

Synopse Bewertungen TG / BPtK

Nr.	Zitat Studie	Abschließende Bewertung TG Gesprächspsychotherapie des G-BA	Abschließende Bewertung Expertenkommission der BPtK
1	<p>Cooper PJ, Murray L, Wilson A, Romaniuk H. Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression. I. Impact on maternal mood. Br J Psychiatry 2003; 182: 412-9.</p>	<p>Gute Studienqualität. Einschränkungen: Es wurde „non-direktive counselling“ durchgeführt, Vergleichbarkeit dieses Verfahrens mit Gesprächspsychotherapie nach Rogers muss bezweifelt werden; gleiches gilt für die beiden anderen Verfahren: Vergleichbarkeit mit Richtlinienverfahren VT und tiefenpsychologischen Verfahren muss in Frage gestellt werden. Weiterbildung der „Spezialisten“-Therapeuten sowie Rolle der „angelegerten“ Therapeuten unklar. Nicht vergleichbares therapeutisches Setting: Hausbesuche</p> <p>Bedeutung für die ambulante Versorgung (Feld 24): Spezifisches Krankheitsbild, dessen Behandlung in der ambulanten Regelversorgung nach PT-RL eher selten vorkommen wird und vorzugsweise bereits stationär bzw. psychiatrisch medikamentös behandelt wird. Keine Relevanz für Fragestellung der Themengruppe</p>	<p>Allgemeine methodische Qualität der Studie Gut: hinreichende Fallzahlen, Randomisierung, langfristige Katamnesen, Operationalisierung des Outcomes durch Selbst- und Fremdeinschätzungsverfahren, Kontrolle von Störvariablen, komplexe Datenanalysen Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken) Non-direktive Beratung ist einer gewöhnlichen medizinischen Primärversorgung hinsichtlich der Reduktion von Depressivität bei Post-Partum-Depressionen überlegen. Allerdings sind die Effekte längerfristig dem Auftreten von Spontanremissionen nicht überlegen und erreichen kurzfristig auch kein Ausmaß, dass sich in einer signifikanten Reduktion der Diagnosestellung einer Major Depression zeigt. Interne Validität (Sicherheit kausaler Schlussfolgerungen) Gut Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem Gut für Post-Partum-Depressionen; Allgemein eingeschränkt durch Einschränkung auf Post-Partum-Depressionen Stärken der Studie: langer Katamnesezeitraum, geringe Drop-Out-Rate, komplexe multivariate statistische Verfahren Schwächen der Studie:</p>
			<p>Unpräzise Beschreibung der Treatment- und Kontrollbedingungen Keine Berücksichtigung begleitender, nicht randomisierter Interventionen (z.B. Pharmakotherapie) Bedeutung der Studie für die Bewertung der Psychotherapiemethode Relevant: Non-direktive Beratung ist als Gesprächspsychotherapie zu werten (GPT I), unklar ist dabei welche Bedeutung der Durchführung der GT durch Laien - Nicht-Psychologen und Nicht-Ärzte – zukommt, ob die Effektivität der GT dadurch, gerade auch im Vergleich zu den anderen angewandten Psychotherapiemethoden, geringer</p>

Tragende Gründe – Supplement 1

			<p>ausfiel.</p> <ul style="list-style-type: none"> - kurzfristig signifikant stärkere Reduktion der depressiven Symptomatik im Vergleich zur Primärversorgung und vergleichbare Effektivität wie etablierte psychotherapeutische Verfahren (psychodynamische und kognitiv-behaviorale Interventionen). - keine langfristige Überlegenheit bei der Reduktion der depressiven Symptomatik sowie keine kurzfristige oder langfristige Reduktion bezüglich klinischer Depressionsdiagnosen gegenüber üblicher Primärversorgung
2	<p>Ward E, King M, Lloyd M, Bower P, Sibbald B, Farelli S, Gabbay M, Tarrier N & Addington-Hall J (2000). Randomised controlled trial of non-directive counselling, cognitive behavior therapy, and usual general practitioner care for patients with depression I: Clinical effectiveness. <i>BMJ</i>, 321, 1383-1388</p> <p>Siehe auch Bower et al. 2000, King et al. 2000</p>	<p>Kurzfristig erzielt Psychotherapie (CBT u. non-directive counselling) bessere Ergebnisse als general care, nach 12 Monaten gleicht sich general care an die Psychotherapieergebnisse an. Der Effekt der Gleichheit der Ergebnisse nach 12 Monaten kann auf berichtete Medikation etc. (49% Antidepressiva, 18% angstlösende Medikation, 34% Überweisung an psychotherapeutische Versorgung) für die General-Practise-Gruppe nach „Behandlungsende“ zurückzuführen sein.</p>	<p>Die Studie belegt unter Feldbedingungen, die auf Deutschland übertragbar sind, die Wirksamkeit von nicht-direktiver Kurzzeitpsychotherapie auf die Reduktion depressiver Symptomatik bei depressiven Störungen. Die erreichten Verbesserungen sind klinisch relevant, der allgemeinärztlichen Versorgung signifikant überlegen, unterscheiden sich nicht signifikant von der Wirksamkeit kognitiv-behaviouraler Kurzzeittherapie und sind bis zur Katamnese zeitlich stabil.</p>
3	<p>Watson JC, Gordon LB, Stermac L, Kalogerakos F & Steckley P (2003). Comparing the effectiveness of process-experiential with cognitive behavioral psychotherapy in the treatment of depression. <i>JCCP</i>, 71 (4), 773-781</p>	<p>Beide Verfahren (CBT u. PET) sind ähnlich wirksam, PET scheint besser die interpersonellen Probleme behandeln zu können als CBT (Feststellung der Autoren). Die Studie lässt eine leichte Überlegenheit der PET in Bezug auf die interpersonellen Probleme von depressiven Patienten mittlerer Ausprägung erkennen.</p> <p>Kritik:</p> <ul style="list-style-type: none"> - keine Katamnese! - es ist strittig, ob sich es sich bei dem angewandten Verfahren (PET) um GT handelt. 	<p>Hinsichtlich des primären Zielkriteriums BDI-Wert zeigt die „Process-Experiential Psychotherapie“ (klientenzentrierte Beziehungsgestaltung mit zusätzlichen gestalttherapeutischen Elementen) gleichwertige Effektivität in Pre-Post-Vergleich und allen erhobenen sekundären Zeitkriterien.</p> <p>Als Schwäche der Studie werden insbesondere hervorgehoben keine Katamnese; hohe Drop-Out-Rate.</p>
4	<p>Grawe K. Differentielle Psychotherapie I. Indikation und spezifische Wirkung von Verhaltenstherapie und Gesprächspsychotherapie. Eine Untersuchung an phobischen Patienten. (1. Auflage). Bern: Huber. 1976.</p>	<p>Aussage VT und GPT-Einzeltherapie im Vergleich zu einer unbehandelten Kontrollgruppe sei als wirksam erwiesen ist vorsichtig zu interpretieren wegen deutlicher methodischer Schwächen: Randomisierung nicht adäquat; sehr kleine Stichprobe; Diagnosestellung nicht operationalisiert; viele Veränderungsmesswerte (75 Variablen im 3fachen Vergleich getestet; bei 53 Patienten); Auswahl der Patienten/ Studienteilnehmer durch die beiden Untersucher nicht nach standardisierten Kriterien (betrifft aber alle Forschungsgruppen); keine differenzierte Darstellung der Charakteristika der Stichprobe; Supervision der Therapeuten durch den Untersucher. Behandlung durch Studenten. Hinweis auf differenzielle Indikation. Zielsetzung der Studie (differenzielle Psychotherapieeffekte) für</p>	<p>Allgemeine methodische Qualität der Studie</p> <p>Die allgemeine methodische Qualität der Studie ist mit Einschränkungen gut. Problematisch ist, dass die Diagnosestellung basierend auf dem klinischen Urteil erfolgt. Die zur Veränderungsmessung eingesetzten Verfahren sind z.T. veraltet. Die Stichprobe und daraus resultierende Power sind für differentielle Vergleiche zu gering. Z.T. wäre eine Korrektur für multiples Testen nötig gewesen und nicht durchgeführt; die Hauptergebnisse würden allerdings bei einer Korrektur bestehen bleiben.</p> <p>Als positiv ist zu bewerten, dass die Wartegruppe in der schon länger erfolglosen medikamentösen Behandlung blieb, da das völlige Absetzen der Medikation sicherlich nicht eingehalten worden wäre.</p> <p>Zusammenfassung der Ergebnisse</p>

Tragende Gründe – Supplement 1

		<p>die damalige Zeit innovativ, aber aufgrund der multiplen Testungen wären Bestätigungsstudien notwendig gewesen.</p> <p>Bedeutung für die ambulante Versorgung (Feld 24):</p> <p>Aufgrund der methodischen Mängel ist keine eindeutige Aussage zur Wirksamkeit oder vergleichenden Wirksamkeit von GT oder VT möglich.</p>	<p>(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken)</p> <p>Verglichen mit der Wartegruppe führen sowohl VT als auch GT zu einer signif. Verbesserung der phobischen Symptomatik, die zur Katamnese anhält, weiterhin zu positiven Persönlichkeitsveränderungen und Verbesserungen der Eigenbefindlichkeit. Bei den gängigen Mittelwertsvergleichen unterscheiden sich die Verfahren nicht in ihrer Wirksamkeit. Bei Betrachtung der Veränderungsvarianz schneidet die VT besser ab als die GT.</p>
			<p>Der Medikamentenverbrauch verringert sich nur bei VT signifikant (Mittelwertsvergleich). Die Ergebnisse bleiben zur Katamnese im Wesentlichen stabil.</p> <p>Bei VT und GT ergeben sich unterschiedliche Veränderungsmuster; bei VT hängen die erzielten Veränderungswerte enger miteinander zusammen als bei GT und Wartegruppe.</p> <p>Hinsichtlich einer differentiellen Indikation ergibt sich, dass die VT bei Patienten mit starken Symptomen/ hohem phobischen Leidensdruck zu besonders deutlichen Verbesserungen führt. Für GT Wirksamkeit findet sich kein spezifischer Indikator, sie wirkt also auch bei schwächerem phobischen Leidensdruck.</p> <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen)</p> <p>Die interne Validität ist insgesamt gut.</p> <p>Problematisch ist, dass die Supervision für beide Gruppe unterschiedlich abläuft (Einzel für GT, Gruppe und Trainingstherapien vor der Studie für VT). In einigen wenigen der zahlreichen Maße unterscheiden sich die Gruppen zum Zeitpunkt Prä, was aber nicht so problematisch ist, da mit Veränderungswerten gerechnet wird.</p> <p>Externe Validität</p> <p>(Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)</p> <p>Die externe Validität hinsichtlich der Patienten ist gut, der Intervention, der Therapiefrequenz und -länge ist gut.</p> <p>Die Behandler (insb. VT) sind recht unerfahren (fortgeschrittene Studenten). Dem Ungleichgewicht versucht man durch eine stärkere Supervision/ Training bei VT entgegenzuwirken.</p> <p>Stärken der Studie:</p> <ul style="list-style-type: none"> - Zwei Therapiegruppen + Wartegruppe. Beide Therapieformen wurden als „best-practise“ realisiert. - randomisierte Zuordnung

Tragende Gründe – Supplement 1

			<p>Schwächen der Studie:</p> <ul style="list-style-type: none"> - N und resultierenden Power sind zu gering für differentielle Fragestellungen. Die externe Validität der Behandler ist gering.
			<p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode</p> <p>Die Studie wird als geeignet für den Nachweis der Wirksamkeit der GT bei Angststörungen angesehen. Obwohl der Vergleich mit der VT einem Power-Problem unterliegt, und d.h. nur als deutlicher Hinweis interpretiert werden sollte, wird ein Wirksamkeitsnachweis durch den Vergleich mit der Wartegruppe erbracht. Diese Studie ist eine von wenigen zum Thema Angststörungen, bei der GT nicht nur eine unspezif. Vergleichsbedingung darstellt, sondern die Realisierung einer guten GT angestrebt wird.</p>
5	<p>Teusch, L., Böhme, H. & Gastpar, M. (1997). The benefit of insight-oriented and experiential approach on panic and agoraphobic symptoms. <i>Psychotherapy and Psychosomatics</i>, 66, 293-301</p>	<p>An sich gute Studie, bis auf:</p> <ul style="list-style-type: none"> - fehlende Angabe zu den stationären Rahmenbedingungen (unkontrollierbare Wirkvariablen) - fehlende systematische Erfassung und Berücksichtigung weiterführender Therapien nach Ende der stationären Behandlung - mit der Kombinationsbehandlung scheinen sich bestimmte Zielsymptome schneller zu bessern als bei alleiniger CCT-Behandlung (siehe auch Ref ID 3236). Allerdings muss auch hier des stationären Settings wegen die Generalisierbarkeit dieser Aussage in Frage gestellt werden. - für eine valide vergleichende Wirksamkeitsüberprüfung von CCT wäre auch die Überprüfung von EXP allein im stationären Setting von Nöten. 	<p>Die Studie ist aufgrund methodischer Einschränkungen nicht für einen Nachweis der Wirksamkeit der GT bei Angststörungen geeignet. Das Design lässt Aussagen über die Wirksamkeit der GT nur auf Evidenzstufe IV zu, da es für die hier interessierende Fragestellung keine geeignete Vergleichsgruppe gibt. Sie liefert jedoch Hinweise auf eine Wirksamkeit der GT in einem stationären Setting.</p> <p>Die spezifische Wirksamkeit kann nicht beurteilt werden, da es 1. keine echte Kontrollgruppe gibt und 2. stationär in jeder Gruppe ein Bündel von Maßnahmen realisiert wurde. Die allgemeine methodische Qualität und die externe Validität bzgl. Therapiesetting, -frequenz sind eingeschränkt.</p>
6	<p>Teusch, L. & Böhme, H. (1999). Is the exposure really crucial in agoraphobia? The influence of client-centered "nonprescriptive" treatment in exposure. <i>Psychotherapy Research</i>, 9 (1), 115-123</p>	<ul style="list-style-type: none"> - multiple t-Tests und ANOVAs ohne Testplanung und Alpha-Adjustierung ist als Methode der Wahl stark anzuzweifeln. - Für die verwendeten Verfahren sind die Gruppen CCT und EXP+CCT sehr klein, dafür ist die EXP-Gruppe fast doppelt so groß wie die beiden anderen zusammen. - Die klinische Relevanz ist aufgrund der beschriebenen Mängel, die gerade in der Nicht-Vergleichbarkeit der Behandlungsgruppen liegen (siehe 19), anzuzweifeln. - Für eine valide vergleichende Wirksamkeitsüberprüfung von CCT wäre auch die Überprüfung von EXP allein im stationären Setting vonnöten. <p>EXP + GPT war schneller wirksam als GPT alleine. Die Studie lässt keine Aussage über Wirksamkeit der GPT zu (Feld 24)</p>	<p>Die Studie ist aufgrund methodischer Einschränkungen nicht für einen Nachweis der Wirksamkeit der GT bei Angststörungen geeignet, liefert jedoch Hinweise auf eine Wirksamkeit der GT in einem stationären Setting.</p>
7	<p>Teusch, L., Böhme, H. & Finke, J. (2001). Konfliktzentrierte Monotherapie oder Methodenintegration? Verän-</p>	<p>Bitte Text aus Feld 22 (23?) kopieren Stationäres Setting: hier ist die Wirkung immer die eines „Ge-</p>	<p>Die Studie ist aufgrund methodischer Einschränkungen nicht für einen Nachweis der Wirksamkeit der GT bei Angststörungen geeig-</p>

Tragende Gründe – Supplement 1

	<p>derungsprozesse von Gesprächspsychotherapie mit und ohne verhaltenstherapeutische Reizkonfrontation bei Agoraphobie mit Panikstörung. Nervenarzt, 72 (31), 31-39</p>	<p>samtpaketes“. Kritik am Design: GPT vs. GPT+EXP nicht zielführend, um Wirksamkeit von GPT zu untersuchen. Zielführend für die Fragestellung... wäre gewesen EXP vs. GPT+EXP</p>	<p>net. Das Design lässt Aussagen über die Wirksamkeit der GT nur auf Evidenzstufe IV zu, da es für die hier interessierende Fragestellung keine geeignete Vergleichsgruppe gibt. Sie liefert jedoch Hinweise auf eine Wirksamkeit der GT in einem stationären Setting.</p>
8	<p>Altenhöfer, A., Schulz W. Schwab, R., Eckert J. Psychotherapie von Anpassungsstörungen. Ist eine auf 12 Stunden begrenzte Gesprächspsychotherapie ausreichend wirksam? Psychotherapeut (Springer) Manuskript 2006 („acceptetd“)</p>	<p><u>Nachteile:</u> Bis 48,4 % der untersuchten Patienten erhielten begleitend eine „Medikation“, ohne diese zu spezifizieren. Vermutlich handelt es sich um Psychopharmaka. Es wird nicht berichtet, welche Substanzen oder Substanzgruppen in welcher Dosierung zum Einsatz kamen und ob sich die Substanzgruppen auf beide Gruppen gleich verteilen.</p> <p>Werte der Behandlungsgruppe liegen im HADS und ADS knapp über dem Schwellenwert für den auffälligen Bereich. Der Schweregrad der Symptomatik der Untersuchungsgruppen – gemessen mit den drei Hauptskalen - bewegt sich im mittleren Bereich um im Verlauf der Therapie deutlich abfällt.</p> <p>Widersprüchliche Angaben in Text und Tabelle bezogen auf Ausgangswerte der ADS und HADS-Mittelwerte in den beiden Gruppen (siehe S. 6, 3. Absatz)</p> <p><u>Vorteile:</u> Naturalistisches Design erlaubt Übertragbarkeit auf Population in der Versorgung.</p> <p>Hinsichtlich der gewählten Methodik ist anzumerken, dass die Diagnose mit einem strukturierten und standardisierten klinischen Interview vorgenommen wurde. Wichtig ist die Erwähnung, dass Therapeuten eingesetzt wurden, die über eine im Durchschnitt sehr lange Berufserfahrung verfügten.</p>	<p>Abschließende Bewertung Allgemeine methodische Qualität der Studie (Altenhöfer) Die allgemeine methodische Qualität der Studie ist hoch. Es erfolgt eine objektive und reliable Diagnosestellung. Die statistische Power ist für den Vergleich mit einer unbehandelten Gruppe als hoch einzuschätzen. Die gewählten Outcome-Verfahren sind reliabel und valide. Alle Ergebnisse werden vollständig berichtet. Es gibt statistische Mängel; Therapiegruppe und Wartegruppe haben ein unterschiedliches Ausgangsniveau in den Outcome-Variablen. Dem hätte man nicht nur durch eine korrigierte Effektgröße, sondern auch durch die Bildung entsprechender Kovariaten gerecht werden können. Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken) # konsistente Befunde # hochsignifikante Verbesserungen der Therapiegruppe im Vergleich zur Wartegruppe in mehreren Maßen # Verbesserungen sind auch klinisch relevant # Verbesserungen sind zur 3-Monats-Katamnese/ und nach 2 Jahren stabil # keine Nebenwirkungen/ Risiken bekannt Interne Validität (Sicherheit kausaler Schlussfolgerungen) Ein- und Ausschlusskriterien sind klar spezifiziert; es gibt keine Selektionseffekte durch selektiven Ein- und Ausschluss. Es gibt ein ausführliches Therapiemanual, die Manualtreue wird allerdings nicht belegt. Problematisch sind die Unterschiede in den Ausgangswerten durch den Verzicht auf randomisierte Zuordnung. Eine angemessene statistische Kontrolle dieser Unterschiede fehlt (z.B. Prä-Werte als Kovariate). Die Katamnese erfolgt mit 3 Monaten sehr zeitnah zum Therapieende, was aber bei Anpassungsstörungen möglicherweise. (<i>weitere Katamnese nach 2 Jahren, Ergebnisse liegen vor</i>)</p>
		<p>Positiv ist auch, dass die methodischen Probleme beim Einschluss der Patienten in die Studie und Zuweisung zur Therapie- oder Wartegruppe offen diskutiert wurden. Die Beurteilung der Therapieeffekte erfolgte mit einem umfangreichen Instrumentarium standardisierter Testverfahren bzw. Skalen.</p>	<p>ein gut gewählter Zeitraum ist. Externe Validität (Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)</p>

Tragende Gründe – Supplement 1

		<p>Zur Erfassung der Kernsymptomatik des ängstlichen und/oder depressiven Syndroms kamen standardisierte Skalen zum Einsatz, die überwiegend im psychologischen Forschungskontext zum Einsatz kommen, so die HADS-D, der STAI sowie die ADS.</p>	<p>Die externe Validität ist als hoch zu bewerten. Die Patientenstichprobe ist nicht selektiv, die Repräsentativität der Behandler und der Intervention ist hoch. Das Therapeutenverhalten wird nicht durch Supervision verändert. Die Frequenz (2x/ Woche) der Intervention ist repräsentativ und gut auf ein ambulantes Setting im dt. Versorgungssystem übertragbar.</p> <p>Stärken der Studie: Die methodische Qualität und externe Validität sind hoch. Die Therapeuten sind erfahren und verfügen über einen Abschluss in GT.</p> <p>Schwächen der Studie: Die Zuordnung zu den Gruppen erfolgt aus ethischen Gründen nicht-randomisiert, sondern nach Reihenfolge des Einschlusses und der Zumutbarkeit einer Wartezeit, daher ist die Wartegruppe zum Zeitpunkt prä im Mittel weniger stark betroffen als die Therapiegruppe.</p> <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode Die Studie wird als geeignet für den Nachweis der Wirksamkeit der GT bei Anpassungsstörungen angesehen. Die Behandlungsgruppe zeigt statistisch und klinisch bedeutsame Verbesserungen der Störung im Vergleich zu einer Wartegruppe. Bei ansonsten hohem methodischen Niveau, ist die nicht erfolgte Randomisierung der Zuordnung problematisch.</p>
9	<p>Dirks P., Grimm F, Tausch AM, Wittern JO. Förderung der seelischen Gesundheit von Krebspatienten durch personenzentrierte Gruppengespräche. Z Klin Psychol 1982; 11 (4): 241-52.</p>	<p>unklar, ob und welche Störungen von Krankheitswert behandelt wurden; eher Maßnahmen zur Förderung von seelischer Gesundheit und Lebensqualität bei Krebspatienten. Für die Fragestellung nicht relevant, da keine Behandlung von psychischen Störungen, sondern psychosoziale Unterstützung bei somatischem Krankheitsbild.</p> <p>Methodische Kritik: Statistische Auswertung z.T. nicht nachvollziehbar, zahlreiche Mängel in der statistischen Auswertung</p> <p>Bedeutung für die ambulante Versorgung (Feld 24): Studie ist nicht geeignet Wirksamkeit und Nutzen der GT zu belegen.</p>	<p>Es gibt starke Einschränkungen der allgemeinen methodischen Qualität der Studie. Es erfolgt keine objektive und reliable Diagnosestellung einer klinisch relevanten Störung z.B. einer Anpassungsstörung (Patienten mit Krebserkrankung). Die Angaben über sozio-demographische Daten, Ein- & Ausschlusskriterien, Zugangswege zu der Studie, Dropout – Raten sind unzureichend (Für Details wird auf die Dissertationsschrift von Kalliner, 1981 verwiesen.) Es ist unklar, wie die Patienten der Gesprächsgruppe bzw. der Vergleichsgruppe zugeordnet werden. Veränderungswerte der Kontrollgruppe werden nicht direkt angegeben. Die Beschreibung der Statistik ist äußerst vage („bei Veränderungsmessung gängige Verfahren“) und lässt sich nur durch Hinzuziehen der Dissertation besser nachvollziehen. Es ist unklar, inwiefern sich Therapie- und Vergleichsgruppe voneinander unterscheiden, alle Ergebnisse werden lediglich in % angegeben ohne Varianz- und Signifikanzangaben.</p> <p>Zusammenfassung der Ergebnisse</p>

