der 300- bis 400-Meter-Subgruppe einfach aufgrund des Studiendesigns und aufgrund der neuen Erfahrungen über den natürlichen Krankheitsverlauf haben wir versucht, hier in einen Kontext zu bringen. Dass man in der Phase am besten eine Wirkung bzw. einen Effekt in einer Ein-Jahres-Studie zeigen kann, beweist aber weiß Gott nicht, dass es in den anderen Gruppen nicht nachzuweisen ist.

Insgesamt kann ich nur sagen: Wir betrachten das ganze Verfahren als eine zweite Chance für kleine Kinder und für Jungen, für die es keine andere therapeutische zugelassene Option in Deutschland gibt. Wir würden uns freuen, wenn die weiteren Überlegungen nach den methodischen Standards des G-BA und IQWiG zu Ende geführt werden, und wir würden uns sicherlich sehr freuen, wenn damit Rahmenbedingungen geschaffen werden, dass wir das Produkt in Deutschland wieder anbieten können. – Auf jeden Fall möchten wir uns recht herzlich bedanken.

Herr Prof. Hecken (Vorsitzender): Ganz herzlichen Dank, dass Sie da waren und auch zur Diskussion zur Verfügung gestanden haben. Dann können wir diese Anhörung beenden. Ich unterbreche dann auch die Sitzung bis morgen früh, 10 Uhr, und mache Sie noch einmal darauf aufmerksam, dass wir um 11 Uhr in Abweichung von der Tagesordnung die Anhörung zum Tagesordnungspunkt 4.1.4 nachholen, die wegen der flugtechnischen Verspätungen heute Nachmittag nicht stattfinden können. – Damit ist die Sitzung für heute geschlossen. Dann sehen wir uns morgen früh wieder. – Danke, dass Sie da waren.

Schluss der Anhörung: 17.40 Uhr