Tragende Gründe – Supplement 1

			<p>(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken) # 27 % der Gesprächspatienten längerfristig in verschiedenen Bereichen ihrer Persönlichkeit „deutlich verbessert“, 37 % „leicht verbessert“, 30% „längerfristig unverändert“, 6 % „leicht verschlechtert“ (Gesamteinschätzung der individ. Patienten von 3 Experten) # mehr (soziale) Aktivitäten, Verbesserung (familiärer) Beziehungen, Verringerung der Nervosität (FPI), weniger Krebsangst, höhere körperliche und seelische Belastbarkeit, günstigere Einstellung zu der Krankheit und zum Leben # 97 % der Gesprächspatienten halten personenzentrierte Gesprächsgruppen im Zusammenhang mit der Krankheit für wichtig und hilfreich Interne Validität (Sicherheit kausaler Schlussfolgerungen) Die interne Validität ist schwer abzuschätzen, da etliche Angaben fehlen. Insbesondere ist unklar, wie die Patienten den beiden Gruppen zugeteilt wurden, ob die Gruppen zum Zeitpunkt prä äquivalent waren, ob es zusätzliche nicht randomisierte Einflüsse gab.</p>
			<p>Externe Validität (Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem) Die externe Validität ist bzgl. der Behandler gut und bzgl. der Repräsentativität der Intervention mittelmäßig (Gruppentherapie, 3 von 11 Sitzungen erstrecken sich über einen ganzen Tag). Die Frequenz (1x/ Woche) der Intervention ist repräsentativ und gut auf ein ambulantes Setting im dt. Versorgungssystem übertragbar Stärken der Studie: Der Patient evaluiert das Therapeutenverhalten (Realisierung der drei Basisvariablen nach Rogers). Schwächen der Studie: Es wird keine klare Diagnose einer psychischen Störung und komorbider Störungen gestellt. Es kann sicherlich nur bei einem Teil der Krebspatienten von einer Störung im Sinne einer Anpassungsstörung ausgegangen werden. Es besteht das Risiko prognostisch wichtige komorbide Erkrankungen zu übersehen. Das Design erlaubt nicht, die Effekte „personenzentrierter Gruppengespräche“ von denen „nicht-personenzentrierter Gruppengespräche,“ zu unterscheiden. Es gibt etliche methodische Mängel, z.B. ist aus dem Artikel nicht ersichtlich,</p>

Tragende Gründe – Supplement 1

			welche Unterschiede statistisch und/oder klinisch relevant sind Bedeutung der Studie für die Bewertung der Psychotherapiemethode
			Die Studie ist aufgrund methodischer Mängel nicht für einen Nachweis der Wirksamkeit der GT geeignet, liefert jedoch durchaus Hinweise auf eine Wirksamkeit der GT bei Reaktionen auf schwere Belastungen/ Anpassungsstörungen. Es gibt starke methodische Mängel der Studie. Einige Unklarheiten lassen sich durch das Studium der Dissertationsschrift von Kalliner, 1981 beseitigen. Problematisch ist aber, dass das Design nicht erlaubt, die Effekte „personenzentrierter Gruppengespräche“ von denen „nicht-personenzentrierter Gruppengespräche“ zu unterscheiden. Möglicherweise ist somit zumindest ein Teil der berichteten Effekte auf eine unspezifische „Heilsamkeit“ des Austausches mit anderen Betroffenen zurückzuführen.
10	<p>Paivio SC, Nieuwenhuis JA. Efficacy of emotion focused therapy for adult survivors of child abuse: A preliminary study. J Trauma Stress 2001; 14 (1): 115-33.</p>	<p>Methodisch hochwertige Studie.</p> <p>In der Studie wird keine klassische GPT angewandt, sondern die EFT, (process-experiential Therapie mit Elementen der Gestalttherapie).</p> <p>Nachgewiesene Effekte für diagnostisch homogene Gruppe mit Kindesmissbrauch in der Vorgeschiedten (keine Multimorbidität). Das Vorgehen ist laut Autoren wahrscheinlich weniger geeignet für Patienten mit multiplen und dauerhaften Problemen</p>	<p>Allgemeine methodische Qualität der Studie</p> <ul style="list-style-type: none"> • Design stellt die Vergleichbarkeit der Gruppen nur eingeschränkt sicher • Verwendung reliabler Messinstrumente <p>Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken)</p> <ul style="list-style-type: none"> • EFT-AS signifikant wirksamer als unbehandelte Gruppe • auch im 9-Monats-Follow-up bleiben Effekte erhalten <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen)</p> <ul style="list-style-type: none"> • sehr gut • Überprüfung der internen Reliabilität durch Monitoring der Intervention <p>Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem</p> <ul style="list-style-type: none"> • Insgesamt ausreichend • Ein- und Ausschlusskriterien teilweise uneindeutig, dadurch schwer abschätzbar, wie stark die Selektivität der Patienten ist • Starkes Monitoring <p>Stärken der Studie:</p> <ul style="list-style-type: none"> • Mehrere Behandler, dadurch kein Personen-Effekt (wurde auch statistisch belegt) • Randomisierte Zuteilung

Tragende Gründe – Supplement 1

			<ul style="list-style-type: none"> • Manualisierte Behandlung <p>Schwächen der Studie:</p> <ul style="list-style-type: none"> • Nur Selbsteinschätzung
			<ul style="list-style-type: none"> • Am Ende relativ kleine Gruppen (im Follow-up) <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode</p> <p>Zeigt, dass für diese Population EFT-AS eine wirksame Behandlungsmethode ist und signifikante Effekte im Vergleich zur unbehandelten Warte-Kontroll-Gruppe bewirkt, welche auch zur Katamnese erhalten bleiben.</p>
11	<p>Eckert J, Biermann-Ratjen EM, Wuchner M.[Long-term changes in borderline symptoms of patients after client-centered group psychotherapy] Die langfristigen Veränderungen der Borderline-Symptomatik bei Patienten nach klientenzentrierter Gruppenpsychotherapie. Psychother Psychosom Med Psychol 2000; 50 (3-4): 140-6.</p>	<p><i>Methodische Einwände der Autoren selbst: „... es muss offen bleiben, welcher Anteil der Veränderungen auf spontane Remissionen bzw. statistische Regression zurückzuführen ist. Weiterhin muss offen bleiben, worauf die Veränderungen in den Kontrollgruppen zurückzuführen sind.“</i></p> <p><i>Bei fehlender Untersuchung direkt nach der Gruppentherapie und dem langen Katamnesezeitraum unklar, inwieweit die berichteten Veränderungen auf die Gruppentherapie zurückzuführen sind oder anderen Wirkfaktoren (z.B. weitere ambulante und/oder stationäre Behandlungen im Katamnesezeitraum eine Rolle spielen.) Auch ein möglicher Spontanverlauf lässt sich nicht differenzieren (keine „treatment as usual“-Gruppe)</i></p> <p>Bedeutung für die ambulante Versorgung (Feld 24): <i>Studie ist nicht geeignet, Wirksamkeit der GPT in der Gruppe nachzuweisen.</i></p>	<p>Allgemeine methodische Qualität der Studie Die Diagnosestellung einer Borderline-Persönlichkeitsstörung erfolgte objektiv und reliabel (DIB, zwei Beurteiler). Methodische Mängel ergeben sich daraus, dass es keine Post-Untersuchung gibt (nur Voruntersuchung und Katamnese). Das Studiendesign ist unklar: Während alle Borderline-Patienten Gruppentherapie erhalten, bekommen die Patienten der anderen Störungsgruppen zum Teil „klinische Standardintervention“. Für die Borderline-Patienten existiert somit keine echte Vergleichsgruppe.</p> <p>Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken) # deutliche Verbesserung der borderline-spezifischen Symptomatik bei Patienten mit Borderline-Störung, Schizophrenie, Depression (Vergleich Prä vs. Katamnese) # stärkste Verbesserung bei Patienten mit Borderline-Störung</p>
			<p># der Großteil der Patienten erfüllt zur Katamnese nicht mehr die Kriterien einer Borderline-Störung # Anteil einer Spontanremission an der Veränderung nicht ausgeschlossen</p> <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen) Die interne Validität ist Mangelhaft. Die Zuordnung der Patienten zu den Gruppen erfolgt nicht randomisiert. Alle Borderline-Patienten erhalten Gruppentherapie, während Patienten der anderen Störungsgruppen zum Teil die „klinische Standardintervention“ erhalten. Bei den Borderline-Patienten gibt es keine unbehandelte Vergleichsgruppe. Die Vergleichsbedingung „klinische Standardintervention“ wird nicht beschrieben. Für die statistischen Vergleiche werden Patienten jeweils eines</p>

Tragende Gründe – Supplement 1

			<p>Störungsbildes mit und ohne Gruppentherapie gepoolt. Die Qualität der Durchführung der Therapie wird nicht überprüft; die Qualifikation der Behandler ist nicht angegeben</p> <p>Externe Validität (Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem) Die Patientenstichprobe ist nicht selektiv, die Repräsentativität der Behandler ist nicht bekannt. Die Frequenz (2x/ Woche) der Intervention ist repräsentativ und gut auf ein ambulantes Setting im dt. Versorgungssystem übertragbar.</p> <p>Stärken der Studie: Diagnose und Verlaufsmessung u.a. über DIB. Katamnesezeitraum von ca. 1,5 J..</p>
			<p>Schwächen der Studie: Es ist kein klares Versuchsdesign zu erkennen. Nur schizophrene oder depressive Patienten erhalten teilweise Gruppentherapie und teilweise Standardintervention. Diese beiden Interventionen werden nicht verglichen (es sind sehr kleine Gruppen). Bei den Borderline Patienten existiert keine echte Vergleichsgruppe. Beim Vergleich der Veränderungswerte der drei Störungsgruppen werden schizophrene und depressive Patienten jeweils über beide Interventionsbedingungen gepoolt. Die Kontrollbedingung „klinische Standardintervention“ wird nicht näher beschrieben. Auf eine Datenerhebung unmittelbar nach Ende der Therapie wird verzichtet.</p> <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode Die Studie ist mit Einschränkungen geeignet für einen Nachweis der Wirksamkeit der GT bei Persönlichkeitsstörungen. Methodische Einschränkungen ergeben sich insb. aus dem Design, das keine echte Vergleichsgruppe bereitstellt, um die Therapiewirkung von einer Remissionswirkung zu trennen. Zudem gibt es keinen Messzeitpunkt unmittelbar nach dem Therapiezeitraum, sondern nur eine ca. 1,5 Jahres-Katamnese.</p>
12	Wuchner M. Behandlungsergebnisse von Borderline-Patienten nach klientenzentrierter Psychotherapie. Treatment results of borderline patients after client-centered psychotherapy. Hamburg: Universität. 1997.	Die Studie ist aufgrund methodischer Mängel nicht geeignet, die Wirksamkeit von GPT-Gruppentherapie zu belegen. Methoden-Kritik: Untersuchung von Borderline-Patienten zu zwei unterschiedlichen Nachbefragungszeitpunkten:	<p>Allgemeine methodische Qualität der Studie Die Studie entspricht in ihrer allgemeinen Qualität dem Stand der Forschung zu Zeiten ihrer Publikation. Die Q-Sort-Technik erlaubt differenzierte inhaltliche Aussagen zu Prozessforschung und Therapieoutcome, die</p>

Tragende Gründe – Supplement 1

		<p>Beide Drop-outs sind in die POST-Messung miteinbezogen. Nicht alle Vergleichsgruppenpatienten (Schizophrenie, Depression) erhielten Gruppentherapie. Vergleich der Gruppen mit Rückschluss auf die Wirksamkeit von GPT-Gruppentherapie daher nicht zulässig.</p> <p>Die Behandlungsziele der Gruppentherapie waren auf die BPS-Patienten ausgerichtet.</p> <p>Mangelnde Vergleichbarkeit der Gruppen.</p> <p>Verbesserung der Patienten kann nicht sicher auf die GPT-Gruppentherapie zurückgeführt werden, wegen zu vieler anderer unkontrollierter Wirkfaktoren:</p> <p>so suchten bspw. etwa die Hälfte aller BPS-Patienten über die ambulante Gruppentherapie hinaus weitere psychotherapeutische Hilfen;</p> <p>Störvariablen, wie z.B. Medikation nicht dokumentiert und kontrolliert;</p> <p>Keine Kontrolle des Spontanverlaufs.</p> <p>Geringe Fallzahl.</p>	<p>mittels moderner üblicher Messinstrumente und Auswertungstechniken nicht möglich sind. Die Ergebnisse der Q-Sort-Technik erlauben allerdings keine Rückschlüsse auf die spezifische (Alkohol-) Problematik der untersuchten Patienten. Da es sich um stationäre Psychotherapie handelt, wird der Alkoholkonsum vermutlich verboten gewesen sein und die Abstinenz kontrolliert worden sein, so dass diese Schwäche nicht wichtig erscheint.</p> <p>Das bei dieser Art von Störung wichtige Outcome sind Rückfälle nach stationärer Entlassung, diese sind erfasst worden. Zu den Katamnesezeitpunkten gibt es Daten über stationäre Rückfälle, die allein für sich genommen die Wirksamkeit der untersuchten Verfahren hinreichend belegen können.</p> <p>Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken)</p> <p>Es handelt sich um stationäre Psychotherapie der Alkoholabhängigkeit. Auch wenn der Schwerpunkt in der Auswertung der Studie ein anderer war, belegen die Daten eindrücklich die durch klientenzentrierte Gruppenpsychotherapie erreichte Reduktion von Rückfällen nach stationärer Entlassung, die den Raten einer Kontrollgruppe signifikant überlegen ist (25 vs. 53%).</p>
			<p>Interne Validität (Sicherheit kausaler Schlussfolgerungen) Die Qualität der Behandlung erscheint vergleichbar. Als Einschränkungen der internen Validität wurden begleitende Interventionen nicht untersucht.</p> <p>Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem</p> <p>Die Stichprobe ist mit Ausnahme der Begrenzung auf das männliche Geschlecht kaum selektiv, es handelt sich im Wesentlichen um die Inanspruchnahmepopulation eines Krankenhauses. Als Einschränkung der externen Validität wurden keine Drop-out und ITT-Analysen durchgeführt (die zum damaligen Stand der Forschung auch noch nicht vorausgesetzt wurden).</p> <p>Es wurden Standardverfahren der Therapie eingesetzt, das stationäre Kurzzeit-Gruppensetting ist bei Patienten mit Alkoholabhängigkeit auch in Deutschland üblich. Alle Therapeuten praktizierten in der vorliegenden</p>
			<p>Studie alle Therapieverfahren; dies wurde gemacht, um die Ver-</p>

Tragende Gründe – Supplement 1

			<p>gleichbarkeit der Qualität der Behandlung zu gewährleisten. In der Versorgungspraxis würden die Therapien von Therapeuten der jeweiligen Therapierichtung durchgeführt werden, was die Wirksamkeit der Therapie vermutlich erhöhen würde, zumindest die Übertragbarkeit der Ergebnisse jedenfalls nicht einschränkt.</p> <p>Stärken der Studie: Die Vergleichbarkeit der durchgeführten Therapien, der lange Katamnesezeitraum, das Outcome-Kriterium der Rückfälle nach Entlassung.</p> <p>Schwächen der Studie: Es liegen keine Diagnosen nach gängigen Klassifikationssystemen vor, der Krankheitswert der Störungen kann jedoch angenommen werden, da die Patienten stationär wegen Alkoholkonsums aufgenommen wurden. Für die Vergleiche vor und nach Therapieende liegen keine heute üblichen Messverfahren vor, das bei dieser Art von Störungen wichtigste Outcome-Kriterium der Rückfälle nach Entlassung wurde jedoch erfasst. Es fehlen Angaben zu den Drop-outs.</p> <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode Die Studie zeigt, dass stationäre Gesprächspsychotherapie im Gruppensetting einer Kontrollgruppe signifikant überlegen ist hinsichtlich der Reduktion von Rückfällen bei Patienten mit Alkoholabhängigkeit.</p>
13	<p>Ends EJ, Page CW. A study of three types of group psychotherapy with hospitalized male inebriates. Quart J Stud Alcohol 1957; 8:263-77.</p>	<p>Abschließende Bewertung (Feld 23): Nicht verwertbar, da hier die Veränderung von theoretischen Konstrukten (Selbst, Selbst-Ideal) in Gruppen untersucht wurde. Keine Verknüpfung dieser Konstrukte mit klinisch relevanten Veränderungen. Keine Diagnostik von konkreten Krankheitsbildern. Keine Untersuchung von individuellen und krankheitsspezifischen Symptomen und deren Veränderung. Nachuntersuchung durch anderes Instrument als in der eigentlichen Untersuchung. Keine Informationen über die „sonstige“ stationäre Behandlung</p> <p>Bedeutung für die ambulante Versorgung (Feld 24): Studie wurde im stationären Bereich durchgeführt. Keine Übertragbarkeit auf ambulante Versorgung, da Gruppenpsychotherapie nicht alleinige Behandlung.</p>	<p>Allgemeine methodische Qualität der Studie Die Studie entspricht in ihrer allgemeinen Qualität dem Stand der Forschung zu Zeiten ihrer Publikation. Die Q-Sort-Technik erlaubt differenzierte inhaltliche Aussagen zu Prozessforschung und Therapieoutcome, die mittels moderner üblicher Messinstrumente und Auswertungstechniken nicht möglich sind. Die Ergebnisse der Q-Sort-Technik erlauben allerdings keine Rückschlüsse auf die spezifische (Alkohol-) Problematik der untersuchten Patienten. Da es sich um stationäre Psychotherapie handelt, wird der Alkoholkonsum vermutlich verboten gewesen sein und die Abstinenz kontrolliert worden sein, so dass diese Schwäche nicht wichtig erscheint. Das bei dieser Art von Störung wichtige Outcome sind Rückfälle nach stationärer Entlassung, diese sind erfasst worden. Zu den Katamnesezeitpunkten gibt es Daten über stationäre Rückfälle, die allein für sich genommen die Wirksamkeit der untersuchten Verfahren hinreichend belegen können.</p>

Tragende Gründe – Supplement 1

			<p>Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken)</p>
			<p>Es handelt sich um stationäre Psychotherapie der Alkoholabhängigkeit. Auch wenn der Schwerpunkt in der Auswertung der Studie ein anderer war, belegen die Daten eindrücklich die durch klientenzentrierte Gruppenpsychotherapie erreichte Reduktion von Rückfällen nach stationärer Entlassung, die den Raten einer Kontrollgruppe signifikant überlegen ist (25 vs. 53%).</p> <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen) Die Qualität der Behandlung erscheint vergleichbar. Als Einschränkungen der internen Validität wurden begleitende Interventionen nicht untersucht.</p> <p>Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem Die Stichprobe ist mit Ausnahme der Begrenzung auf das männliche Geschlecht kaum selektiv, es handelt sich im Wesentlichen um die Inanspruchnahmepopulation eines Krankenhauses. Als Einschränkung der externen Validität wurden keine Drop-out und ITT-Analysen durchgeführt (die zum damaligen Stand der Forschung auch noch nicht vorausgesetzt wurden). Es wurden Standardverfahren der Therapie eingesetzt, das stationäre Kurzzeit-Gruppensetting ist bei Patienten</p>
			<p>mit Alkoholabhängigkeit auch in Deutschland üblich. Alle Therapeuten praktizierten in der vorliegenden Studie alle Therapieverfahren; dies wurde gemacht, um die Vergleichbarkeit der Qualität der Behandlung zu gewährleisten. In der Versorgungspraxis würden die Therapien von Therapeuten der jeweiligen Therapierichtung durchgeführt werden, was die Wirksamkeit der Therapie vermutlich erhöhen würde, zumindest die Übertragbarkeit der Ergebnisse jedenfalls nicht einschränkt.</p> <p>Stärken der Studie: Die Vergleichbarkeit der durchgeführten Therapien, der lange Kamnesezeitraum, das Outcome-Kriterium der Rückfälle nach Entlassung.</p> <p>Schwächen der Studie: Es liegen keine Diagnosen nach gängigen Klassifikationssystemen vor, der Krankheitswert der Störungen</p>

Tragende Gründe – Supplement 1

			<p>kann jedoch angenommen werden, da die Patienten stationär wegen Alkoholkonsums aufgenommen wurden.</p> <p>Für die Vergleiche vor und nach Therapieende liegen keine heute üblichen Messverfahren vor, das bei dieser Art von Störungen wichtigste Outcome-Kriterium der Rückfälle nach Entlassung wurde jedoch erfasst. Es fehlen Angaben zu den Drop-outs.</p> <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode</p> <p>Die Studie zeigt, dass stationäre Gesprächspsychotherapie im Gruppensetting einer Kontrollgruppe signifikant überlegen ist hinsichtlich der Reduktion von Rückfällen bei Patienten mit Alkoholabhängigkeit.</p>
14	<p>Markowitz JC. Treatment of depressive Symptoms in human immunodeficiency virus-positive patients. Arch Gen psychiatry 1998; 55(5):452-57</p>	<p>Fazit der Autoren: Trotz methodischer Grenzen (geringe Fallzahl, fehlender Nachweis affektiver Störungen im SKID, zum Teil nur klinisch diagnostizierte Depression, überwiegend männliche Patienten) sprechen die Ergebnisse dafür, psychotherapeutischen Interventionen bei HIV-positiven Patienten mehr Aufmerksamkeit zu schenken. Selbst bei HIV-positiven Patienten stellt Depression eine Behandlungsoption dar. Eine Replikation der Studienergebnisse wird empfohlen.</p> <p>Entspricht dem Fazit der Autoren mit einer Einschränkung: es wird nicht deutlich, ob die Depression durch die HIV-Erkrankung ausgelöst wird (was der Kommentar der Autoren zur IPT nahe legt) oder neben der HIV-Erkrankung besteht, zumindest z. T., da ein hoher Prozentsatz der Patienten (zwischen 73 und 60 % über Major depression-Episoden in der Vergangenheit berichtet. IPT und Pharmakotherapie sind im Hinblick auf die Zielsymptomatik der GT überlegen.</p> <p><i>Definition und Durchführung der ND-Bedingung lässt keine Rückschlüsse zu, dass es sich um GPT handelt, wie sie in Deutschland angewandt wird</i></p> <p>Studie ist nicht geeignet Aus-sagen über die Wirksamkeit und den Nutzen von GPT zu machen</p>	<p>Abschließende Bewertung</p> <p>Allgemeine methodische Qualität der Studie</p> <ul style="list-style-type: none"> o insgesamt mittelmäßig o gut: ITT-Analyse; verschiedene Therapeuten, Monitoring o Nachteil: o Manches ist nicht beschrieben, was methodische Mängel beinhalten könnte, z.B. Zeitraum (hat den Anschein, als könnten es mehrere Jahre sein □ Verzerrungen möglich, da z.B. anderer Umgang mit HIV) und Rekrutierung und Dropouts o verschiedene Dauer der Behandlungen, dadurch uneinheitliche Zeitabstände der Messzeitpunkte zur Beendigung der Therapie <p>Zusammenfassung der Ergebnisse</p> <p>(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken)</p> <ul style="list-style-type: none"> o IPT und Supportive Therapie mit Imipramin sind sign. erfolgreicher bei HIV-Patienten mit depressiven Symptomen als CBT und Supportive Therapie ohne Imipramin; aber: kein Follow-Up!
			<ul style="list-style-type: none"> o klinische Relevanz noch fraglich, da nur 53% mit Diagnose affektiver Störung <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen)</p> <ul style="list-style-type: none"> o insgesamt mittel: zwar gutes Monitoring und valide Messinstrumente, aber Gruppen sind nicht vergleichbar hinsichtlich der Dauer und Intensität der Behandlung o teilweise große Unterschiede in der Verteilung der Merkmale zwischen den Gruppen (Schweregrad, Persönlichkeitsstörungen, Diagnosen) – Analyse: Persönlichkeitsstörungen wirken sich nicht

Tragende Gründe – Supplement 1

			<p>aus; jedoch bei höherem Schweregrad mehr Verbesserung</p> <p>Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem ++ viele Behandler, gute Übertragbarkeit aufs Versorgungssystem – einjähriges Training</p> <p>Stärken der Studie:</p> <ul style="list-style-type: none"> ○ verschiedene Therapeuten ○ ITT-Analyse <p>Schwächen der Studie:</p> <ul style="list-style-type: none"> ○ Rekrutierung und Dropout-Fälle nicht beschrieben
			<ul style="list-style-type: none"> ○ verschiedene Intensitäten der Gruppen; Umgang mit Medikamentenverschreibung nicht beschrieben (zusätzlicher Kontakt?) <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode</p> <ul style="list-style-type: none"> ○ Die Studie zeigt eine hohe Wirksamkeit von ST in Kombination mit Imipramin, aber nicht ohne □ spricht hier eher für die medikamentöse Wirkung und wenig Effekte der ST allein. Jedoch ist die Vergleichbarkeit der Gruppen dadurch eingeschränkt, dass die Intensität der Behandlungen stark unterschiedlich ist. Die Intensität der ST geringer als bei den anderen. ○ Die methodischen Voraussetzungen für einen Wirksamkeitsnachweis sind dahingehend nicht gegeben, als dass keine unbehandelte Kontrollgruppe existiert und die anderen Behandlungen intensiver waren.
15	Sachse R. Psychologische Psychotherapie bei chronisch entzündlichen Darmerkrankungen. Göttingen: Hogrefe.2006	<p>Die Studien zeigen dass viele Patienten der drei Krankheitsgruppen von einem solchen Verfahren profitieren und dass die Kosten gesenkt werden könnten. Allerdings ist die spezifische Wirksamkeit des Verfahrens nicht gegeben. Es wurde während der Durchführung massiv verändert. Es gibt keine klinische Kontrollgruppe und die Homogenitätsannahme in Bezug auf die chronisch entzündlichen Darmerkrankungen ist nicht nachgewiesen, muss vielmehr bezweifelt werden. Warum das Verfahren für neurotische Erkrankungen nicht geeignet sein soll bleibt offen. . Die Durchführung einer solchen Behandlung ohne kontinuierliche somatische Diagnostik ist nicht ungefährlich.</p> <p>Die Bedeutung für die Ambulante Versorgung ist gering. Erstens ist das Verfahren definitionsgemäß hochspezifisch auf eine vermeintlich homogenes Krankheitsbild zugeschnitten , die selbst gar</p>	<p>Allgemeine methodische Qualität der Studie Design wäre prinzipiell für Wirksamkeitsnachweis geeignet. Einschränkungen ergeben sich jedoch daraus dass die Bildung bzw. Rekrutierung der verschiedenen Behandlungsgruppen, die miteinander verglichen werden nicht ausreichend beschrieben wird und die Vergleichbarkeit somit nur eingeschränkt beurteilt werden kann. Ferner werden einige wesentliche methodische Aspekte nicht beleuchtet; die Ergebnisse nicht vollständig präsentiert.</p> <p>Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken) es werden nur die Effektstärken und Signifikanzen berichtet – diese</p>

Tragende Gründe – Supplement 1

		<p>nicht nachgewiesen ist. Der Status des Verfahrens in Bezug auf die GT bleibt offen. Ist das nun eine Methode innerhalb des Verfahrens GT oder sind beide gleichrangige Verfahren. Von daher ist die Selbstzuordnung durch den Autor zur kognitiven Verhaltenstherapie durchaus einleuchtend.</p>	<p>zeigen hohe Wirksamkeit von GT im Gegensatz zu den anderen Kontrollgruppen – die Interpretation dieser dargestellten Zusammenhänge ist jedoch vorsichtig zu gestalten, da wesentliche Informationen nicht gegeben werden</p> <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen) schwer zu beurteilen, da relevante Aspekte nicht berichtet werden – Vergleichbarkeit der Gruppen kann nicht sicher bewertet werden.</p> <p>Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem Selektivität der Patienten ist da, wird aber nicht dokumentiert; Intervention an sich scheint übertragbar zu sein auf das deutsche Versorgungssystem; Behandlerrepräsentativität gewährleistet, da mehrere verschiedene Behandler tätig waren</p> <p>Stärken der Studie: speziell auf CEDE ausgerichtete GT wird auf Effektivität hin untersucht verschieden Behandlungsgruppen werden auf Effektivität hin miteinander verglichen</p> <p>Schwächen der Studie: ausgewählte Berichterstattung, Auslassung relevanter Ergebnisse</p> <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode Den Ergebnissen dieser Studie zufolge zeigt die Untersuchung eine deutliche Wirksamkeit von GT bei dieser Stichprobe – ggfs. könnte man die Ergebnisse als Nachweis gelten lassen – notwendig dafür wäre jedoch, dass man die im Text nicht berichteten methodischen Aspekte noch mal nachprüfen könnte (wie kommen Gruppen genau zustande? Was ist mit den nicht berichteten Skalen? Wie vergleichbar sind die Gruppen) – somit können die vorliegenden Veröffentlichungen nur als Hinweis auf eine Wirksamkeit verwendet werden.</p>
16	<p>Meyer et al (1981). The hamburg short-term psychotherapy comparison experiment. <i>Psychother Psychosom</i>, 35, 81-207</p>	<p>Zusammenfassung: Insgesamt lässt sich sowohl aufgrund der methodischen Mängel als auch aufgrund des Designs der Studie keine Aussage über den Nutzen der „Gesprächspsychotherapie“ im Vergleich zu einer „psychodynamischen Kurzzeitpsychotherapie“ treffen...</p>	<p><i>Diskussionsbedarf über Art und Umfang der erforderlichen Darstellung.</i></p> <p>Die abschließende Bewertung der Expertenkommission der BPTK kommt zu vielen kritischen Einschränkungen dieser Studie: - keine Angaben über Diagnosen oder Diagnostik und einzelne</p>

Tragende Gründe – Supplement 1

			<p>Symptome</p> <ul style="list-style-type: none"> - KZT-Therapeuten waren bevorteilt, da Setting und Art der Therapie ungewohnt für PDT-Therapeuten - Intervention...unüblich und nicht in der Praxis realisierbar <p>Gleichwohl stellt die Expertengruppe fest: „Methodisch lässt die Studie Aussagen über die generelle Wirksamkeit im Gegensatz zu einer unbehandelten Stichprobe beschränkt zu. Sie zeigt außerdem, dass KZT genauso wirksam wie PDT ist und in einigen wenigen Bereichen sogar überlegen.“</p>
17	<p>Barrowclough C., King P., Colville J., Russell E., Burns A., Tarrier N. ; Journal of Consulting and Clinical Psychology 2001, Vol. 69, No. 5, 756-762 A Randomized Trial of the Effectiveness of Cognitive-Behavioral Therapy and Supportive Counseling for Anxiety Symptoms in Older Adults</p>	<p>Insgesamt zu geringe Fallzahl bei heterogener Population. Kein signifikanter Beweis für den Effekt beider Therapieverfahren. .Wenn überhaupt, dann Anhaltspunkte für die Überlegenheit der CBT. Der Einfluss der medikamentösen Behandlung ist nicht einschätzbar kann aber nicht vernachlässigt werden.</p> <p>Studie ist nicht geeignet, Aussagen über die Wirksamkeit und den Nutzen von GT zu machen</p>	<p>Allgemeine methodische Qualität der Studie Die Methodik ist allgemein hinreichend gut (Diagnosestellung, Wahl der Messzeitpunkte, Wahl der Messinstrumente und statistische Auswertung sind befriedigend), allerdings wird die CBT eindeutig bevorzugt.</p> <p>Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken) Störungsspezifische CBT ist supportiver Beratung bei der Reduktion von Angst- und depressiver Symptomatik bei älteren Angstpatienten zu Behandlungsende in 2 von 5 Outcome-Maßen überlegen sowie zur Katamnese in 4 von 5 Outcome-Maßen.</p> <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen) Deutlich eingeschränkt. Alle Patienten nahmen vor Behandlungsbeginn auch Psychopharmaka, keine weiteren Angaben zur Medikation während der Behandlung, keine differenzierte Darstellung nach Gruppen. Die Supportive Beratung wurde von nur 1 Therapeut durchgeführt, die CBT nur von 2 Therapeuten. Die Behandlungen waren vergleichbar in Dauer und Intensität, die Gruppen waren auch zu Baseline vergleichbar, unterscheiden sich aber deutlich in den Drop-out Raten (bei CBT doppelt so hoch wie bei SC). Die CBT war störungsspezifisch, die supportive Beratung nicht.</p> <p>Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem Deutlich eingeschränkt. Die Stichprobe ist aufgrund des Alters (>= 55 Jahre) und der überwiegend geringen Schulbildung (95% verließen die Schule im 14.-16.Lj.) sehr selektiv. Die Therapien wurden bei den Patienten zu Hause durchgeführt, was in Deutschland nicht üblich ist.</p> <p>Stärken der Studie: Der Schwerpunkt auf Psychotherapie älterer Menschen ist interessant und generell ein wenig beforschtes Gebiet.</p> <p>Schwächen der Studie: Einige methodische Mängel schränken die Aussagekraft der Studie ein. Dies gilt insbesondere für die parallele Pharmakotherapie und</p>

Tragende Gründe – Supplement 1

			<p>die unterschiedliche hohen drop-outs zwischen den beiden Treatmentbedingungen</p> <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode</p> <p>Eingeschränkt relevant, da die Intervention lediglich als GPT II gewertet werden kann. Darüber hinaus erlaubt das Design und die methodischen Mängel keinen Wirksamkeitsnachweis.</p>
18	<p>Goldman RN, Greenberg LS, Angus L. The effects of adding emotion-focused interventions to the client-centered relationship conditions in the treatment of depression. <i>Psychotherapy Research</i> 2006; 16 (5): 536-46.</p>	<p>Kein Darstellung der basic characteristics getrennt nach Gruppen.</p> <p>Die numerische Ergebnisdarstellung der Veränderung der Depression (siehe Feld 20) ist nicht plausibel ($\Sigma n >$ Zahl der behandelten Patienten).</p> <p>Zusammenfassung von Daten zweier Studien (Studie 1998 kombiniert mit Studie 2006) zu einem Pool (combined Analysis) methodisch fragwürdig, da Unklarheiten bezüglich der Zusammenführung der Daten. Beobachtungsgleichheit nicht gewährleistet wegen zeitlichem Abstand.</p> <p>Viele Zielkriterien ohne Korrektur des alpha-Niveaus (Problem des multiplen Testens)</p> <p>Fallzahl auch bei combined analysis zu klein</p> <p>Die Aussage beide Verfahren seien hoch wirksam, kann nicht belegt werden, da Vergleich mit einer unbehandelten Kontrollgruppe fehlt.</p> <p>Keine Aussage über die Stabilität der Veränderungen möglich, da keine Katamnese</p> <p>Bedeutung für die ambulante Versorgung (Feld 24): Studie ist nicht geeignet, Aussagen über die Wirksamkeit und den Nutzen von GT zu machen</p>	<p>Allgemeine methodische Qualität der Studie</p> <p>Die allgemeine methodische Qualität der Studie ist mit wenigen Einschränkungen gut.</p> <p>Diagnosestellung und Veränderungsmessung erfolgt mittels reliabler und valider Verfahren. Die Qualität der Durchführung der beiden untersuchten Verfahren wird mittels geeigneter Verfahren überprüft. Die Stichprobe und daraus resultierende Power sind für differentielle Vergleiche etwas zu gering. Es gibt keine Katamnese-Erhebung.</p> <p>Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)</p> <p>Beide Therapien führen gleichermaßen zu klinisch bedeutsamen Verbesserungen der Symptomatik. In einzelnen Outcome-Maßen schneidet die EFT etwas besser ab (BDI, SCL-90-R)</p> <p>Anzumerken ist, dass die Studie zusätzlich integrativ mit der Kohorte von Greenberg & Watson, 1998 analysiert wurde. Diese Ergebnisse wurden bei der Bewertung nicht berücksichtigt.</p> <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen)</p> <p>Die interne Validität ist insgesamt gut. Die Effekte zur Post-Erhebung sind deutlich und können nicht allein auf Spontanremission zurückgeführt werden. Hervorzuheben ist, dass jeder Therapeut beide Verfahren durchführt; somit werden differentielle Effekte der Person des Therapeuten vermieden.</p>
			<p>Externe Validität (Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)</p> <p>Die externe Validität hinsichtlich der Patienten ist eingeschränkt, da die Patienten unter Annahme vieler Ausschlusskriterien über Anzeigenwerbung akquiriert wurden, was Selektionseffekte nicht ausschließen lässt. Die externe Validität hinsichtlich Dauer und Frequenz der Therapie ist gut. Einschränkungen ergeben sich aus einem umfangreichen Training und Supervision der Therapeuten. Es</p>

Tragende Gründe – Supplement 1

			<p>handelt sich um 14 unterschiedlich erfahrene Therapeuten, was zu einer guten Gesamtrepräsentativität führt.</p> <p>Stärken der Studie:</p> <ul style="list-style-type: none"> - Ausschaltung von event. Therapeuteneffekten, da jeder Therapeut beide Therapien durchführen - randomisierte Zuordnung <p>Schwächen der Studie:</p> <ul style="list-style-type: none"> - N und resultierenden Power sind zu gering für differentielle Fragestellungen. - keine Katamnese - viele Ausschlusskriterien d.h. Selektivität der Stichprobe anzunehmen - umfangreiche Trainingsmaßnahmen und Supervision der
			<p>Therapeuten</p> <p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode</p> <p>Die Studie wird als geeignet für den Nachweis der Wirksamkeit der GT bei depressiven affektiven Störungen angesehen.</p> <p>Obwohl der Vergleich der beiden Methoden einem Power-Problem unterliegt, sind die Prä-Post Effekte in beiden Gruppen überzeugend.</p>
19	<p>Machado LAC, Azevedo DC, Capanema MB, Neto TN, Cerceau DM. Client-Centered Therapy vs Exercise Therapy for Chronic Low Back Pain: A Pilot Randomized Controlled Trial in Brazil. Pain Medicine 2006; 8 (3): 251-58</p>	<p>Im Hinblick auf die Vergleichbarkeit der Behandlungsgruppen, differenzierte Darstellung der Ergebnisse und Dropout Raten sind die Ergebnisse der Studie mit großer Einschränkung zu interpretieren. Weiterhin war der Mittelwert der CCT lediglich bei 6 Sitzungen und der Übungsbehandlung bei 8.</p> <p>Also ist eine sehr geringe Therapieadhärenz zu konstatieren.</p> <p>Die Studie liefert keine Anhaltspunkte für eine Wirksamkeit der GT für die untersuchte Indikation</p>	<p>Allgemeine methodische Qualität der Studie schwach</p> <ul style="list-style-type: none"> - zu kleine Gruppengrößen (unzureichende statistische Power; offenbar substantielle wenn auch nicht statistisch signifikante Baseline-Unterschiede zwischen den Gruppen) <p>Der Anteil der Patienten, die die therapeutischen Angebote vollständig durchgeführt haben, ist extrem gering. Hinsichtlich der Outcome-variablen wurden keine signifikanten Unterschiede gefunden. Allerdings besteht konsistent der Trend, dass die Sport-/Bewegungsgruppe einen höheren Therapie- Benefit erzielte. - (ferner: keinerlei Angaben zu begleitenden Interventionen irrelevant: keine Überlegenheit des gesprächspsychotherapeutischen Ansatzes gegenüber Vergleichsbedingung nachweisbar; GT-Behandlung wurde allein von der Therapieadhärenz völlig unzureichend realisiert. zu kleine Datenbasis; zu hohe Abbruchquote.</p>
20	<p>Tarrier N, Yusupoff L, Kinney C, McCarthy E, Gledhill A, Haddock G, Morris J. Randomised controlled trial of intensive cognitive behaviour therapy for patients with chronic schizophrenia. BMJ 1998; 317 (7154) : 303-7.</p>	<p>Die Medikation, d.h. die korrekte Einnahme wurde während der Studie nicht kontrolliert. (methodischer Mangel)</p> <p>Die kognitive VT wurde als ergänzende, unterstützende Methode in der Behandlung der chronischen Schizophrenie angesehen.</p> <p>Nur 3 Therapeuten, die beide Therapieformen durchführten.</p> <p>Behandlungseffekte sind nicht eindeutig auf die Therapien zurückzuführen (Ansprechen der Medikation, Therapeutenmerkmale, Phase der Erkrankung, schwierige Einschätzung der Rückfälle)</p>	<p>Allgemeine methodische Qualität der Studie</p> <p>Hoher Aufwand, ordentliche Methodik: umfangreiches Screening und Auswahl von Patienten; klar definierte Ein- und Ausschlusskriterien; Intention-to-treat-Analysen durchgeführt</p> <p>Zusammenfassung der Ergebnisse</p> <p>(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken)</p>

Tragende Gründe – Supplement 1

		<p>Nur 18.5 % der gescreenten Patienten gingen in die Randomisierung ein</p> <p>Kein Wirksamkeitsnachweis</p>	<p>Die Behandlungsgruppen unterscheiden sich hinsichtlich der Reduktion der Positivsymptomatik statistisch signifikant. Im Einzelvergleich erweist sich aber nur die kognitiv-behaviorale Therapie der psychiatrischen Routinebehandlung als überlegen. Immerhin verbessern sich auch die Patienten mit supportiver Unterstützung signifikant im Gegensatz zu den ausschließlich Routinebehandelten, allerdings fällt der entsprechende Gruppenvergleich nicht signifikant aus.</p> <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen) etwas eingeschränkt gute Vergleichbarkeit der Gruppen und plausible Operationalisierung der Veränderungen aber unklar, ob unterschiedliche Effektivität der kognitiv-behavioralen versus supportiven Therapie wirklich auf unterschiedliche Behandlungskonzepte zurückzuführen ist, da die gesamten Therapien von nur 3 Therapeuten durchgeführt wurden und es unklar ist, wie die Behandlungen verteilt wurden und mit welcher Qualifikation und welchen Präferenzen die Behandlungen erfolgten.</p> <p>Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem etwas eingeschränkt relativ hohe Selektivität der Patienten (nur 18,5% der gescreenten Patienten gingen in die Randomisierung ein: Einschränkung auf psychotherapiebereite chronisch Schizophrene)</p> <p>Stärken der Studie: Umfassende Stichprobenerhebung und -beschreibung</p> <p>Schwächen der Studie: Nur 3 Therapeuten, die beide Psychotherapiearme abdecken; unklar ob Unterschiede der Behandlungseffektivität auf Therapeutenmerkmale oder Behandlungskonzepte zurückzuführen sind; fehlende Angaben zu Post-Hoc-Analysen.</p>
			<p>Bedeutung der Studie für die Bewertung der Psychotherapiemethode</p> <p>Studie liefert keinen Nachweis der Wirksamkeit, da Ergebnisse zum Behandlungsende nicht für klientenzentrierten Ansatz (GPT II) sprechen, sondern für kognitiv-behaviorale Therapie, die bezüglich der Reduktion der Positivsymptomatik überlegen scheint; allerdings entsteht der</p>

Tragende Gründe – Supplement 1

			Eindruck, dass diese Überlegenheit auf Präferenzen der Forscher zurück geführt werde könnte (im Einleitungsteil wird nur über kognitiv-behaviorale Therapie referiert).
21	<p>Tarrier N, Kinney C, McCarthy E, Humphreys L, Wittkowski A, Morris J. Two-year follow-up of cognitive-behavioral therapy and supportive counseling in the treatment of persistent symptoms in chronic schizophrenia. J Consult Clin Psychol 2000; 68 (5) : 917-22.</p>	<p>Hohe Anzahl von drop outs Relativ kleine Gruppen Wenig differenzierte Datenanalysen Eingeschränkte Verwertbarkeit - Supportive Therapie zeigt in 2-Jahreskatamnese höhere Effektivität, es ist eine unstrukturierte Therapieform; der individuelle Einfluss der Therapeuten ist nicht abschätzbar; ob die Ergebnisse auf die spezielle Therapieform zurückzuführen sind oder der Zeitfaktor die größere Rolle spielt ist nicht sicher zu sagen</p>	<p>Allgemeine methodische Qualität der Studie unbefriedigend: zwar hoher Aufwand bzgl. Der Stichprobenrekrutierung aber geringe Gruppengrößen, undifferenzierte Datenanalysen, unklare Ergebnisinterpretationsmöglichkeiten</p> <p>Zusammenfassung der Ergebnisse (Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken) Über den 2Jahres-Katamnesenzeitraum hinweg (3-Monats + 1-Jahres + 2-Jahreskatamnese) zeigten sich sowohl supportive Beratung als auch kognitiv-behaviorale Therapie der medizinisch-psychiatrischen Routinebehandlung überlegen (hinsichtlich Positiv- sowie Negativsymptomatik), wohingegen sich die beiden psychotherapeutischen Interventionen nicht signifikant voneinander unterschieden. Keine Unterschiede zwischen den Behandlungsgruppen fanden sich dagegen in Bezug auf die Häufigkeit von Rückfällen und die Medikamentendosierung. Hinsichtlich der Unterschiede von Prä nach 2-Jahres-Katamnese in Hinblick auf BDI, BHI und SFS zeigte die supportive Beratung die höchsten Effektstärken im Vergleich zu kognitiv-behavioraler Therapie und mehr noch zur medizinischen Routinebehandlung. Diese Überlegenheit der supportiven Beratung zum Zeitpunkt der 2-Jahres-Katamnese zeigte sich augenscheinlich (nach den Verlaufsgrafiken) auch in Bezug auf die schizophrene Positiv- und Negativsymptomatik</p> <p>Interne Validität (Sicherheit kausaler Schlussfolgerungen) Einschränkung: Verbesserung der Gruppe der supportiven Beratung zur 2-Jahres-Katamnese kann nicht eindeutig auf das Treatment zurückgeführt werden - Verbesserungen finden vor allem zwischen der 1-Jahres-Katamnese und der 2-Jahres-Katamnese statt - Kleine Gruppen, fehlende Drop-out-Analysen</p> <p>Externe Validität Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem</p> <p>Eingeschränkt: relativ hohe Selektivität der Patienten (nur 18,5% der gescreenten Patienten gingen in die Randomisierung und 13% in die Follow-up-Auswertungen ein: Ein-</p>

Tragende Gründe – Supplement 1

			<p>schränkung auf psychotherapiebereite, medikamentenstabile chronisch Schizophrene)</p> <p>Stärken der Studie: Umfassende Stichprobenerhebung und –beschreibung; multimodale Erfassung von Outcome-Kriterien (Fremdeinschätzungen, Selbsteinschätzungen, Informationen aus medizinischer Dokumentation: (Krankenhausaufenthaltstage, Medikamentendosierung)</p>
			<p>Schwächen der Studie: Nur 3 Therapeuten, die beide Psychotherapiearme abdecken: unklar ob Unterschiede der Behandlungseffektivität auf Therapeutenmerkmale oder Behandlungskonzepte zurückzuführen sind. Relativ kleine Gruppengrößen zur 2-Jahreskatamnese (20, 20, 21) Wenig differenzierte Datenanalysen (Summierung der Follow-up-Zeitpunkte, entweder nur deskriptive oder nur Inferenzstatistik. Bedeutung der Studie für die Bewertung der Psychotherapiemethode Eingeschränkt relevant: supportive Beratung (GPT II) zeigt sich zur 2-Jahres-Katamnese sowohl der Routinebehandlung als auch in einigen sekundären Zielkriterien der kognitiv-behavioralen Therapie überlegen, doch diese Effekte scheinen sich erst nach der 1-Jahres-Katamnese einzustellen und können daher nicht eindeutig auf das Treatment zurückgeführt werden. Zudem sind die Gruppen recht klein (n=20) und es wurden keine Drop-out-Analysen (immerhin 30%) vorgenommen)</p>
22	<p>Tarrier N, Haddock G, Lewis S, Drake R, Gregg L. Suicide behaviour over 18 months in recent onset schizophrenic patients: the effects of CBT. Schizophr Res.2006 Mar;83(1):15-27.Epub 2006 Feb 7.</p>	<p><i>Kein signifikanter Wirksamkeitsnachweis</i> Es bleibt fraglich, ob es sich hier um GT im Sinne der zu prüfenden Intervention handelt.</p>	<p>Stärken der Studie: Große Stichprobe, randomisierte Zuteilung, Concealment Schwächen der Studie: Nur 5 Therapeuten, die beide Psychotherapiearme abdecken; lediglich regional arbeitende Fremdeinschätzer, regional unterschiedliche Behandlungsdauer. Bedeutung der Studie für die Bewertung der Psychotherapiemethode Keine über die Publikation von Tarrier et al. (2004) hinausgehende Aussagekraft: unklare Datenanalysen, unklare Ein-/Ausschlusskriterien, hinsichtlich Outcome-Kriterium verübte Suizide/Suizidalität keine Überlegenheit für Supportives Counseling (+ Treatment-As-Usual) gegenüber Treatment-As-Usual</p>
23	<p>Tarrier N, Lewis S, Haddock G, Bentall R, Drake R, Kinderman P, Kingdon D, Siddler R, Everitt J,</p>	<p>Große Stichprobe, randomisierte Zuteilung, Concealment Nur 5 Therapeuten davon 2 Krankenschwestern, die beide Psy-</p>	<p>Stärken der Studie: Große Stichprobe, randomisierte Zuteilung, Concealment</p>

Tragende Gründe – Supplement 1

<p>Leadley K, Benn A, Grazebrook K, Haley C, Akhtar S, Davies L, Palmer S, Dunn G. Cognitive-behavioural therapy in first-episode and early schizophrenia. 18-month follow-up of a randomised controlled trial. Br J Psychiatry. 2004 Mar; 184:231-9.</p>	<p>chotherapiearme abdecken; lediglich regional arbeitende Fremdeinschätzer, regional unterschiedliche Behandlungsdauer. Einschränkungen ergeben sich aus den regionalen Unterschieden (nur in einer der drei Regionen waren die psychologischen Behandlungen effektiver als die Routinebehandlung), welches die Interpretation der Befunde erschwert. Interpretation der unterschiedlichen Ergebnisse zwischen den verschiedenen Regionen ist unklar. Nur bei einer der drei Regionen zeigten sich die Behandlungsgruppen der Routinebehandlung überlegen. Die Ausgangssymptomatik zwischen den Regionen war unterschiedlich; die Behandlungsdauer war zwischen den Regionen unterschiedlich. Unterschiede in der tatsächlichen Symptomschwere (Ausgangsniveau) und Unterschiede in der Behandlungsdauer. Die Analyse von Regionunterschieden erbrachte, dass signifikante Unterschiede im Behandlungserfolg nur für eine (Liverpool) der drei Regionen bestanden. Weiterhin wird das SC von den Autoren selbst als unspezifische und unstrukturierte Kontrollbedingung bezeichnet.</p> <p>Aufgrund der Limitierungen kaum belastbare Hinweise auf Wirksamkeit eines unstrukturierten SC auf psychotische Symptomatik bei nicht chronischer Schizophrenie.</p> <p>Es bleibt fraglich, ob es sich hier um GT im Sinne der zu prüfenden Intervention handelt.</p>	<p>Schwächen der Studie: Nur 5 Therapeuten, die beide Psychotherapiearme abdecken; lediglich regional arbeitende Fremdeinschätzer, regional unterschiedliche Behandlungsdauer. Bedeutung der Studie für die Bewertung der Psychotherapiemethode Hinweis auf Wirksamkeit der klientenzentrierten Intervention (GPZ II). Leichte Einschränkungen ergeben sich aus den regionalen Unterschieden (nur in einer der drei Regionen waren die psycholog. Behandlungen effektiver als die Routinebehandlung), welches die Interpretation der Befunde erschwert. Deutliche Effekte zeigen sich in einem Fremdeinschätzungsverfahren, hier gibt es für die GPT II signifikant stärkere Symptomverbesserungen. Für die dichotomen Outcome-Kriterien Relapse und Re Hospitalisierungsraten lassen sich keine Effekte nachweisen. Hierfür war vermutlich die statistische Power zu gering. Für die Wirksamkeit der klientenzentrierten Intervention spricht zudem, dass der klientenzentrierte Ansatz im Vergleich zur CBT nicht unterlegen ist, obwohl die Autoren (und Therapeuten!) Vertreter der CBT sind.</p>
--	--	---

Anhang

Auswertungsbögen der Expertenkommission der BPtK

(Auswertungsbögen der Themengruppe finden sich im Anhang des Themengruppenberichts bzw. des update-Berichts).

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

MT

Studie:

Autoren:

Altenhöfer et al.

Titel:

Psychotherapie von Anpassungsstörungen
Ist eine auf 12 Sitzungen begrenzte Gesprächspsychotherapie ausreichend wirksam?

Quelle:

Psychotherapeut, 2007, 52: 24-34

1. Basisinformationen

1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Das geschichtspsychotherapeutische Behandlungsangebot basierte für beide Patientengruppen auf dem „klientenzentrierten Konzept“ (Eckert 1996) nach Rogers, sowie auf einem eigens hierfür von Biermann-Ratjen und Eckert (Biermann-Ratjen u. Eckert 2001) ausgearbeiteten „Manual zur geschichtspsychotherapeutischen Behandlung von Anpassungsstörungen“ (Eckert 1996, Kap. 23).

1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Wirksamkeit von GT (12 Sitzungen) bei Anpassungsstörungen im Vergleich zu einer Wartegruppe.

Fragestellungen:

1. Wirkt sich die Behandlung mit einer zeitlich begrenzten Gesprächspsychotherapie für Patienten mit Anpassungsstörungen positiv aus?
2. Bleiben positive Veränderungen, die aus der psychotherapeutischen Behandlung resultieren, über den Therapiezeitraum hinaus erhalten, oder verschlechtern sich diese Patienten nach Abschluss der Therapie wieder deutlich?

1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Anpassungsstörung nach DSM-IV (SKID)

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews)	1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnosestellung	1 A2
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	1 C1

2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschlusskriterien:

- Anpassungsstörung nach DSM-IV als Reaktion auf 2 umschriebene belastende Lebensereignisse (Verlust einer wichtigen Bezugsperson, z. B. durch Todesfall, Trennung; Scheidung; Leistungsversagen mit finalen Konsequenzen in Studium oder Beruf, wie etwa Exmatrikulation oder Verlust des Arbeitsplatzes)
- Lebensalter zwischen 18 und 65 Jahren

Ausschlusskriterien:

- Anpassungsstörung als Reaktion auf andere (vergl. oben) Ereignisse
- akute Suizidalität, gravierende, komorbide psychische Erkrankungen, z. B. Psychosen, Essstörungen oder Abhängigkeitserkrankungen
- da Anpassungsstörungen nach DSM-IV nur diagnostiziert werden dürfen, wenn keine weiteren Störungen auf Achse I, affektive und Angststörungen ausgeschlossen

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	1 B1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	1 B2
K5. (C.3.)	Selektivität der Stichprobe aufgrund der Ausschlusskriterien	1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten 2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen) 3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien	1 C3

2.3 () Setting/Patientenzugang

- Patientenrekrutierung von Juni 2001 bis Mai 2004 deutschlandweit durch die Vermittlung bzw. Anmeldung von Patienten aus Beratungsstellen, Institutsambulanzen und niedergelassene Psychotherapeuten

	Kriterium	Operationalisierung	Rating
K6. (C.2.)	Art der Rekrutierung der Stichprobe	1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege 2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege 3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung)	1 C

2.4 (7) Zentrale Patientenmerkmale (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- 40 Frauen, 10 Männer
 - mittleres Alter 39,3 Jahre (SD ca. 11 Jahre)

- 47 deutsch, 3 andere
 - aus verschiedenen Bundesländern, am häufigsten Hamburg und Bayern
 - Familienstand (24 ledig, 8 verheiratet, 6 getrennt lebend, 4 geschieden, 8 verwitwet)
 - höchster Schulabschluss (Hauptschule 7, Realschule 10, Gymnasium 33)
 - aktuelle Berufstätigkeit: am häufigsten Vollzeit (19), Teilzeit (15) oder Ausbildung (8)

- ICD-10 Diagnose:
F43.20 Anpassungsstörung mit kurzer depressiver Reaktion (3)
F43.21 Anpassungsstörung mit längerer depressiver Reaktion (25)
F43.22 Anpassungsstörung mit Angst und depressiver Reaktion gemischt (13)
F43.23 Anpassungsstörung mit vorwiegender Beeinträchtigung von anderen Gefühlen (6)
F43.28 Anpassungsstörung mit sonstigen vorwiegend genannten Symptomen (3)

- Auslösendes Ereignis (Leistungsversagen 7, Todesfall 17, Trennung 26)
 - Behandlungsvorerfahrungen (keine 34, ambulante Langzeittherapie 4, ambulante Kurzzeittherapie 11, Psychiatrieaufenthalt 1)
 - Medikamente (nein 26, ja 24)

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Keine Angaben

	Kriterium	Operationalisierung	Rating
K7. (A.16.)	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	9 A16

2.6 (9) Anzahl gescreener, eingeschlossener, zugewiesener und ausgewerteter Patienten (Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- 59 Patienten gescreent
 - Zwei Patienten konnten aufgrund einer abweichenden Diagnose nicht in die Studie aufgenommen werden. Ein Patient schied innerhalb der Wartezeit aus, weil er an anderer Stelle sofort eine ambulante Psychotherapie beginnen konnte.

- 56 Patienten werden in die Studie eingeschlossen
 - keine Randomisierung aus ethischen Gründen, stattdessen wird zu-

nächst die Behandlungs- und im Anschluss die Wartegruppe aufgefüllt

- 6 Patienten brachen die Behandlung zu unterschiedlichen Zeitpunkten ab (unmittelbar nach der Erstdiagnostik, nach ein und nach 4 Therapie-
stunden).

- 50 Patienten können zum Prä-Post-Vergleich vollständig analysiert
werden (31 Behandlungsgruppe vs. 19 Wartegruppe)

- 16 Wartegruppenpatienten nehmen nach Ablauf der Wartezeit das
Therapieangebot an (2 Patienten entschieden sich für eine reguläre, von
den Krankenkassen finanzierte Psychotherapie, 1 Patientin wollte keine
psychotherapeutische Behandlung in Anspruch nehmen)

- zu einer Katamneseuntersuchung (16 Wochen nach Behandlung-
sende) erklären sich 41 Patienten bereit

- keine IIT-Analyse, Drop-out-Analyse

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 2
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 2
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A7 2
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	1 A3
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 2

2.7 (10) Vergleichbarkeit der Gruppen

<p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen: # Gruppen weitestgehend vergleichbar bzgl. der erhobenen sozialen und demographischen Faktoren, 2 signif. Unterschiede - 1. Zusammensetzung nach Herkunftsbundesland verschieden (Behandlungsgruppe mehr Hamburger, Wartegruppe mehr Bayern) - 2. Behandlungsvorerfahrung (Behandlungsgruppe überwiegend keine, Hälfte der Wartegruppe schon ambulante Therapie)</p> <p># Gruppen nicht vergleichbar bzgl. der Ausgangswerte in den Outcome-Maßen - Wartegruppe ist Prä weniger ängstlich und depressiv (auf zwei HADS-D Skalen), die Autoren berechnen d.h. korrigierte Effektstärken nach Klauer (1993)</p> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer): - ist gegeben (in beiden Gruppen 12 Therapiestunden, Minimum 8 Stunden, mittlere Stundenzahl 11,7 Stunden, Dauer im Mittel 16,1 Wochen, 1x/Woche)</p> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen): - ist gegeben</p> <p>Vergleichbarkeit der Therapieadhärenz: - mehr Dropouts in der Wartegruppe</p>

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	<p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p>	B9 3
K14. (A.18.)	Drop-out-Analysen	<p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>	A18 2

3. Intervention

**3.1 (14) Studiendesign /
Behandlungsarme**

Behandlungsgruppe mit Wartegruppe, Gesprächspsychotherapie

**3.2 (11) Behandlung Experimental-
gruppe
(Art, Dauer, Frequenz)**

Das gesprächspsychotherapeutische Behandlungsangebot basierte für beide Patientengruppen auf dem „klientenzentrierten Konzept“ (Eckert 1996) nach Rogers, sowie auf einem eigens hierfür von Biermann-Ratjen und Eckert (Biermann-Ratjen u. Eckert 2001) ausgearbeiteten „Manual zur gesprächspsychotherapeutischen Behandlung von Anpassungsstörungen“ (Eckert 1996, Kap. 23).

- in beiden Gruppen 12 Therapiestunden, Minimum 8 Stunden, mittlere Stundenzahl 11,7 Stunden, Dauer im Mittel 16,1 Wochen, 1x/Woche

**3.3 (12) Behandlung Kontroll-
gruppe
(Art, Dauer, Frequenz)**

**3.4 (13) Evtl. weitere Kontroll-
gruppen
(Art, Dauer, Frequenz)**

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 1
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	B4
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	B5 1
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	B6 3
K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	C4 1

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantziellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 2
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	C6 1
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	C7a 1
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	C7b 1
	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	C7c 1

**3.5 (19) Beobachtungszeitraum/
Messzeitpunkte/Katamnese**

Die diagnostische Einschätzung aller Psychotherapie- und Wartegruppenpatienten erfolgte mit Ausnahme des SKID und der Erfassung der soziodemographischen Daten, die nur zum ersten Erhebungszeitpunkt (prä) erhoben wurden, vor Behandlungs- bzw. Wartezeitbeginn (prä), am Ende der Behandlung bzw. Wartezeit (post) und 3 Monate nach Behandlungsende.

- in beiden Gruppen 12 Therapiestunden, Minimum 8 Stunden, mittlere Stundenzahl 11,7 Stunden, Dauer im Mittel 16,1 Wochen, 1x/Woche

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 1
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	B10
K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	B11 2

3.6 (15) Anzahl der Zentren

Keine Angabe

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...)

Keine Randomisierung aus ethischen Gründen; zunächst Behandlungs- dann Wartegruppe aufgefüllt

3.8 (17) ggf. Concealment der Randomisierung

Keine Randomisierung

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 9
K27. (B.8.)	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	B8 3

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

Belastung/ Beeinträchtigung durch psychische Symptomatik/ Subjektives Erleben/ allgemeines Funktionsniveau

- Allgemeine psychische Belastungen und Beeinträchtigungen „Brief symptom inventory“ (BSI; Franke 2000)
- Beeinträchtigungs-Schwere-Score (BSS; Schepank 1995)
- Skala zur globalen Erfassung des Funktionsniveaus (GAF-Skala): Achse V des DSM-IV
- Allgemeine Lebensqualität Fragebogen zur Lebenszufriedenheit (FLZ; Fahrenberg et al. 2000)

4.2 (21) Sekundäre Zielkriterien

Angst

- „State-trait-Angstinventar“ (STAI; Laux et al. 1981)
- Skala Ängstlichkeit - Allgemeine psychische Belastungen und Beeinträchtigungen „Brief symptom inventory“ (BSI; Franke 2000)
- Angstskala - „Hospital anxiety and depression scale“ (HADS; Herrmann et al. 1995)

Depressivität

- Skala Depressivität - Allgemeine psychische Belastungen und Beeinträchtigungen „Brief symptom inventory“ (BSI; Franke 2000)
- Allgemeine Depressionsskala (ADS; Hautzinger u. Bailer 1993)
- Depressionsskala - „Hospital anxiety and depression scale“ (HADS; Herrmann et al. 1995)

4.3 () Verblindung der Erhebung der Zielkriterien

Nicht eindeutig bekannt, aber anzunehmen (externe Diagnostiker)

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	<ol style="list-style-type: none"> 1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien, 	A7 2
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	<ol style="list-style-type: none"> 1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend 	1A8
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	<ol style="list-style-type: none"> 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung) 	1C9
K31. (A.9.)	Multiple Informationsquellen	<ol style="list-style-type: none"> 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle 	1A9

--	--	--	--

4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)

- Kontrolle Einfluss der Medikation: obwohl in Bezug auf die Häufigkeit der Einnahme von Medikamenten kein signifikanter Unterschied zwischen den Gruppen, wurde der mögliche Einfluss des Faktors Medikamenteneinnahme auf das Therapieergebnis mithilfe einer dreifaktoriellen Varianzanalyse kontrolliert

- Wirksamkeit der zeitlich begrenzten geschäftspsychotherapeutischen Behandlungen wurde aus 3 unterschiedlichen „Perspektiven“ betrachtet und berechnet:

1. durch gruppenstatistische Vergleiche (Varianzanalysen),
2. in Form von Effektstärken (korrigiert wg. unterschiedlicher Ausgangswerte) und
3. auf der Basis von individuellen, klinisch und statistisch signifikanten Verbesserungen

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	1 A6
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung	A13 2
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	A14 3

4.5 (22) Ergebnisse

signifikanter Einfluss der Medikamenteneinnahme auf den Therapieerfolg lediglich für das „rating“ auf der GAF-Skala des DSM-IV: Die externen Diagnostiker schätzten die positiven Veränderungen im allgemeinen Funktionsniveau bei den Patienten ohne Medikamenteneinnahme am Ende der Therapie höher ein als bei den Patienten mit Medikamenteneinnahme.

Prä-Post
Varianzanalyse mit Messwiederholung (2 Versuchsgruppen, 2 Messzeitpunkte); Interaktion zwischen Faktor Versuchsgruppe (Behandlungs- vs. Wartegruppe) und Faktor Messzeitpunkt:
- hochsignifikante Verbesserungen der Therapiegruppe im Vergleich zur Wartegruppe in den folgenden Maßen: dem Beeinträchtigungsschwere-Score (BSS), der Skala zur globalen Erfassung des Funktionsniveaus (GAF-Skala des DSMIV), „global severity index“ (GSI) des „brief symptom inventory“ (BSI), allgemeine Depressionsskala (ADS), HADS-Depression

durchgängig mittlere bis große Effektstärken

Follow-up
Für die Gesamtpatientengruppe und auch für die reinen Behandlungsgruppenpatienten wurden über alle drei Erhebungszeitpunkte hinweg signifikante lineare Trends in allen Fragebögen und Rating-Skalen gefunden, d. h. die positiven Behandlungseffekte blieben auch über die Therapie hinaus stabil. Signifikante Unterschiede zwischen dem ersten

(prä) und zweiten (post) sowie zwischen dem ersten und dritten Erhebungszeitpunkt (Katamnese). Für keinen der Tests wurden jedoch signifikante Unterschiede zwischen dem zweiten und dritten Erhebungszeitpunkt gefunden.

individueller Therapieerfolg

- Auch auf der individuellen Ebene verbesserten sich die Psychotherapiepatienten in allen Erhebungsinstrumenten statistisch und klinisch signifikant, während dies bei den Patienten der Wartegruppe nur in ganz geringem Umfang der Fall war. Die meisten nach Klauer (1993) berechneten korrigierten Effektstärken, die die teilweise unterschiedlichen Ausgangswerte der Psychotherapie- und Wartegruppenpatienten berücksichtigten, lassen sich als große Effekte interpretieren.

	Kriterium	Operationalisierung	Rating
K36. (A.11.)	Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten	1) für alle Outcome-Variablen berichtet 2) ausschließlich für die primären Zielkriterien berichtet 3) nicht für alle primären Zielkriterien berichtet	A11 1
K37. (B.12.)	Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte)	1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe 2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig 3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG	B12 1

4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen

- keine

	Kriterium	Operationalisierung	Rating
K38. (A.12.)	Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen	1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen 3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet	A12 1

4.7 (24) Fazit/Schlussfolgerungen der Autoren

Eine auf 12 Sitzungen begrenzte Gesprächspsychotherapie hat sich in dieser Untersuchung als wirksame und effektive Behandlung von Patienten mit Anpassungsstörungen erwiesen. Trotz des begrenzten Stundenumfangs erzielten die Psychotherapiepatienten im gruppenstatistischen Vergleich mit den Patienten der Wartegruppe deutliche Verbesserungen im Hinblick auf ihre psychische Symptomatik, ihr subjektives Erleben und ihr allgemeines Funktionsniveau. Diese positiven Veränderungen blieben auch im Katamnesezeitraum stabil. Auch auf der individuellen Ebene verbesserten sich die Psychotherapiepatienten in allen Erhebungsinstrumenten statistisch und klinisch signifikant, während dies bei den Patienten der Wartegruppe nur in ganz geringem Umfang der Fall war.

	Kriterium	Operationalisierung	Rating
K39. (A.1.)	Manipulation der Daten	1) keine Hinweise auf Ergebnismanipulation 3) Hinweise auf Ergebnismanipulation	A1

5. Beurteilung der Ergebnisse der Studie		
D1	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual". (Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt) (ja = 1; nein = 3; nicht zutreffend = freilassen)	D1 1
D2	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D2
D3	In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D3

6. Praxistransfer

	Kriterium	Operationalisierung	Rating
K40. (C.10.)	Spezifikation und Herstellbarkeit notwendiger Settingbedingungen	1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team) 2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team) 3) Notwendige Settingbedingungen nicht herstellbar	C10 1
K41. (C.11.)	Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation	1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar 2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar 3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar	C1 1
K42. (C.12.)	Spezifikation und Erfassbarkeit relevanter Patientenmerkmale	1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar 2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar 3) Relevante Patientenmerkmale praktisch nicht erfassbar	C12 1
K43. (C.13.)	Spezifikation und Herstellbarkeit relevanter Treatmentaspekte	1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual) 2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung) 3) Relevante Treatmentmerkmale praktisch nicht herstellbar	C13 1

7. Gesamtbewertung der Studie

Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)			
Durchschnittswert für B) Interne Validität (Summe / 12)			
Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)			
Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4)			

Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

<p>Kriterien für 'A. Allgemeine methodische Qualität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde <p>und</p> <ol style="list-style-type: none"> wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist. 	
<p>Kriterien für 'B. Interne Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist. 	
<p>Kriterien für 'C. Externe Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde. 	
<p>Die Ergebnisse der Studie belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>	

Abschließende Bewertung

Allgemeine methodische Qualität der Studie

Die allgemeine methodische Qualität der Studie ist hoch. Es erfolgt eine objektive und reliable Diagnosestellung. Die statistische Power ist für den Vergleich mit einer unbehandelten Gruppe als hoch einzuschätzen. Die gewählten Outcome-Verfahren sind reliabel und valide. Alle Ergebnisse werden vollständig berichtet.

Es gibt statistische Mängel; Therapiegruppe und Wartegruppe haben ein unterschiedliches Ausgangsniveau in den Outcome-Variablen. Dem hätte man nicht nur durch eine korrigierte Effektgröße, sondern auch durch die Bildung entsprechender Kovariaten gerecht werden können.

Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

konsistente Befunde

hochsignifikante Verbesserungen der Therapiegruppe im Vergleich zur Wartegruppe in mehreren Maßen

Verbesserungen sind auch klinisch relevant

Verbesserungen sind zur 3-Monats-Katamnese stabil

keine Nebenwirkungen/ Risiken bekannt

Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Ein- und Ausschlusskriterien sind klar spezifiziert; es gibt keine Selektionseffekte durch selektiven Ein- und Ausschluss. Es gibt ein ausführliches Therapiemanual, die Manualtreue wird allerdings nicht belegt.

Problematisch sind die Unterschiede in den Ausgangswerten durch den Verzicht auf randomisierte Zuordnung. Eine angemessene statistische Kontrolle dieser Unterschiede fehlt (z.B. Prä-Werte als Kovariate). Die Katamnese erfolgt mit 3 Monaten sehr zeitnah zum Therapieende, was aber bei Anpassungsstörungen möglicherweise ein gut gewählter Zeitraum ist.

Externe Validität

(Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)

Die externe Validität ist als hoch zu bewerten. Die Patientenstichprobe ist nicht selektiv, die Repräsentativität der Behandler und der Intervention ist hoch. Das Therapeutenverhalten wird nicht durch Supervision verändert. Die Frequenz (2x/ Woche) der Intervention ist repräsentativ und gut auf ein ambulantes Setting im dt. Versorgungssystem übertragbar.

Stärken der Studie:

Die methodische Qualität und externe Validität sind hoch. Die Therapeuten sind erfahren und verfügen über einen Abschluss in GT.

Schwächen der Studie:

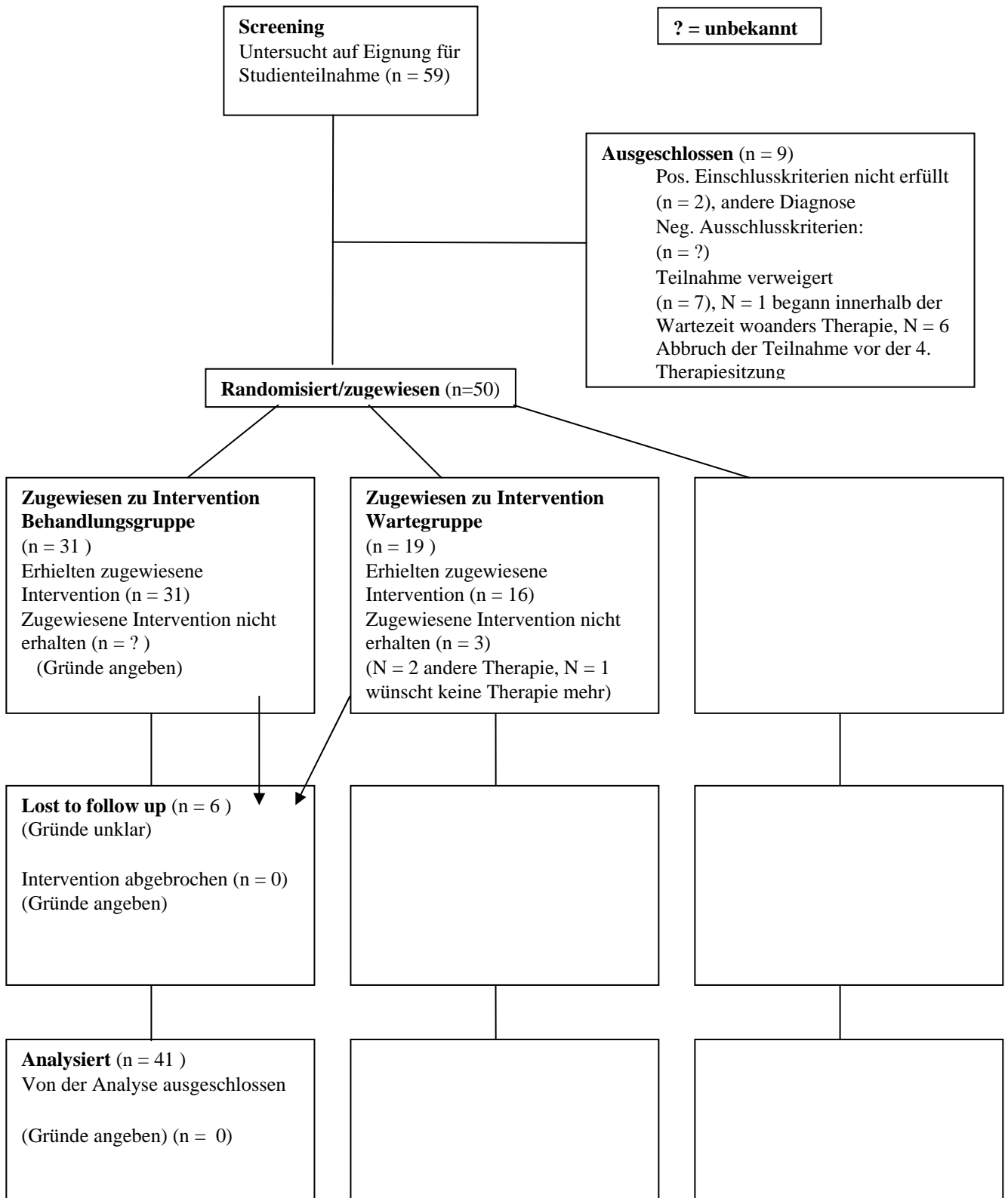
Die Zuordnung zu den Gruppen erfolgt aus ethischen Gründen nicht-randomisiert, sondern nach Reihenfolge des Einschlusses und der Zumutbarkeit einer Wartezeit, daher ist die Wartegruppe zum Zeitpunkt prä im Mittel weniger stark betroffen als die Therapiegruppe.

Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie wird als geeignet für den Nachweis der Wirksamkeit der GT bei Anpassungsstörungen angesehen.

Die Behandlungsgruppe zeigt statistisch und klinisch bedeutsame Verbesserungen der Störung im Vergleich zu einer Wartegruppe. Bei ansonsten hohem methodischen Niveau, ist die nicht erfolgte Randomisierung der Zuordnung problematisch.

8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

K5 (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

K14 (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

K33 (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

K39 (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

Anmerkungen zur durchgeführten Studienbewertungen

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

MT

Studie:

Autoren:

Altenhöfer et al.

Gorschenek et al. (Katamnese)

Titel:

Psychotherapie von Anpassungsstörungen
Ist eine auf 12 Sitzungen begrenzte Gesprächspsychotherapie ausreichend wirksam?

Psychotherapie von Anpassungsstörungen
Ist eine auf 12 Sitzungen begrenzte Gesprächspsychotherapie auch langfristig ausreichend wirksam?

Quelle:

Psychotherapeut, 2007, 52: 24-34

Psychother Psych Med, 2007, 57: 1-8

1. Basisinformationen

1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Das gesprächspsychotherapeutische Behandlungsangebot basierte für beide Patientengruppen auf dem „klientenzentrierten Konzept“ (Eckert 1996) nach Rogers, sowie auf einem eigens hierfür von Biermann-Ratjen und Eckert (Biermann-Ratjen u. Eckert 2001) ausgearbeiteten „Manual zur gesprächspsychotherapeutischen Behandlung von Anpassungsstörungen“ (Eckert 1996, Kap. 23).

1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Wirksamkeit von GT (12 Sitzungen) bei Anpassungsstörungen im Vergleich zu einer Wartegruppe.

Fragestellungen:

1. Wirkt sich die Behandlung mit einer zeitlich begrenzten Gesprächspsychotherapie für Patienten mit Anpassungsstörungen positiv aus?
2. Bleiben positive Veränderungen, die aus der psychotherapeutischen Behandlung resultieren, über den Therapiezeitraum hinaus erhalten, oder verschlechtern sich diese Patienten nach Abschluss der Therapie wieder deutlich?
3. Bleiben die Effekte zur 2-Jahres-Katamnese stabil und gibt es einen posttherapeutischen Gewinn?

1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Anpassungsstörung nach DSM-IV (SKID)

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews)	1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnosestellung	1 A2
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	1 C1

2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschlusskriterien:

- Anpassungsstörung nach DSM-IV als Reaktion auf 2 umschriebene belastende Lebensereignisse (Verlust einer wichtigen Bezugsperson, z. B. durch Todesfall, Trennung; Scheidung; Leistungsversagen mit finalen Konsequenzen in Studium oder Beruf, wie etwa Exmatrikulation oder Verlust des Arbeitsplatzes)
- Lebensalter zwischen 18 und 65 Jahren

Ausschlusskriterien:

- Anpassungsstörung als Reaktion auf andere (vergl. oben) Ereignisse
- akute Suizidalität, gravierende, komorbide psychische Erkrankungen, z. B. Psychosen, Essstörungen oder Abhängigkeitserkrankungen
- da Anpassungsstörungen nach DSM-IV nur diagnostiziert werden dürfen, wenn keine weiteren Störungen auf Achse I, affektive und Angststörungen ausgeschlossen

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	1 B1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	1 B2
K5. (C.3.)	Selektivität der Stichprobe aufgrund der Ausschlusskriterien	1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten 2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen) 3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien	1 C3

2.3 () Setting/Patientenzugang

- Patientenrekrutierung von Juni 2001 bis September 2003 deutschlandweit durch die Vermittlung bzw. Anmeldung von Patienten aus Beratungsstellen, Institutsambulanzen und niedergelassene Psychotherapeuten

	Kriterium	Operationalisierung	Rating
K6. (C.2.)	Art der Rekrutierung der Stichprobe	1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege 2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege 3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung)	1 C

2.4 (7) Zentrale Patientenmerkmale (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- 40 Frauen, 10 Männer
- mittleres Alter 39,3 Jahre (SD ca. 11 Jahre)

- 47 deutsch, 3 andere
- aus verschiedenen Bundesländern, am häufigsten Hamburg und Bayern
- Familienstand (24 ledig, 8 verheiratet, 6 getrennt lebend, 4 geschieden, 8 verwitwet)
- höchster Schulabschluss (Hauptschule 7, Realschule 10, Gymnasium 33)
- aktuelle Berufstätigkeit: am häufigsten Vollzeit (19), Teilzeit (15) oder Ausbildung (8)

- ICD-10 Diagnose:
F43.20 Anpassungsstörung mit kurzer depressiver Reaktion (3)
F43.21 Anpassungsstörung mit längerer depressiver Reaktion (25)
F43.22 Anpassungsstörung mit Angst und depressiver Reaktion gemischt (13)
F43.23 Anpassungsstörung mit vorwiegender Beeinträchtigung von anderen Gefühlen (6)
F43.28 Anpassungsstörung mit sonstigen vorwiegend genannten Symptomen (3)

- Auslösendes Ereignis (Leistungsversagen 7, Todesfall 17, Trennung 26)
- Behandlungsvorerfahrungen (keine 34, ambulante Langzeittherapie 4, ambulante Kurzzeittherapie 11, Psychiatrieaufenthalt 1)
- Medikamente (nein 26, ja 24)

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Keine Angaben

	Kriterium	Operationalisierung	Rating
K7. (A.16.)	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	9 A16

2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten (Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- 59 Patienten gescreent
- Zwei Patienten konnten aufgrund einer abweichenden Diagnose nicht in die Studie aufgenommen werden. Ein Patient schied innerhalb der Wartezeit aus, weil er an anderer Stelle sofort eine ambulante Psychotherapie beginnen konnte.
- 56 Patienten werden in die Studie eingeschlossen

- keine Randomisierung aus ethischen Gründen, stattdessen wird zunächst die Behandlungs- und im Anschluss die Wartegruppe aufgefüllt
- 6 Patienten brachen die Behandlung zu unterschiedlichen Zeitpunkten ab (unmittelbar nach der Erstdiagnostik, nach ein und nach 4 Therapiestunden).
- 50 Patienten können zum Prä-Post-Vergleich vollständig analysiert werden (31 Behandlungsgruppe vs. 19 Wartegruppe)
- 16 Wartegruppenpatienten nehmen nach Ablauf der Wartezeit das Therapieangebot an (2 Patienten entschieden sich für eine reguläre, von den Krankenkassen finanzierte Psychotherapie, 1 Patientin wollte keine psychotherapeutische Behandlung in Anspruch nehmen)
- zu einer Katamneseuntersuchung (16 Wochen nach Behandlungsende) erklärten sich 41 Patienten bereit
- weitere Katamneseerhebung (ohne Fremdurteil) nach 2 Jahren (Gorschenek et al); Rücklauf 35 Patienten; 25 davon analysiert (10 Patienten hatten im Katamneseintervall weitere Behandlung in Anspruch genommen)
- keine IIT-Analyse, Drop-out-Analyse

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 2
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 2
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A7 2
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	1 A3
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 2

2.7 (10) Vergleichbarkeit der Gruppen

<p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen:</p> <p># Gruppen weitestgehend vergleichbar bzgl. der erhobenen sozialen und demographischen Faktoren, 2 signif. Unterschiede</p> <ul style="list-style-type: none"> - 1. Zusammensetzung nach Herkunftsbundesland verschieden (Behandlungsgruppe mehr Hamburger, Wartegruppe mehr Bayern) - 2. Behandlungsvorerfahrung (Behandlungsgruppe überwiegend keine, Hälfte der Wartegruppe schon ambulante Therapie) <p># Gruppen nicht vergleichbar bzgl. der Ausgangswerte in den Outcome-Maßen</p> <ul style="list-style-type: none"> - Wartegruppe ist Prä weniger ängstlich und depressiv (auf zwei HADS-D Skalen), die Autoren berechnen d.h. korrigierte Effektstärken nach Klauer (1993) <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer):</p> <ul style="list-style-type: none"> - ist gegeben (in beiden Gruppen 12 Therapiestunden, Minimum 8 Stunden, mittlere Stundenzahl 11,7 Stunden, Dauer im Mittel 16,1 Wochen, 1x/Woche) <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen):</p> <ul style="list-style-type: none"> - ist gegeben <p>Vergleichbarkeit der Therapieadhärenz:</p> <ul style="list-style-type: none"> - mehr Dropouts in der Wartegruppe

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	<p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p>	B9 3
K14. (A.18.)	Drop-out-Analysen	<p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>	A18 2

3. Intervention

**3.1 (14) Studiendesign /
Behandlungsarme**

Behandlungsgruppe mit Wartegruppe, Gesprächspsychotherapie

**3.2 (11) Behandlung Experimental-
gruppe
(Art, Dauer, Frequenz)**

Das gesprächspsychotherapeutische Behandlungsangebot basierte für beide Patientengruppen auf dem „klientenzentrierten Konzept“ (Eckert 1996) nach Rogers, sowie auf einem eigens hierfür von Biermann-Ratjen und Eckert (Biermann-Ratjen u. Eckert 2001) ausgearbeiteten „Manual zur gesprächspsychotherapeutischen Behandlung von Anpassungsstörungen“ (Eckert 1996, Kap. 23).

- in beiden Gruppen 12 Therapiestunden, Minimum 8 Stunden, mittlere Stundenzahl 11,7 Stunden, Dauer im Mittel 16,1 Wochen, 1x/Woche

**3.3 (12) Behandlung Kontroll-
gruppe
(Art, Dauer, Frequenz)**

**3.4 (13) Evtl. weitere Kontroll-
gruppen
(Art, Dauer, Frequenz)**

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 1
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	B4
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	B5 1
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	B6 3
K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	C4 1

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantziellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 2
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	C6 1
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	C7a 1
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	C7b 1
	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	C7c 1

**3.5 (19) Beobachtungszeitraum/
Messzeitpunkte/Katamnese**

Die diagnostische Einschätzung aller Psychotherapie- und Wartegruppenpatienten erfolgte mit Ausnahme des SKID und der Erfassung der soziodemographischen Daten, die nur zum ersten Erhebungszeitpunkt (prä) erhoben wurden, vor Behandlungs- bzw. Wartezeitbeginn (prä), am Ende der Behandlung bzw. Wartezeit (post) und 3 Monate nach Behandlungsende. Zwei Jahr nach Therapieende erfolgte eine weitere Katamneseuntersuchung (Gorschenek et al.).

- in beiden Gruppen 12 Therapiestunden, Minimum 8 Stunden, mittlere Stundenzahl 11,7 Stunden, Dauer im Mittel 16,1 Wochen, 1x/Woche

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 1
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	B10
K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	B11 1

3.6 (15) Anzahl der Zentren

Keine Angabe

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...)

Keine Randomisierung aus ethischen Gründen; zunächst Behandlungs- dann Wartegruppe aufgefüllt

3.8 (17) ggf. Concealment der Randomisierung

Keine Randomisierung

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 9
K27. (B.8.)	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	B8 3

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

Belastung/ Beeinträchtigung durch psychische Symptomatik/ Subjektives Erleben/ allgemeines Funktionsniveau

- Allgemeine psychische Belastungen und Beeinträchtigungen „Brief symptom inventory“ (BSI; Franke 2000)
- Beeinträchtigungs-Schwere-Score (BSS; Schepank 1995)
- Skala zur globalen Erfassung des Funktionsniveaus (GAF-Skala): Achse V des DSM-IV
- Allgemeine Lebensqualität Fragebogen zur Lebenszufriedenheit (FLZ; Fahrenberg et al. 2000)

4.2 (21) Sekundäre Zielkriterien

Angst

- „State-trait-Angstinventar“ (STAI; Laux et al. 1981)
- Skala Ängstlichkeit - Allgemeine psychische Belastungen und Beeinträchtigungen „Brief symptom inventory“ (BSI; Franke 2000)
- Angstskala - „Hospital anxiety and depression scale“ (HADS; Herrmann et al. 1995)

Depressivität

- Skala Depressivität - Allgemeine psychische Belastungen und Beeinträchtigungen „Brief symptom inventory“ (BSI; Franke 2000)
- Allgemeine Depressionsskala (ADS; Hautzinger u. Bailer 1993)
- Depressionsskala - „Hospital anxiety and depression scale“ (HADS; Herrmann et al. 1995)

4.3 () Verblindung der Erhebung der Zielkriterien

Nicht eindeutig bekannt, aber anzunehmen (externe Diagnostiker)

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	<ol style="list-style-type: none"> 1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien, 	A7 2
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	<ol style="list-style-type: none"> 1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend 	1A8
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	<ol style="list-style-type: none"> 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung) 	1C9
K31. (A.9.)	Multiple Informationsquellen	<ol style="list-style-type: none"> 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle 	1A9

--	--	--	--

4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)

- Kontrolle Einfluss der Medikation: obwohl in Bezug auf die Häufigkeit der Einnahme von Medikamenten kein signifikanter Unterschied zwischen den Gruppen, wurde der mögliche Einfluss des Faktors Medikamenteneinnahme auf das Therapieergebnis mithilfe einer dreifaktoriellen Varianzanalyse kontrolliert

- Wirksamkeit der zeitlich begrenzten geschäftspsychotherapeutischen Behandlungen wurde aus 3 unterschiedlichen „Perspektiven“ betrachtet und berechnet:
 1. durch gruppenstatistische Vergleiche (Varianzanalysen),
 2. in Form von Effektstärken (korrigiert wg. unterschiedlicher Ausgangswerte) und
 3. auf der Basis von individuellen, klinisch und statistisch signifikanten Verbesserungen

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	1 A6
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung	A13 2
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	A14 1

4.5 (22) Ergebnisse

signifikanter Einfluss der Medikamenteneinnahme auf den Therapieerfolg lediglich für das „rating“ auf der GAF-Skala des DSM-IV: Die externen Diagnostiker schätzten die positiven Veränderungen im allgemeinen Funktionsniveau bei den Patienten ohne Medikamenteneinnahme am Ende der Therapie höher ein als bei den Patienten mit Medikamenteneinnahme.

Prä-Post
 Varianzanalyse mit Messwiederholung (2 Versuchsgruppen, 2 Messzeitpunkte); Interaktion zwischen Faktor Versuchsgruppe (Behandlungs- vs. Wartegruppe) und Faktor Messzeitpunkt:
 - hochsignifikante Verbesserungen der Therapiegruppe im Vergleich zur Wartegruppe in den folgenden Maßen: dem Beeinträchtigungsschwere-Score (BSS), der Skala zur globalen Erfassung des Funktionsniveaus (GAF-Skala des DSMIV), „global severity index“ (GSI) des „brief symptom inventory“ (BSI), allgemeine Depressionsskala (ADS), HADS-Depression

durchgängig mittlere bis große Effektstärken

Follow-up
 Für die Gesamtpatientengruppe und auch für die reinen Behandlungsgruppenpatienten wurden über alle drei Erhebungszeitpunkte hinweg signifikante lineare Trends in allen Fragebögen und Rating-Skalen gefunden, d. h. die positiven Behandlungseffekte blieben auch über die Therapie hinaus stabil. Signifikante Unterschiede zwischen dem ersten

(prä) und zweiten (post) sowie zwischen dem ersten und dritten Erhebungszeitpunkt (Katamnese). Für keinen der Tests wurden jedoch signifikante Unterschiede zwischen dem zweiten und dritten Erhebungszeitpunkt gefunden.

Die Effekte bleiben im wesentlichen stabil zur 2-Jahres-Katamnese.

individueller Therapieerfolg

- Auch auf der individuellen Ebene verbesserten sich die Psychotherapiepatienten in allen Erhebungsinstrumenten statistisch und klinisch signifikant, während dies bei den Patienten der Wartegruppe nur in ganz geringem Umfang der Fall war. Die meisten nach Klauer (1993) berechneten korrigierten Effektstärken, die die teilweise unterschiedlichen Ausgangswerte der Psychotherapie- und Wartegruppenpatienten berücksichtigten, lassen sich als große Effekte interpretieren.

	Kriterium	Operationalisierung	Rating
K36. (A.11.)	Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten	1) für alle Outcome-Variablen berichtet 2) ausschließlich für die primären Zielkriterien berichtet 3) nicht für alle primären Zielkriterien berichtet	A11 1
K37. (B.12.)	Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte)	1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe 2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig 3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG	B12 1

4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen

- keine

	Kriterium	Operationalisierung	Rating
K38. (A.12.)	Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen	1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen 3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet	A12 1

4.7 (24) Fazit/Schlussfolgerungen der Autoren

Eine auf 12 Sitzungen begrenzte Gesprächspsychotherapie hat sich in dieser Untersuchung als wirksame und effektive Behandlung von Patienten mit Anpassungsstörungen erwiesen. Trotz des begrenzten Stundenumfangs erzielten die Psychotherapiepatienten im gruppenstatistischen Vergleich mit den Patienten der Wartegruppe deutliche Verbesserungen im Hinblick auf ihre psychische Symptomatik, ihr subjektives Erleben und ihr allgemeines Funktionsniveau. Diese positiven Veränderungen blieben auch im Katamnesezeitraum stabil. Auch auf der individuellen Ebene verbesserten sich die Psychotherapiepatienten in allen Erhebungsinstrumenten statistisch und klinisch signifikant, während dies bei den Patienten der Wartegruppe nur in ganz geringem Umfang der Fall war. 10 Patienten suchten nach Abschluss der 12 Sitzungen weitere Therapie auf. Für ca. ein Drittel der Patienten war eine Therapielänge von 12 Sitzungen also nicht ausreichend.

	Kriterium	Operationalisierung	Rating
K39. (A.1.)	Manipulation der Daten	1) keine Hinweise auf Ergebnismanipulation 3) Hinweise auf Ergebnismanipulation	A1 1

5. Beurteilung der Ergebnisse der Studie		
D1	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual". (Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt) (ja = 1; nein = 3; nicht zutreffend = freilassen)	D1 1
D2	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D2
D3	In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D3

6. Praxistransfer

	Kriterium	Operationalisierung	Rating
K40. (C.10.)	Spezifikation und Herstellbarkeit notwendiger Settingbedingungen	1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team) 2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team) 3) Notwendige Settingbedingungen nicht herstellbar	C10 1
K41. (C.11.)	Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation	1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar 2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar 3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar	C1 1
K42. (C.12.)	Spezifikation und Erfassbarkeit relevanter Patientenmerkmale	1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar 2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar 3) Relevante Patientenmerkmale praktisch nicht erfassbar	C12 1
K43. (C.13.)	Spezifikation und Herstellbarkeit relevanter Treatmentaspekte	1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual) 2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung) 3) Relevante Treatmentmerkmale praktisch nicht herstellbar	C13 1

7. Gesamtbewertung der Studie

Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)			
Durchschnittswert für B) Interne Validität (Summe / 12)			
Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)			
Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4)			

Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

<p>Kriterien für 'A. Allgemeine methodische Qualität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde <p>und</p> <ol style="list-style-type: none"> wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist. 	
<p>Kriterien für 'B. Interne Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist. 	
<p>Kriterien für 'C. Externe Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde. 	
<p>Die Ergebnisse der Studie belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>	

Abschließende Bewertung

Allgemeine methodische Qualität der Studie

Die allgemeine methodische Qualität der Studie ist hoch. Es erfolgt eine objektive und reliable Diagnosestellung. Die statistische Power ist für den Vergleich mit einer unbehandelten Gruppe als hoch einzuschätzen. Die gewählten Outcome-Verfahren sind reliabel und valide. Alle Ergebnisse werden vollständig berichtet.

Es gibt statistische Mängel; Therapiegruppe und Wartegruppe haben ein unterschiedliches Ausgangsniveau in den Outcome-Variablen. Dem hätte man nicht nur durch eine korrigierte Effektgröße, sondern auch durch die Bildung entsprechender Kovariaten gerecht werden können.

Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

konsistente Befunde

hochsignifikante Verbesserungen der Therapiegruppe im Vergleich zur Wartegruppe in mehreren Maßen

Verbesserungen sind auch klinisch relevant

Verbesserungen sind zur Katamnese (3 Monate, 2 Jahre stabil)

keine Nebenwirkungen/ Risiken bekannt

Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Ein- und Ausschlusskriterien sind klar spezifiziert; es gibt keine Selektionseffekte durch selektiven Ein- und Ausschluss. Es gibt ein ausführliches Therapiemanual, die Manualtreue wird allerdings nicht belegt.

Problematisch sind die Unterschiede in den Ausgangswerten durch den Verzicht auf randomisierte Zuordnung. Eine angemessene statistische Kontrolle dieser Unterschiede fehlt (z.B. Prä-Werte als Kovariate).

Externe Validität

(Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)

Die externe Validität ist als hoch zu bewerten. Die Patientenstichprobe ist nicht selektiv, die Repräsentativität der Behandler und der Intervention ist hoch. Das Therapeutenverhalten wird nicht durch Supervision verändert. Die Frequenz (2x/ Woche) der Intervention ist repräsentativ und gut auf ein ambulantes Setting im dt. Versorgungssystem übertragbar.

Stärken der Studie:

Die methodische Qualität und externe Validität sind hoch. Die Therapeuten sind erfahren und verfügen über einen Abschluss in GT.

Schwächen der Studie:

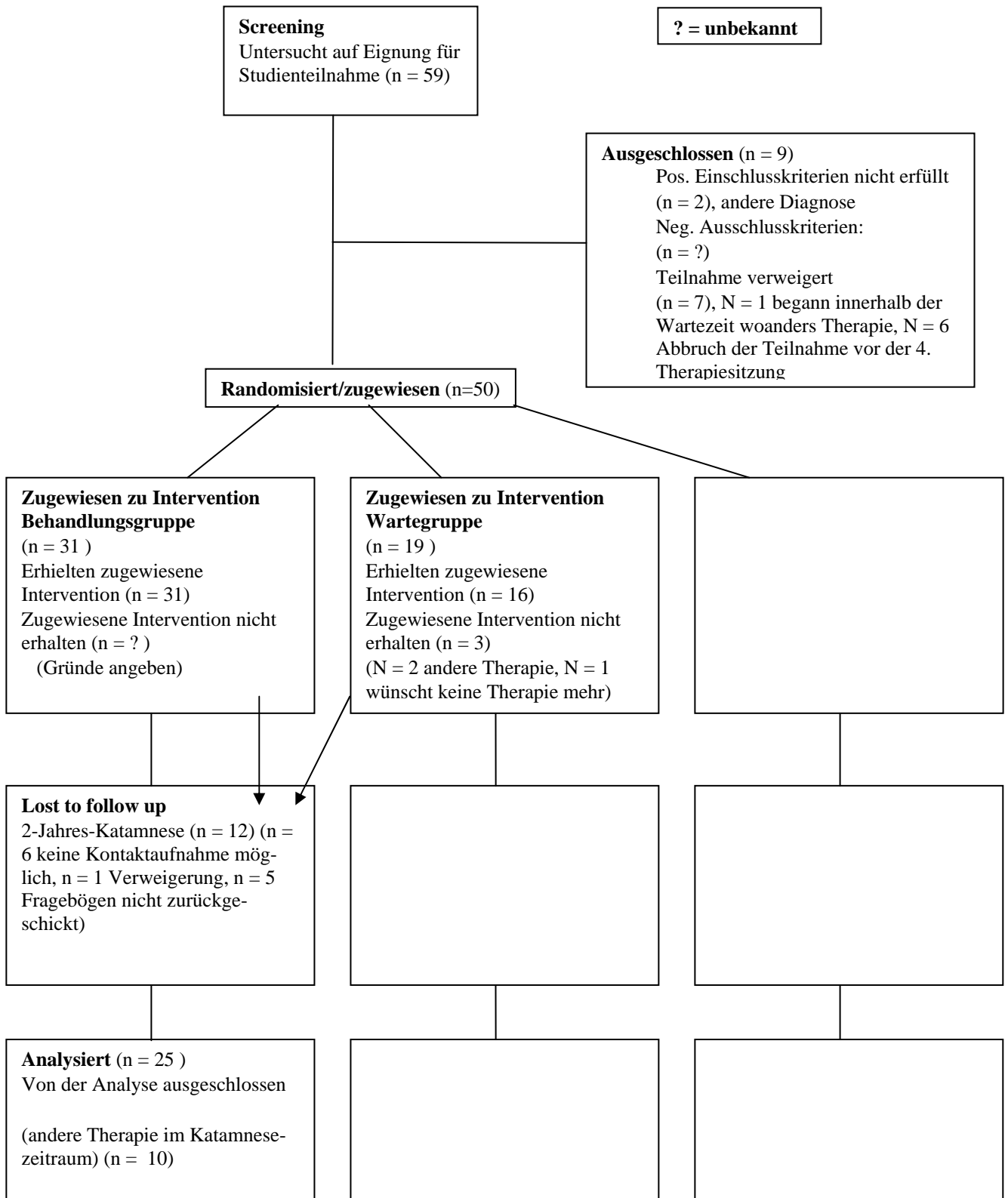
Die Zuordnung zu den Gruppen erfolgt aus ethischen Gründen nicht-randomisiert, sondern nach Reihenfolge des Einschusses und der Zumutbarkeit einer Wartezeit, daher ist die Wartegruppe zum Zeitpunkt prä im Mittel weniger stark betroffen als die Therapiegruppe.

Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie wird als geeignet für den Nachweis der Wirksamkeit der GT bei Anpassungsstörungen angesehen.

Die Behandlungsgruppe zeigt statistisch und klinisch bedeutsame Verbesserungen der Störung im Vergleich zu einer Wartegruppe. Bei ansonsten hohem methodischen Niveau, ist die nicht erfolgte Randomisierung der Zuordnung problematisch.

8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

K5 (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

K14 (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

K33 (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

K39 (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

Anmerkungen zur durchgeführten Studienbewertungen

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

MS

Studie:

Autoren:

Barrowclough et al.

Titel:

A randomized trial of the effectiveness of cognitive-behavioral therapy and supportive counseling for anxiety symptoms in older adults

Quelle:

Journal of Consulting and Clinical Psychology 2001; 69: 756-762

1. Basisinformationen

1.1 () **Benennung der Vorgehensweise**
(Verfahren/Methode/Interventionstechnik)

Supportive Beratung

1.2 (6) **Fragestellung / Zielsetzung der Studie / Forschungshypothese**

Wirksamkeit von kognitiv-behavioraler Psychotherapie bei älteren Menschen mit Angststörungen

1.3 (2) **Design**

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) **Evidenzstufe**

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Angststörungen nach DSM-IV (überprüft mit SCID)

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnosestellung (mittels (teil-)standardisierter Interviews)	1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnosestellung	A2 1
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	C1 1

2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschlusskriterien:

- Alter mindestens 55 Jahre
- Angststörung nach strukturiertem klinischen Interview für DSM-IV (Panikstörung mit oder ohne Agoraphobie, soziale Phobie, generalisierte Angststörung, Angststörung NOS).
- Schweregrad der Angststörung im Beck Anxiety Inventory außerhalb des nicht-klinischen Normalbereichs

Ausschlusskriterien:

- medizinische Krankheitsfaktoren, die die Angstsymptomatik erklären könnten
- kognitive Leistungseinbußen (Mini Mental State Examination <24 Punkte)
- Psychopharmaka waren nur erlaubt, wenn sie konstant für 3 Monate vor Behandlungsbeginn gegeben wurden und die Dosis während der Behandlung auch gleich bleibend geplant war
- Anderweitige psychotherapeutische Behandlung in den letzten 6 Monaten vor Behandlungsbeginn

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	B1 1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	B2 1

			2
--	--	--	---

2.3 () Setting/Patientenzugang

Gängige klinische Zugangswege sowie zu geringem Teil auch über Anzeigen in lokalen Zeitungen.

	Kriterium	Operationalisierung	Rating
K6.	Art der Rekrutierung der Stichprobe	1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege 2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege 3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung)	C2 2
(C.2.)			

2.4 (7) Zentrale Patientenmerkmale
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

Alter: 72 Jahre (SD 6.2);
23% Männer;
49% verheiratet, 39% verwitwet,
95% haben die Schule zwischen dem 14.-16. Lj. verlassen.

Alle nahmen Psychopharmaka (detailliert angegeben).
1 Patient lebte in einer Pflegeeinrichtung, der Rest lebte in eigener Wohnung.
81% hatten körperliche Beschwerden/Krankheiten.

Diagnosen:
51% Panikstörung,
2% soziale Phobie (n=1),
19% generalisierte Angststörung,
28% Angststörung NOS.

Die Angststörung bestand im Mittel seit 20 Jahren (SD 20.0), in der CBT-Gruppe 25 Jahre (SD 22), in der SC-Gruppe 16 Jahre (SD 18).
10 Patienten (23%) hatten komorbide psychiatrische Diagnosen: 7 eine affektive Störung als Folge der Angststörung und 3 eine weitere Angststörung.

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Keine Angabe

	Kriterium	Operationalisierung	Rating
K7.	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	A16 3
(A.16.)			

2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

2stufiges Screening:
225 Patienten wurden initial gescreent,
46 davon wurden ausgeschlossen,
179 waren potentiell geeignet.
55 erfüllten alle Kriterien im 2. ausführlichen Screening und wurden randomisiert zugewiesen. Die meisten wurden ausgeschlossen, da sie nicht die DSM-IV Kriterien erfüllten, keine weiteren Angaben.

Stichprobengrößen:

CBT: 27
SC: 28

9 Drop-outs während der Baseline-Phase, 7 bei CBT, 2 bei SC. Die Gründe werden detailliert angegeben (5 bei CBT und 1 bei SC ziehen Einverständnis zurück; 2 bei CBT und 1 bei SC bekamen eine schwere körperliche Erkrankung).

3 weitere Drop-outs während der ersten 4 Sitzungen (1 bei CBT, 2 bei SC), es liegen keine Post-Daten vor.

Gesamtanzahl der Drop-outs: 29.6% bei CBT, 14.3% bei SC

43 Patienten beenden die Behandlung (CBT: 19; SC: 24).
Von 91% gibt es 3-Monats-Katamnesedaten (CBT: 16; SC: 23);
ebenso zur 6-Monats-Katamnese,
zur 12-Monats-Katamnese gibt es von 93% Daten (von 1 CBT-Patienten mehr).

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 2
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 2
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A17 2
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A3 2
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 2

2.7 (10) Vergleichbarkeit der Gruppen

Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen
Keine signifikanten Gruppenunterschiede (sorgfältig geprüft).

Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)

Keine Unterschiede in der Selbsteinschätzung der Patienten hinsichtlich der Glaubwürdigkeit der Therapien. Beide Gruppen erhielten vergleichbar viele Sitzungen.

Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)

Alle Patienten nahmen Psychopharmaka, Unterschiede zwischen den Gruppen wurden nicht dargestellt, daher nicht beurteilbar.

Vergleichbarkeit der Therapieadhärenz

Die CBT-Gruppe hat etwa doppelt so viele Drop-outs wie die SC-Gruppe.

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen 2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert 3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanten Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt	B9 1
K14. (A.18.)	Drop-out-Analysen	1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen 2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt 3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage	A18 2

3. Intervention

**3.1 (14) Studiendesign /
Behandlungsarme**

Verglichen wurden 2 Behandlungsarme: kognitiv-behaviorale Psychotherapie vs. supportive Beratung

Beide Behandlungen gingen über 16 Wochen (8-12 wöchentliche Sitzungen) und wurden bei den Patienten zu Hause durchgeführt (bis auf 4 Ausnahmen, wo die Therapie in der Klinik stattfand). Die Behandlungen wurden jeweils von qualifizierten Vertretern des jeweiligen Verfahrens durchgeführt.

**3.2 (11) Behandlung Experimental-
gruppe
(Art, Dauer, Frequenz)**

Supportive Beratung (SC). Manualbasiert, das Vorgehen ist detailliert beschrieben. Empathie, Wertschätzung und Authentizität als Grundvariablen, kein expliziter Bezug zur GPT.

**3.3 (12) Behandlung Kontroll-
gruppe
(Art, Dauer, Frequenz)**

Störungsspezifische Kognitiv-behaviorale Psychotherapie (CBT), keine spezifische Änderung des therapeutischen Vorgehens für die ältere Patientengruppe. Manualisiert, das Vorgehen wird detailliert beschrieben.

**3.4 (13) Evtl. weitere Kontroll-
gruppen
(Art, Dauer, Frequenz)**

Keine

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 1
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	B4 1
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	B5 1
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	B6 1
K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	C4 2

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantiellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 3
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	C6 1
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	C7a 1
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	C7b 9

	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	C7c 1
--	---	---	----------

3.5 (19) Beobachtungszeitraum/ Messzeitpunkte/Katamnese

Nach der Baseline-Erhebung und der randomisierten Zuweisung gab es eine 6-wöchige Baseline-Phase ohne psychotherapeutische Behandlung, dann erfolgte die Prä-Untersuchung. Postmessung direkt nach Therapieende (nach 16 Wochen), Katamnesen nach 3, 6 und 12 Monaten.

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 2
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	B10 1
K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	B11 1

3.6 (15) Anzahl der Zentren

1

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...)

Randomisierung nach

- Schweregrad der Symptomatik (kategorisiert am Cut-off von 19 Punkten im Beck Anxiety Inventory,
- Vorliegen komorbider Erkrankungen (dichotomisiert ja/nein)
- Art der Angststörung

3.8 (17) ggf. Concealment der Randomisierung

Die Randomisierung erfolgte über ein Computerprogramm und wurde von unabhängigen Dritten durchgeführt.

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 3
K27.	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung	B8

(B.8.)		hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	2
--------	--	--	---

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

Angstsymptomatik:

- Selbsteinschätzung: Beck Anxiety Inventory und State-Trait Anxiety Inventory Trait form
- Fremdeinschätzung: Hamilton Rating Scale for Anxiety

4.2 (21) Sekundäre Zielkriterien

Depressive Symptomatik (beides Selbsteinschätzungen)

- Beck Depression Inventory
- Geriatric Depression Scale

4.3 () Verblindung der Erhebung der Zielkriterien

Ja

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,	A7 1
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend	A8 1
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)	C9 1
K31. (A.9.)	Multiple Informationsquellen	1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle	A9 2
K32. (A.10.)	Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)	1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler 2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern 3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten	A10 1

4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)

Die Anwendungsvoraussetzungen für parametrische Verfahren wurden geprüft. Für Baseline-Gruppenunterschiede wurden T-Tests verwendet. T-Tests für abhängige Stichproben wurden verwendet, um Veränderungen pro Gruppe über die Zeit zu untersuchen. Gruppenunterschiede nach Behandlungsende wurden mittels Kovarianzanalysen geprüft (mit den Baseline-Werten als Kovariaten). Für die Katamnesedaten wurden Messwiederholungsvarianzanalysen gerechnet (mit Post-Werten und den 3 Katamnesezeitpunkten) mit den Baselinewerten als Kovariaten.

Zusätzlich 3 Klassifizierungen als Treatment Responder:

- Anxiety Responder: 20% Reduktion der Angstsymptomatik (im Beck Anxiety Inventory und in der Hamilton Anxiety Rating Scale von Baseline bis zur 12-Monats-Katamnese)
- Depression Responder: 20% Reduktion der depressiven Symptomatik (im Beck Depression Inventory und in der Geriatric Depression Scale von Baseline bis zur 12-Monats-Katamnese)
- Responder nach Endstate-Functioning: Outcome-Werte zur 12-Monatskatamnese im nicht-klinischen Normalbereich (< 10 im Beck Anxiety Inventory, <= 10 in der Hamilton Anxiety Rating Scale, <= 15 in der Geriatric Depression Scale, <10 im Beck Depression Inventory).

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	A6 1
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substantiellem Risiko für deren Verletzung	A13 1
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	A14 2

4.5 (22) Ergebnisse

Baseline-Phase: keine signifikanten Veränderungen.

Prä-Post-Veränderungen:
Bis auf 1 Ausnahme bessern sich alle Outcome-Maße in beiden Gruppen signifikant (Ausnahme: die Geriatric Depression Scale in der SC-Gruppe).

Bei Behandlungsende:
CBT ist SC signifikant in der Besserung von 2
Selbsteinschätzungsverfahren überlegen: im Beck Anxiety Inventory und in der Geriatric Depression Scale.

Katamnese:
Die erreichten Veränderungen bleiben stabil. In der Messwiederholungsvarianzanalyse zeigt die CBT-Gruppe im Vergleich zur SC-Gruppe in 4 der 5 Outcome-Maße signifikant höhere

Verbesserungen (nur im Beck Depression Inventory nicht).

Treatment Responder:
 Anxiety Responder:
 Signifikant mehr in der CBT-Gruppe mit 71% vs 39% der SC-Gruppe;
 Depression Responder:
 59% der CBT-Gruppe vs. 39% der SC-Gruppe (ns);
 Endstate-Functioning Anxiety:
 41% der CBT-Gruppe vs. 26% der SC-Gruppe (ns);
 Endstate-Functioning Depression:
 59% der CBT-Gruppe vs. 39% der SC-Gruppe (ns).

Drop-out-Analysen:
 Drop-outs waren bei Baseline signifikant depressiver und jünger.

	Kriterium	Operationalisierung	Rating
K36. (A.11.)	Vollständige Darstellung der Ergebnisse zu allen Outcomemaßen und zu allen relevanten Messzeitpunkten	1) für alle Outcome-Variablen berichtet 2) ausschließlich für die primären Zielkriterien berichtet 3) nicht für alle primären Zielkriterien berichtet	A11 1
K37. (B.12.)	Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte)	1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe 2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig 3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG	B12 1

4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen

Keine Angabe

	Kriterium	Operationalisierung	Rating
K38. (A.12.)	Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen	1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet	A12 3

4.7 (24) Fazit/Schlussfolgerungen der Autoren

Überzeugende Wirksamkeit von CBT bei Angststörungen älterer Menschen, deutliche Überlegenheit von CBT gegenüber der supportiven Beratung.

	Kriterium	Operationalisierung	Rating
K39. (A.1.)	Manipulation der Daten	1) keine Hinweise auf Ergebnismanipulation 3) Hinweise auf Ergebnismanipulation	A1 1

5. Beurteilung der Ergebnisse der Studie		
D1	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual". (Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt) (ja = 1; nein = 3; nicht zutreffend = freilassen)	D1 3
D2	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D2 3
D3	In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D3 3

6. Praxistransfer

	Kriterium	Operationalisierung	Rating
K40. (C.10.)	Spezifikation und Herstellbarkeit notwendiger Settingbedingungen	1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team) 2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team) 3) Notwendige Settingbedingungen nicht herstellbar	C10 2
K41. (C.11.)	Spezifikation und Herstellbarkeit der notwendigen Behandlerqualifikation	1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar 2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar 3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar	C11 1
K42. (C.12.)	Spezifikation und Erfassbarkeit relevanter Patientenmerkmale	1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar 2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar 3) Relevante Patientenmerkmale praktisch nicht erfassbar	C12 1
K43. (C.13.)	Spezifikation und Herstellbarkeit relevanter Treatmentaspekte	1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual) 2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung) 3) Relevante Treatmentmerkmale praktisch nicht herstellbar	C13 1

7. Gesamtbewertung der Studie

Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)			
Durchschnittswert für B) Interne Validität (Summe / 12)			
Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)			
Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4)			

Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

<p>Kriterien für 'A. Allgemeine methodische Qualität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> 2. wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde <p>und</p> <ol style="list-style-type: none"> 3. wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist. 	
<p>Kriterien für 'B. Interne Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> 2. der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist. 	
<p>Kriterien für 'C. Externe Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> 2. das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> 3. bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde <p>und</p> <ol style="list-style-type: none"> 4. bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde. 	
<p>Die Ergebnisse der Studie belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>	

Abschließende Bewertung

Allgemeine methodische Qualität der Studie

Die Methodik ist allgemein hinreichend gut (Diagnosestellung, Wahl der Messzeitpunkte, Wahl der Messinstrumente und statistische Auswertung sind befriedigend), allerdings wird die CBT eindeutig bevorzugt.

Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Störungsspezifische CBT ist supportiver Beratung bei der Reduktion von Angst- und depressiver Symptomatik bei älteren Angstpatienten zu Behandlungsende in 2 von 5 Outcome-Maßen überlegen sowie zur Katamnese in 4 von 5 Outcome-Maßen.

Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Deutlich eingeschränkt. Alle Patienten nahmen vor Behandlungsbeginn auch Psychopharmaka, keine weiteren Angaben zur Medikation während der Behandlung, keine differenzierte Darstellung nach Gruppen. Die Supportive Beratung wurde von nur 1 Therapeut durchgeführt, die CBT nur von 2 Therapeuten. Die Behandlungen waren vergleichbar in Dauer und Intensität, die Gruppen waren auch zu Baseline vergleichbar, unterscheiden sich aber deutlich in den Drop-out Raten (bei CBT doppelt so hoch wie bei SC). Die CBT war störungsspezifisch, die supportive Beratung nicht.

Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Deutlich eingeschränkt. Die Stichprobe ist aufgrund des Alters (≥ 55 Jahre) und der überwiegend geringen Schulbildung (95% verließen die Schule im 14.-16.Lj.) sehr selektiv. Die Therapien wurden bei den Patienten zu Hause durchgeführt, was in Deutschland nicht üblich ist.

Stärken der Studie:

Der Schwerpunkt auf Psychotherapie älterer Menschen ist interessant und generell ein wenig beforschtes Gebiet.

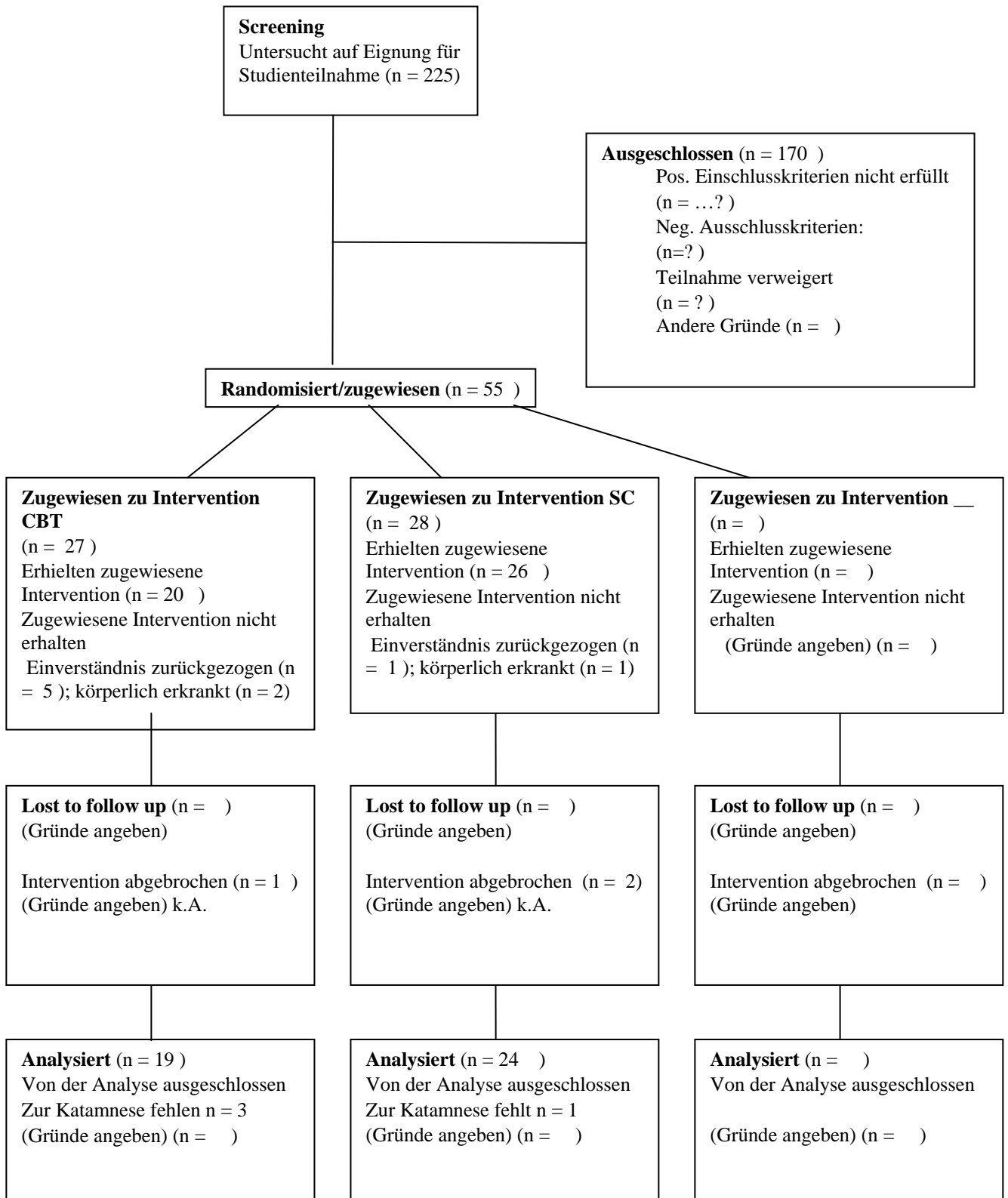
Schwächen der Studie:

Einige methodische Mängel schränken die Aussagekraft der Studie ein. Dies gilt insbesondere für die parallele Pharmakotherapie und die unterschiedliche hohen drop-outs zwischen den beiden Treatmentbedingungen

Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Eingeschränkt relevant, da die Intervention lediglich als GPT II gewertet werden kann. Darüber hinaus erlaubt das Design und die methodischen Mängel keinen Wirksamkeitsnachweis.

8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

K5 (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

K14 (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

K33 (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

K39 (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

Anmerkungen zur durchgeführten Studienbewertungen

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

HK

Studie:

Autoren:

Cooper, P. J., Murray, L., Wilson, A. & Romaniuk, H.

Titel:

Controlled trial of the short- and long term effect of psychological treatment of post-partum depression

Quelle:

British Journal Psychiatry, 2003, 182, 412-419

1. Basisinformationen

1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Non-direktive Beratung nach Holden et al, 1989

1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Vergleich der Effektivität verschiedener psychotherapeutischer Interventionen (non-direktiv, kognitiv-behavioral, psychodynamisch) gegenüber der üblichen medizinische Behandlung der Primärversorgung (Kontrollgruppe) bei Post-Partum-Depressionen

1.3 (2) Design

- X Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) Evidenzstufe

- X Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Frauen mit Post-Partum-Depression (Major Depression nach Geburt eines Kindes)

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews)	1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnosestellung	A2 2
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	C1 2

2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschluss: Überschreiten eines cut-poffs bei einem Fragebogen zu Post-Partum-Depressivität; klinische Diagnose einer Major Depression; erstmalige Geburt, Englisch als Muttersprache, wohnhaft in 15-Meilen-Radius um Hospital;
Ausschluss: Frühgeburt, angeborene schwere Krankheiten des Kindes, Zwillingsgeburt, Vorhaben, aus der Stadt weg zu ziehen

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	B1 1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	B2 1
K5. (C.3.)	Selektivität der Stichprobe aufgrund der Ausschlusskriterien	1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten 2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen) 3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien	C3 1

2.3 () Setting/Patientenzugang

Anhand von Geburtsberichten eines Hospitals in Cambridge postalischer Kontakt mit Bitte einen Fragebogen auszufüllen zur Erhebung von Post-Partum depressiven Beschwerden: überschreiten eines Cut-Off-Wertes (>11) der Edinburgh Postnatal Depression Scale

	Kriterium	Operationalisierung	Rating
--	-----------	---------------------	--------

			3
--	--	--	----------

2.4 (7) Zentrale Patientenmerkmale
 (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

Alter Range: 17-42
 (relativ ausführlich weitere Merkmale dargestellt, augenscheinlich weitgehende Vergleichbarkeit der Verteilungen zwischen den Gruppen)

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

n=44 pro Gruppe; insgesamt N=176

	Kriterium	Operationalisierung	Rating
K7. (A.16.)	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	A16 1

2.6 (9) Anzahl gescreener, eingeschlossener, zugewiesener und ausgewerteter Patienten
 (Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

Identifiziert: 206; randomisiert: 193;
 4.5-Monatskatamnese: 170
 9-Monatskatamnese: 180
 18-Monatskatamnese: 178
 60-Monatskatamnese: 138

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 1
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 2
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A17 2
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A3 1
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 1

2.7 (10) Vergleichbarkeit der Gruppen

<p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen</p> <ul style="list-style-type: none"> - gut; Einschränkung lediglich bzgl. Psychodynamischer Treatmentgruppe: geringere Anzahl von Personen mit „high social disadvantage“ <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)</p> <ul style="list-style-type: none"> - gegeben für die Treatmentgruppen <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)</p> <ul style="list-style-type: none"> - keine Angaben <p>Vergleichbarkeit der Therapieadhärenz</p> <ul style="list-style-type: none"> - ähnlich hohe drop-outs

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	<p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanten Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p>	<p>B9</p> <p>2</p>
K14. (A.18.)	Drop-out-Analysen	<p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>	<p>A18</p> <p>2</p>

3. Intervention

3.1 (14) Studiendesign / Behandlungsarme

- Nondirektives „Counselling“ (wie in Edingburgh Study von Holden et al, 1989)
- Kognitiv-behaviorale Therapie (interaction guidance treatment (Mc-Donough, 1993)
- Psychodynamische Therapie (speziell für Mütter, nach Cramer, 1990 & Stern, 1995)
- übliche Primärversorgung

3.2 (11) Behandlung Experimental- gruppe (Art, Dauer, Frequenz)

- Nondirektives „Counselling“ (wie in Edingburgh study von Holden et al, 1989;
- 1*wöchentlich zwischen 8 und 18 Wochen post partum zu Hause bei der jeweiligen Mutter

3.3 (12) Behandlung Kontroll- gruppe (Art, Dauer, Frequenz)

- routinemäßige Primärversorgung, keine Angaben zu Art, Dauer, Frequenz (wahrscheinlich je nach Bedarf)

3.4 (13) Evtl. weitere Kontroll- gruppen (Art, Dauer, Frequenz)

- Kognitiv-behaviorale sowie psychodynamische Therapie ebenfalls 1*wöchentlich zwischen 8 und 18 Wochen post partum zu Hause bei der jeweiligen Mutter

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 3
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	B4 3
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	B5 3
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	B6 1
K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	C4 3

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 3
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	C6 1
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	C7a 3
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	C7b 3
	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	C7c 3

**3.5 (19) Beobachtungszeitraum/
Messzeitpunkte/Katamnese**

Pre-Treatment sowie 4.5-, 9-, 18-, 60-Monats-Katamnesen

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 2
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	B10 1
K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	B11 1

3.6 (15) Anzahl der Zentren

1

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...)

Randomisierung

3.8 (17) ggf. Concealment der Randomisierung

Concealment nicht gegeben

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 3
K27. (B.8.)	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ($n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	B8 1

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

-Reduktion in Post-Partum-Depressivitätsskala *Edinburgh Postnatal Depression Scale*;
 -Häufigkeit der SKID-Diagnose Depression
 - Auftreten mindestens einer weiteren depressiven Episode nach der Post-Partum-Depression

4.2 (21) Sekundäre Zielkriterien

keine

4.3 () Verblindung der Erhebung der Zielkriterien

Interviewer kannten angeblich die Gruppenzuordnung nicht

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,	A7 2
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend	A8 1
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)	C9 2
K31. (A.9.)	Multiple Informationsquellen	1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle	A9 2
K32. (A.10.)	Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)	1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler 2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern 3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten	A10 1

4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)

Aufwändige Verfahren zur simultanen Analyse mehrerer Messgelegenheiten unter Einbeziehung von konfundierenden Variablen

(Recht spärliche Angaben zu den angewandten Verfahren (Nennung von Monografien; wahrscheinlich spezielle Anwendungen von Multi-Level-Analysen für längsschnittliche Daten.)

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	A6 1
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung	A13 2
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	A14 1

4.5 (22) Ergebnisse

Alle drei Treatmentgruppen zeigen kurzfristig (4.5-Monatskatamnese) signifikant bessere Depressivitätsreduktion (erhoben mit *Edinburgh Postnatal Depression Scale*) als die Kontrollgruppe, nicht aber eine höhere Reduktion der SKID-Depressionsdiagnosen.

Die erhöhte Depressivitätsreduktion in den Treatmentgruppen ist zu späteren Messgelegenheiten (9-, 18-, 60-Monats-Katamnesen nicht mehr feststellbar.

	Kriterium	Operationalisierung	Rating
K36. (A.11.)	Vollständige Darstellung der Ergebnisse zu allen Outcomemaßen und zu allen relevanten Messzeitpunkten	1) für alle Outcome-Variablen berichtet 2) ausschließlich für die primären Zielkriterien berichtet 3) nicht für alle primären Zielkriterien berichtet	A11 1
K37. (B.12.)	Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte)	1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe 2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig 3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG	B12 1

4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen

- keine Angaben

	Kriterium	Operationalisierung	Rating
K38. (A.12.)	Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen	1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen 3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet	A12 3

4.7 (24) Fazit/Schlussfolgerungen der Autoren

Psychologische Interventionen verbessern die depressive Stimmung bei Post-Partum-Depressionen, sind aber langfristig dem Auftreten spontaner Remissionen nicht überlegen.

	Kriterium	Operationalisierung	Rating
K39. (A.1.)	Manipulation der Daten	1) keine Hinweise auf Ergebnismanipulation 3) Hinweise auf Ergebnismanipulation	A1 1

5. Beurteilung der Ergebnisse der Studie		
D1	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual". (Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt) (ja = 1; nein = 3; nicht zutreffend = freilassen)	D1 1
D2	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D2
D3	In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D3 3

6. Praxistransfer

	Kriterium	Operationalisierung	Rating
K40. (C.10.)	Spezifikation und Herstellbarkeit notwendiger Settingbedingungen	1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team) 2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team) 3) Notwendige Settingbedingungen nicht herstellbar	C10 1
K41. (C.11.)	Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation	1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar 2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar 3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar	C11 1
K42. (C.12.)	Spezifikation und Erfassbarkeit relevanter Patientenmerkmale	1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar 2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar 3) Relevante Patientenmerkmale praktisch nicht erfassbar	C12 1
K43. (C.13.)	Spezifikation und Herstellbarkeit relevanter Treatmentaspekte	1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual) 2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung) 3) Relevante Treatmentmerkmale praktisch nicht herstellbar	C13 1

7. Gesamtbewertung der Studie

Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)			
Durchschnittswert für B) Interne Validität (Summe / 12)			
Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)			
Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4)			

Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

<p>Kriterien für 'A. Allgemeine methodische Qualität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde <p>und</p> <ol style="list-style-type: none"> wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist. 	
<p>Kriterien für 'B. Interne Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist. 	
<p>Kriterien für 'C. Externe Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde. 	
<p>Die Ergebnisse der Studie belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>	

Abschließende Bewertung

Allgemeine methodische Qualität der Studie

Gut: hinreichende Fallzahlen, Randomisierung, langfristige Katamnesen, Operationalisierung des Outcomes durch Selbst- und Fremdeinschätzungsverfahren, Kontrolle von Störvariablen, komplexe Datenanalysen

Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Non-direktive Beratung ist einer gewöhnlichen medizinischen Primärversorgung hinsichtlich der Reduktion von Depressivität bei Post-Partum-Depressionen überlegen. Allerdings sind die Effekte längerfristig dem Auftreten von Spontanremissionen nicht überlegen und erreichen kurzfristig auch kein Ausmaß, dass sich in einer signifikanten Reduktion der Diagnosestellung einer Major Depression zeigt.

Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Gut

Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Gut für Post-Partum-Depressionen;
Allgemein eingeschränkt durch Einschränkung auf Post-Partum-Depressionen

Stärken der Studie:

langer Katamnesezeitraum, geringe Drop-Out-Rate, komplexe multivariate statistische Verfahren

Schwächen der Studie:

Unpräzise Beschreibung der Treatment- und Kontrollbedingungen
Keine Berücksichtigung begleitender, nicht randomisierter Interventionen (z.B. Pharmakotherapie)

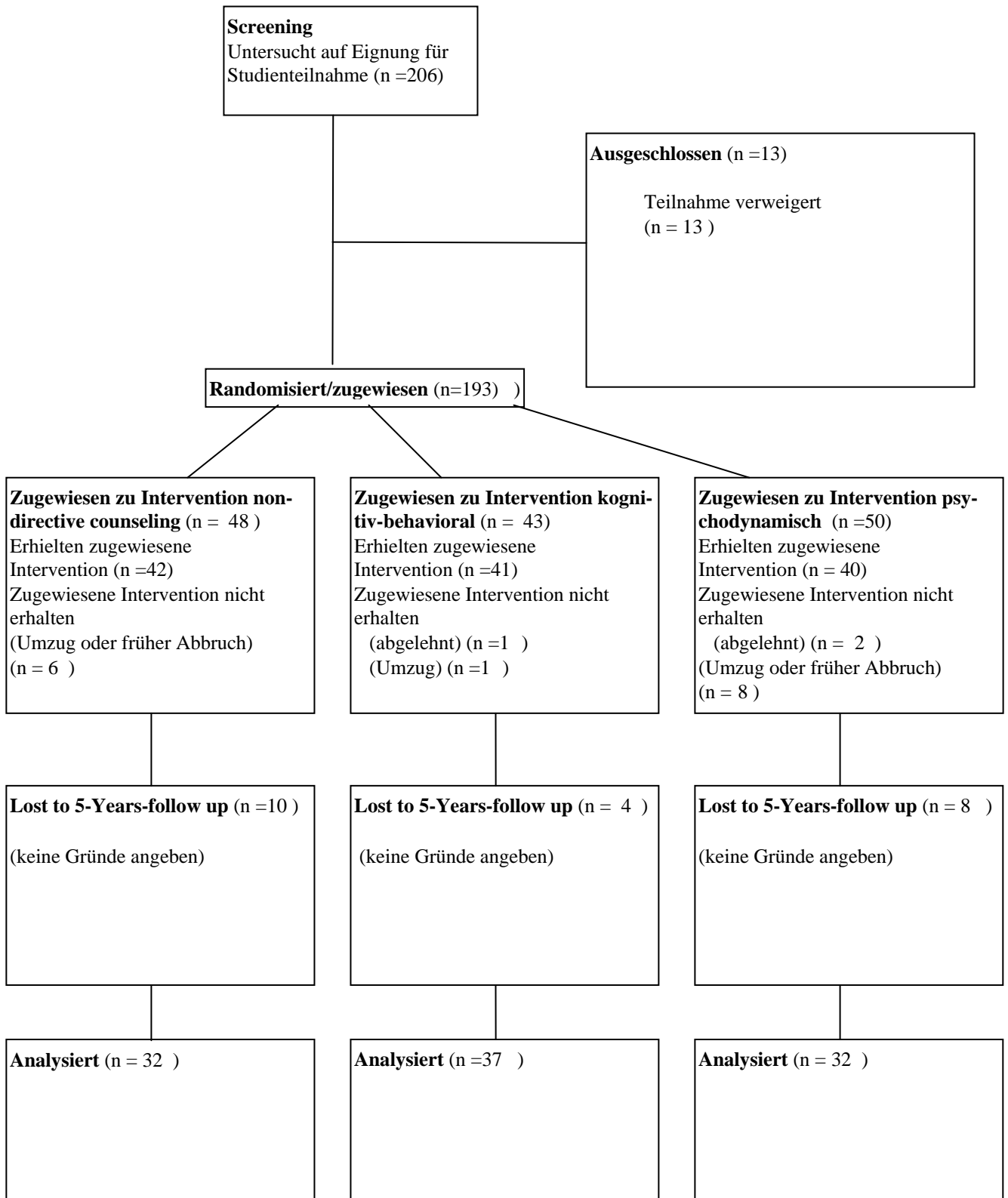
Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Relevant:

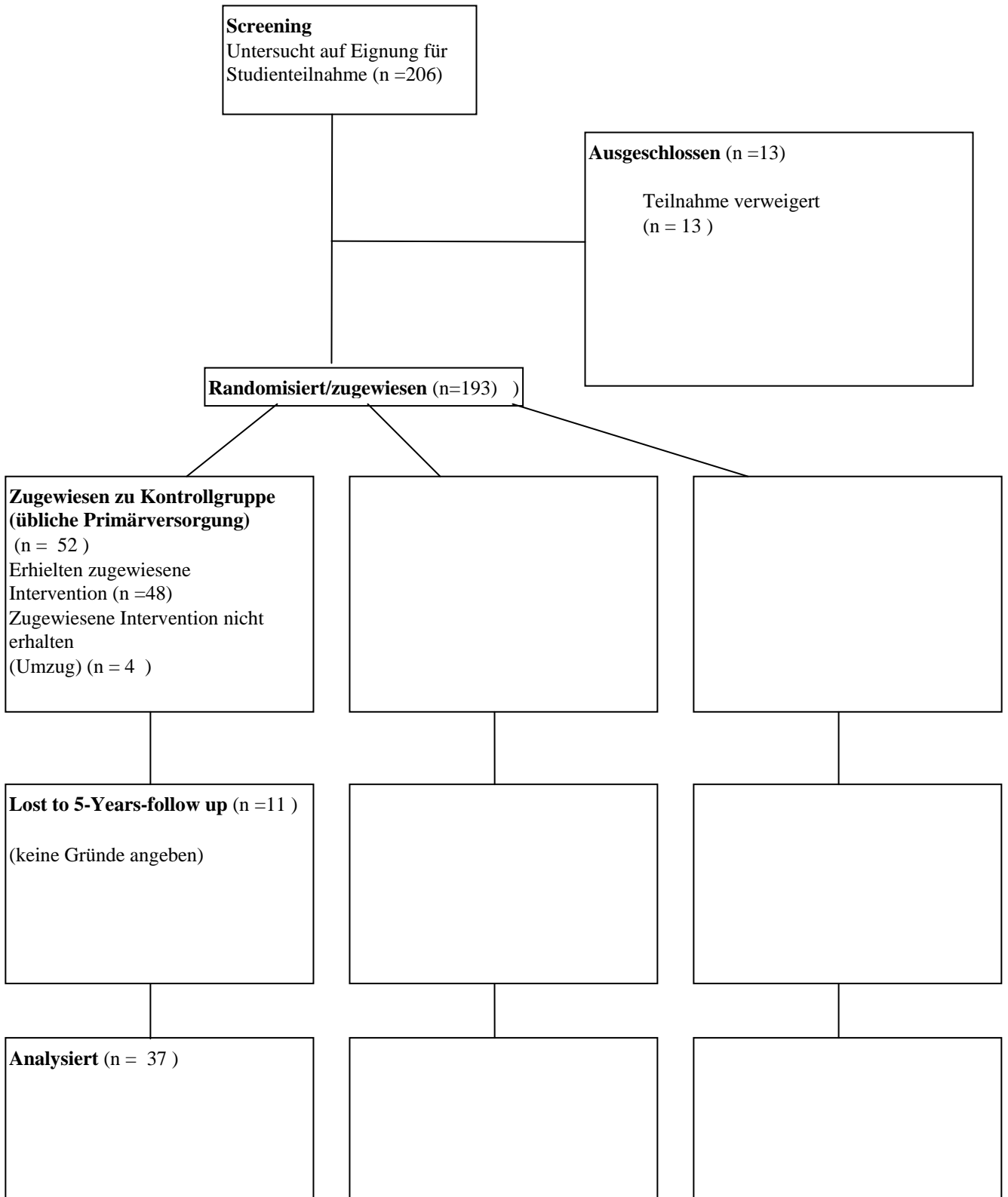
Non-direktive Beratung ist als Gesprächspsychotherapie zu werten (GPT I), unklar ist dabei welche Bedeutung der Durchführung der GT durch Laien - Nicht-Psychologen und Nicht-Ärzte – zukommt, ob die Effektivität der GT dadurch, gerade auch im Vergleich zu den anderen angewandten Psychotherapiemethoden, geringer ausfiel.

- kurzfristig signifikant stärkere Reduktion der depressiven Symptomatik im Vergleich zur Primärversorgung und vergleichbare Effektivität wie etablierte psychotherapeutische Verfahren (psychodynamische und kognitiv-behaviorale Interventionen).
- keine langfristige Überlegenheit bei der Reduktion der depressiven Symptomatik sowie keine kurzfristige oder langfristige Reduktion bezüglich klinischer Depressionsdiagnosen gegenüber üblicher Primärversorgung

8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

K5 (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

K14 (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gec-

reenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

K33 (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

K39 (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

Anmerkungen zur durchgeführten Studienbewertungen

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

MT

Studie:

Autoren:

Dircks et al.

Titel:

Förderung der seelischen Gesundheit von Krebspatienten durch personenzentrierte
Gruppengespräche

Quelle:

Z. f. Klin. Psych., 1982, 11(4): 241-252

1. Basisinformationen

1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Personenzentrierte Gesprächsgruppen
- die Therapeuten haben das Ziel, in ihrem Verhalten Echtheit, Akzeptanz und Empathie (im Sinne von Rogers) auszudrücken

1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Wirksamkeit von personenzentrierter Gruppengesprächspsychotherapie bei Krebspatienten im Vergleich zu einer Gruppe ohne Teilnahme an Gesprächsgruppen

Hypothese: Personenzentrierte Gesprächsgruppen führen bei Krebspatienten zu positiven Veränderungen im Bereich der Persönlichkeit, im Verhältnis zu ihrer sozialen Umwelt sowie ihrer Einstellung zu Krankheit und Tod.

1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Krebserkrankung unterschiedlicher Art (Brust, Gebärmutter-, Nierenkrebs), im Durchschnitt 3 Jahre

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews)	1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnosestellung	3 A2
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	2 C1

2.2 (7) Relevante Ein- und Ausschlusskriterien

- Einschluss: Krebserkrankung unterschiedlicher Art (Brust, Gebärmutter-, Nierenkrebs)

- Ausschluss: keine

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	1 B1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	1 B2
K5. (C.3.)	Selektivität der Stichprobe aufgrund der Ausschlusskriterien	1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten 2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen) 3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien	1 C3

2.3 () Setting/Patientenzugang

Gesprächssetting Hamburg Allgemeines Krankenhaus Altona, Patientenzugang nicht aus dem Artikel ersichtlich (für Details Referenz auf Dissertationsschrift von Kalliner, 1981)

	Kriterium	Operationalisierung	Rating
K6. (C.2.)	Art der Rekrutierung der Stichprobe	1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege 2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege 3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung)	9 C

2.4 (7) Zentrale Patientenmerkmale
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- mittleres Alter 51 Jahre (24-78 Jahre)
- 44 Frauen, 14 Männer
- Erkrankungsdauer durchschnittlich ca. 3,2 Jahre (2 Monate – 13,5 Jahre)

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Keine Angaben

	Kriterium	Operationalisierung	Rating
K7. (A.16.)	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	9 A16

2.6 (9) Anzahl gescreener, eingeschlossener, zugewiesener und ausgewerteter Patienten
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- Anzahl gescreener Patienten nicht bekannt
- 58 Patienten nehmen an der Studie teil (30 Gesprächsgruppenspsychotherapie, 28 Kontrollgruppe ohne Gruppentherapie)
- aus dem Artikel wird nicht ersichtlich, wie die Zuteilung zu den Gruppen erfolgte und, ob sie randomisiert erfolgte
- Vor- und Nachuntersuchung, Katamnese nach 3 Monaten
- da nur prozentuale Werte angegeben werden, ist unklar, ob Drop-outs zur Katamnese

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 2
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 9
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A7 3
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	9 A3
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 9

2.7 (10) Vergleichbarkeit der Gruppen

<p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen:</p> <ul style="list-style-type: none"> - GT-Gruppe ist im Durchschnitt jünger (GT-48 Jahre, KG-54 Jahre), häufiger weiblich (GT-87%, KG-64%) - beide Gruppen geben vor der Behandlung gleich viele (durchschnittlich 7) persönliche, sie stark belastende Beeinträchtigungen an, die sich aber unterschiedlich verteilen: GT gibt deutlich häufiger als KG „Schwierigkeiten in zwischenmenschlichen Beziehungen“ und „emotionale Schwierigkeiten“ an, während KG etwas mehr „körperliche Folgebeschwerden“, „reduzierte körperlich-seelische Leistungsfähigkeit“, „Angst (außer Krebsangst)“, „Selbstwertprobleme“ und „Schwierigkeiten mit äußeren Lebensbedingungen“ angibt <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer):</p> <ul style="list-style-type: none"> - entfällt, da Kontrollgruppe unbehandelt <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen):</p> <ul style="list-style-type: none"> - kann nicht beurteilt werden <p>Vergleichbarkeit der Therapieadhärenz:</p> <ul style="list-style-type: none"> - kann nicht beurteilt werden, da keine Dropout-Zahlen angegeben und Ergebnisse ausschließlich in %
--

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	<p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p>	B9 3
K14. (A.18.)	Drop-out-Analysen	<p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>	A18 9

3. Intervention

**3.1 (14) Studiendesign /
Behandlungsarme**

Therapiestudie mit unbehandelter Vergleichsgruppe (unklar, ob randomisiert), 2 Behandlungsarme: 1. Personenzentrierte Gesprächsgruppen
2. keine Teilnahme an personenzentrierten Gesprächsgruppen

**3.2 (11) Behandlung Experimental-
gruppe
(Art, Dauer, Frequenz)**

Personenzentrierte Gesprächsgruppen
- die Therapeuten haben das Ziel, in ihrem Verhalten Echtheit, Akzeptanz und Empathie (im Sinne von Rogers) auszudrücken

- 11 Sitzungen, 1x/Woche, mehrstündig/ drei Sitzungen ganztägig, ambulant

**3.3 (12) Behandlung Kontroll-
gruppe
(Art, Dauer, Frequenz)**

Kontrollgruppe ohne personenzentrierte Gesprächsgruppen

**3.4 (13) Evtl. weitere Kontroll-
gruppen
(Art, Dauer, Frequenz)**

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 2
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	B4
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	B5 9
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	B6
K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	C4

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantziellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 9
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	C6
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	C7a 1
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	C7b 9
	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	C7c 9

**3.5 (19) Beobachtungszeitraum/
Messzeitpunkte/Katamnese**

Behandlungszeitraum 11 Wochen, Untersuchung am Anfang und am Ende des Behandlungszeitraumes, 2-3 Monate nach Behandlungsende
halbstrukturiertes Interview, nach 3 Monaten Katamnese
- resultierender Beobachtungszeitraum ca. 6 Monate

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 1
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	B10
K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	B11 2

3.6 (15) Anzahl der Zentren

1

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...)

Nicht bekannt

3.8 (17) ggf. Concealment der Randomisierung

entfällt

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 1
K27. (B.8.)	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	B8 3

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

Positive Veränderung a) in verschiedenen Bereichen der Persönlichkeit, b) im Verhältnis zu der sozialen Umwelt, c) in der Einstellung zu Krankheit und Tod

Outcome-Maße:

- Liste individueller Beeinträchtigungen siehe Wittern & Tausch, 1980 (Problemliste) (Liste krankheitspezifischer und anderer persönlicher Probleme + Belastungswert)
- Freiburger Persönlichkeitsinventar (FPI)
- Selbstkonzeptskalen (SKS) (s. Wittern, 1979)
- Tätigkeitsliste (zur Erfassung angenehmer Tätigkeiten und Subskala „Soziale Aktivitäten“) (s. Lewinson & Libet, 1972)

- zur 3-Monats-Katamnese raten 3 Experten auf 5-stufiger Skala mit hinreichender Reliabilität „Veränderungen insg.“ (von deutlich verbessert bis deutlich verschlechtert)

- in dieses Urteil fließen die Werte in den zuvor genannten Maßen ein, zusätzlich die „Einschätzung der Gesamtveränderungen und der Veränderung der Einstellung zur Krankheit“ gemäß Nachbefragungsinterview mit dem Patienten, die Zufriedenheit des Patienten mit der Gesprächsgruppe (gemäß Strupp-Fragebogen, Wittern, 1979)

4.2 (21) Sekundäre Zielkriterien

keine

4.3 () Verblindung der Erhebung der Zielkriterien

Nicht bekannt

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,	A7 2
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend	2A8
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)	1C9
K31. (A.9.)	Multiple Informationsquellen	1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle	1A9
K32. (A.10.)	Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)	1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler 2) validiertes Fremdeinschätzungsverfahren angewendet von	A10 9

	rigkeit)	2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern 3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten	
--	----------	---	--

4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)

- nicht ausführlich beschrieben („Die statistische Verrechnung erfolgte je nach Datenqualität mit in der Veränderungsmessung gängigen Verfahren. Für alle Ergebnisse wurde bei einseitiger Testung ein Signifikanzniveau von $\alpha = 5\%$ festgelegt“)

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	1A6
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung	A13 9
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	A14 9

4.5 (22) Ergebnisse

Gesprächsgruppenpatienten verändern sich in ihrer Person, ihrem Verhalten und im Umgang mit anderen Menschen.

- 27 % der Gesprächspatienten längerfristig in verschiedenen Bereichen ihrer Persönlichkeit „deutlich verbessert“, 37 % „leicht verbessert“, 30% „längerfristig unverändert“, 6 % „leicht verschlechtert“ (Gesamteinschätzung der individ. Patienten von 3 Experten)
- mehr soziale Aktivitäten und Tätigkeiten, die Patienten für sich selbst gut finden (Tätigkeitsliste)
- Verbesserung der familiären Beziehungen, Verminderung von Problemen in zwischenmenschlichen Beziehungen
- Verringerung der Nervosität (FPI)
- höheres Maß an Selbstzufriedenheit (SKS) direkt nach den Gesprächen
- weniger Krebsangst, höhere körperliche und seelische Belastbarkeit
- günstigere Einstellung zu ihrer Krankheit und zum Leben
- 97 % der Gesprächspatienten halten personenzentrierte Gesprächsgruppen im Zusammenhang mit der Krankheit für wichtig und hilfreich

- Anmerkung: es kann nicht beurteilt werden, ob bzw. zu welchem Grad die o.g. Ergebnisse statistisch signifikant sind, da sie in % ohne Angabe der Varianz berichtet werden

	Kriterium	Operationalisierung	Rating
K36. (A.11.)	Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten	1) für alle Outcome-Variablen berichtet 2) ausschließlich für die primären Zielkriterien berichtet 3) nicht für alle primären Zielkriterien berichtet	A11 1
K37. (B.12.)	Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte)	1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe 2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig 3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG	B12 3

**4.6. (23) Ergebnisse zu unerwünschten Therapie-
wirkungen**

- 6% der Patienten erleben dem Gesamturteil der Experten nach eine leichte Verschlechterung

	Kriterium	Operationalisierung	Rating
K38. (A.12.)	Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen	1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen 3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet	A12 1

4.7 (24) Fazit/Schlussfolgerungen der Autoren

Die Befunde der vorliegenden Untersuchung legen nahe, personenzentrierte Gesprächsgruppen als psychologisches Hilfeangebot zur Förderung der seelischen Gesundheit von Krebspatienten in der Nachsorgezeit in Betracht zu ziehen:

Die Krebspatienten unserer Gesprächsgruppen zeigen deutlich mehr seelisches Wohlbefinden im Umgang mit der eigenen Person, im Umgang mit anderen sowie im Umgang mit ihrer Krebserkrankung. „Insgesamt“ haben sich trotz ihrer Erkrankung 27 % längerfristig „deutlich verbessert“, 37 % „leicht verbessert“. Insbesondere konnten wir folgende längerfristig beständigen Auswirkungen der Teilnahme an den Gesprächsgruppen und zum Teil an den Selbsthilfegruppen feststellen: Die Patienten fühlen sich entlastet in ihren Problemen mit zwischenmenschlichen Beziehungen. Ein Teil erlebt das Verhältnis zur Familie verbessert. Sie sind fähiger, gegenseitige Offenheit und Einfühlung zu verwirklichen. Angenehme Aktivitäten, besonders auch soziale, nehmen zu. Krebsangst und Nervosität sind vermindert. Die Einstellung zur eigenen Krankheit und zum Leben ändert sich günstig. Betroffenen gewinnen mehr Selbstvertrauen, sie sind offener für ihre eigenen körperlichen und seelischen Vorgänge, ihr Lebenswille ist gestärkt.

	Kriterium	Operationalisierung	Rating
K39. (A.1.)	Manipulation der Daten	1) keine Hinweise auf Ergebnismanipulation 3) Hinweise auf Ergebnismanipulation	A1

5. Beurteilung der Ergebnisse der Studie		
D1	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual". (Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt) (ja = 1; nein = 3; nicht zutreffend = freilassen)	1 D1
D2	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D2
D3	In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D3

6. Praxistransfer

	Kriterium	Operationalisierung	Rating
K40. (C.10.)	Spezifikation und Herstellbarkeit notwendiger Settingbedingungen	1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team) 2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team) 3) Notwendige Settingbedingungen nicht herstellbar	C10 1
K41. (C.11.)	Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation	1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar 2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar 3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar	C1 1
K42. (C.12.)	Spezifikation und Erfassbarkeit relevanter Patientenmerkmale	1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar 2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar 3) Relevante Patientenmerkmale praktisch nicht erfassbar	C12 1
K43. (C.13.)	Spezifikation und Herstellbarkeit relevanter Treatmentaspekte	1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual) 2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung) 3) Relevante Treatmentmerkmale praktisch nicht herstellbar	C13 1

7. Gesamtbewertung der Studie

Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)			
Durchschnittswert für B) Interne Validität (Summe / 12)			
Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)			
Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4)			

Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

<p>Kriterien für 'A. Allgemeine methodische Qualität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde <p>und</p> <ol style="list-style-type: none"> wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist. 	
<p>Kriterien für 'B. Interne Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist. 	
<p>Kriterien für 'C. Externe Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde. 	
<p>Die Ergebnisse der Studie belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>	

Abschließende Bewertung

Allgemeine methodische Qualität der Studie

Es gibt starke Einschränkungen der allgemeinen methodischen Qualität der Studie. Es erfolgt keine objektive und reliable Diagnosestellung einer klinisch relevanten Störung z.B. einer Anpassungsstörung (Patienten mit Krebserkrankung). Die Angaben über sozio-demographische Daten, Ein- & Ausschlusskriterien, Zugangswege zu der Studie, Dropout – Raten sind unzureichend (Für Details wird auf die Dissertationsschrift von Kalliner, 1981 verwiesen.) Es ist unklar, wie die Patienten der Gesprächsgruppe bzw. der Vergleichsgruppe zugeordnet werden. Veränderungswerte der Kontrollgruppe werden nicht direkt angegeben. Die Beschreibung der Statistik ist äußerst vage („bei Veränderungsmessung gängige Verfahren“) und lässt sich nur durch Hinzuziehen der Dissertation besser nachvollziehen. Es ist unklar, inwiefern sich Therapie- und Vergleichsgruppe voneinander unterscheiden, alle Ergebnisse werden lediglich in % angegeben ohne Varianz- und Signifikanzangaben.

Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/ Risiken)

27 % der Gesprächspatienten längerfristig in verschiedenen Bereichen ihrer Persönlichkeit „deutlich verbessert“, 37 % „leicht verbessert“, 30% „längerfristig unverändert“, 6 % „leicht verschlechtert“

(Gesamteinschätzung der individ. Patienten von 3 Experten)

mehr (soziale) Aktivitäten, Verbesserung (familiärer) Beziehungen, Verringerung der Nervosität (FPI), weniger Krebsangst, höhere körperliche und seelische Belastbarkeit, günstigere Einstellung zu der Krankheit und zum Leben

97 % der Gesprächspatienten halten personenzentrierte Gesprächsgruppen im Zusammenhang mit der Krankheit für wichtig und hilfreich

Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Die interne Validität ist schwer abzuschätzen, da etliche Angaben fehlen. Insbesondere ist unklar, wie die Patienten den beiden Gruppen zugeteilt wurden, ob die Gruppen zum Zeitpunkt prä äquivalent waren, ob es zusätzliche nicht randomisierte Einflüsse gab.

Externe Validität

(Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)

Die externe Validität ist bzgl. der Behandler gut und bzgl. der Repräsentativität der Intervention mittelmäßig (Gruppentherapie, 3 von 11 Sitzungen erstrecken sich über einen ganzen Tag). Die Frequenz (1x/ Woche) der Intervention ist repräsentativ und gut auf ein ambulantes Setting im dt. Versorgungssystem übertragbar

Stärken der Studie:

Der Patient evaluiert das Therapeutenverhalten (Realisierung der drei Basisvariablen nach Rogers).

Schwächen der Studie:

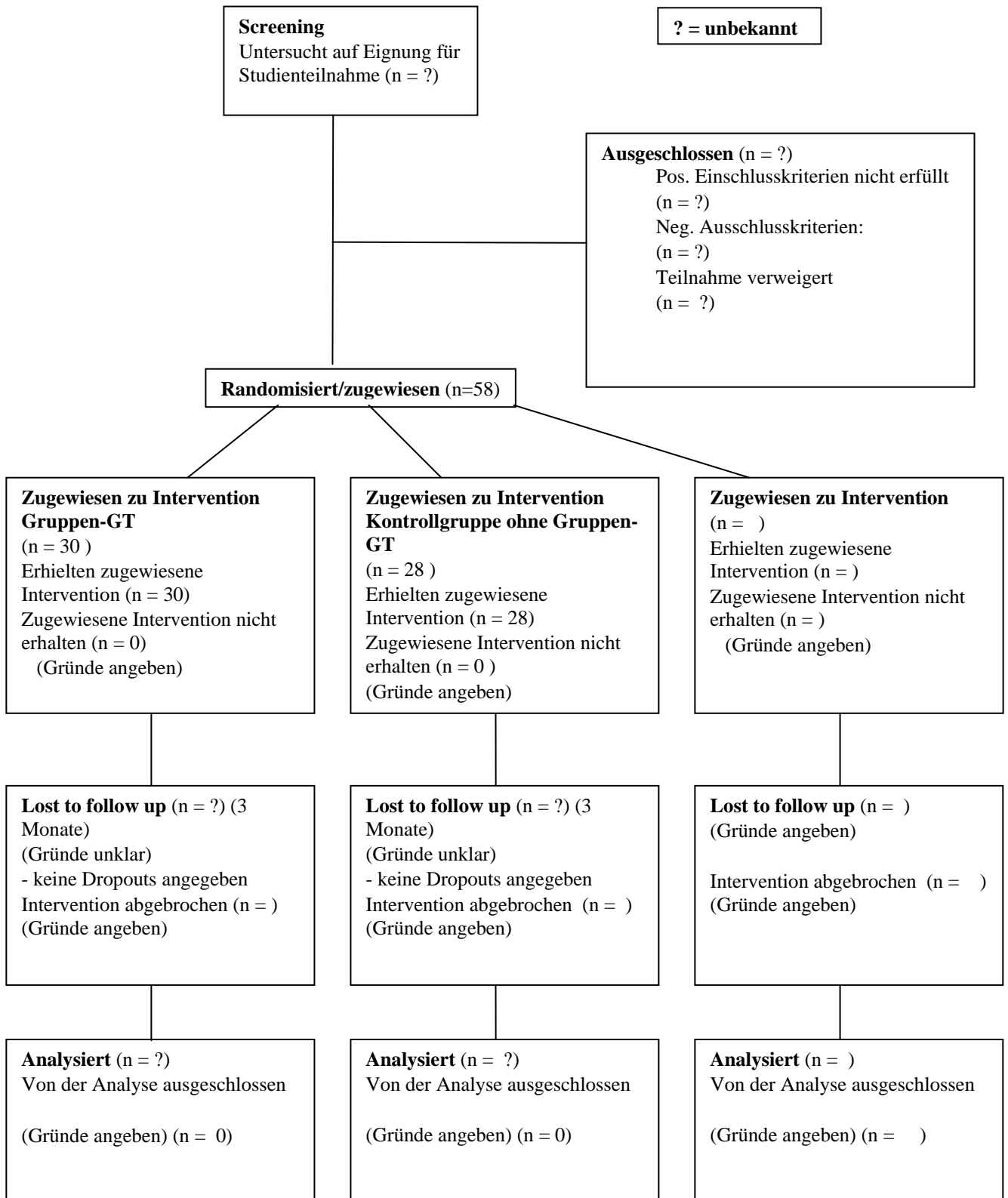
Es wird keine klare Diagnose einer psychischen Störung und komorbider Störungen gestellt. Es kann sicherlich nur bei einem Teil der Krebspatienten von einer Störung im Sinne einer Anpassungsstörung ausgegangen werden. Es besteht das Risiko prognostisch wichtige komorbide Erkrankungen zu übersehen. Das Design erlaubt nicht, die Effekte „personenzentrierter Gruppengespräche“ von denen „nicht-personenzentrierter Gruppengespräche“, zu unterscheiden. Es gibt etliche methodische Mängel, z.B. ist aus dem Artikel nicht ersichtlich, welche Unterschiede statistisch und/oder klinisch relevant sind

Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie ist aufgrund methodischer Mängel nicht für einen Nachweis der Wirksamkeit der GT geeignet, liefert jedoch durchaus Hinweise auf eine Wirksamkeit der GT bei Reaktionen auf schwere Belastungen/ Anpassungsstörungen.

Es gibt starke methodische Mängel der Studie. Einige Unklarheiten lassen sich durch das Studium der Dissertationsschrift von Kalliner, 1981 beseitigen. Problematisch ist aber, dass das Design nicht erlaubt, die Effekte „personenzentrierter Gruppengespräche“ von denen „nicht-personenzentrierter Gruppengespräche“ zu unterscheiden. Möglicherweise ist somit zumindest ein Teil der berichteten Effekte auf eine unspezifische „Heilsamkeit“ des Austausches mit anderen Betroffenen zurückzuführen.

8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

K5 (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

K14 (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

K33 (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

K39 (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

Anmerkungen zur durchgeführten Studienbewertungen

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

MT

Studie:

Autoren:

Eckert et al.

Titel:

Die langfristigen Veränderungen der Borderline-Symptomatik bei Patienten nach klientenzentrierter Gruppenpsychotherapie

Quelle:

Psychother. Psychosom. med. Psychol., 2000, 50: 140-146

1. Basisinformationen

1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

- Gruppenpsychotherapie
- klientenzentriertes Störungs- und Therapiekonzept (Rogers; Biermann-Ratjen et al., 1995)
- modifiziert durch störungsspezifische Behandlungsregeln (Eckert & Wuchner 1996), bzgl. des Settings Gruppe orientiert am Wirkfaktorenmodell von Yalom (Yalom, 1974, Eckert & Biermann-Ratjen, 1985)

1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

- Wie entwickeln sich die Symptome von Patienten mit einer zweifelsfrei diagnostizierten Borderline-Persönlichkeitsstörung im Anschluss an eine ambulante klientenzentrierte Gruppenpsychotherapie?
- Veränderungen des borderlinespezifischen Erlebens und Verhaltens der Borderline-Patienten im Vergleich mit dem von depressiven und schizophrenen Patienten

1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Borderline-Persönlichkeitsstörung, Depression oder Schizophrenie nach ICD-9; dem Diagnostischen Interview für Borderline-Patienten (DIB, Gunderson et al., 1981); dem Rorschach-Test (Kriterien zur Bestimmung der borderlinespezifischen Abwehr, Kwawer et al., 1980)

- Borderline: DIB (Urteil Interviewer und unabh. Videorating); Rorschach
- Depression, Schizophrenie: klinisches Urteil nach ICD-9; DIB (Urteil Interviewer und unabh. Videorating)

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews)	1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnosestellung	1 A2
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	2 C1

2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschlusskriterien:
- Borderline-PSKS, Depression, Schizophrenie (s.o.)
Ausschlusskriterien:
- keine

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	1 B1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	2 B2
K5. (C.3.)	Selektivität der Stichprobe aufgrund der Ausschlusskriterien	1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten 2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen) 3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien	1 C3

2.3 () Setting/Patientenzugang

- ambulante und psychiatrische Patienten zweier psychiatrischer Universitätskliniken

	Kriterium	Operationalisierung	Rating
K6. (C.2.)	Art der Rekrutierung der Stichprobe	1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege 2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege 3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung)	1 C

2.4 (7) Zentrale Patientenmerkmale (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- Borderline (N = 14): mittleres Alter = 27,6 Jahre (SD 6,2), 29% männlich, 71% weiblich
- Schizophrenie (N = 13): mittleres Alter = 28,5 Jahre (SD 8,0), 46% männlich, 54% weiblich
- Depression (N = 16): mittleres Alter = 30,5 Jahre (SD 10,1), 38% männlich, 62% weiblich

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Keine Angaben

	Kriterium	Operationalisierung	Rating
K7. (A.16.)	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	9 A16

2.6 (9) Anzahl gescreener, eingeschlossener, zugewiesener und ausgewerteter Patienten (Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- Anzahl gescreener Patienten unbekannt
- 47 Patienten werden in die Studie eingeschlossen
- Verteilung nach Diagnosen:
N = 14 Borderline, N = 14 Schizophrenie, N = 19 Depression
- der klientenzentrierten Gruppenpsychotherapie zugewiesen:
N = 14 Borderline, N = 6 Schizophrenie, N = 10 Depression
- Dropouts im Verlauf der Therapie:
N = 2 Borderline, N = 1 Schizophrenie, N = 3 Depression
- Vergleichsgruppe ohne Gruppenpsychotherapie (Behandlung nach klinischem Standard):
N = 0 Borderline, N = 8 Schizophrenie, N = 9 Depression

- 43 Patienten werden nach im Mittel 4 Jahre nach der Untersuchung prä (ca. 1,5 Jahre nach Abschluss der Therapie) katamnestisch untersucht

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 3
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 2
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A7 2
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	2 A3
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 1

2.7 (10) Vergleichbarkeit der Gruppen

Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen:

- # Störungsgruppen vergleichbar bzgl. Alter, unter den Borderlinern, wie typisch für die Störung, mehr Frauen als in den anderen Störungsgruppen
- # die Störungsgruppen haben unterschiedliche Ausgangswerte bzgl. fast aller erhobenen DIB Werte („Affekte“, „psychot. Symptome“, „zwischenmenschliche Beziehungen“, „Gesamtwert“)
- # Vergleichbarkeit der Gruppen mit und ohne Gruppenspsychotherapie kann nicht abgeschätzt, da keine separaten Angaben

Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer):

- ist gegeben (2 Sitzungen a 1,5 h pro Woche, ca. 100 Sitzungen insg., Zeitraum: gutes Jahr)

Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen):

- nicht bekannt

Vergleichbarkeit der Therapieadhärenz:

- ist gegeben

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	<p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p>	B9 9
K14. (A.18.)	Drop-out-Analysen	<p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>	A18 2

3. Intervention

**3.1 (14) Studiendesign /
Behandlungsarme**

Behandlungsgruppe Gesprächspsychotherapie, innerhalb der Behandlungsgruppe Vergleich der Entwicklung bei drei verschiedenen Störungsbildern
- eine Teil der Patienten mit den Störungsbildern Schizophrenie oder Depression erhält keine Gruppentherapie, sondern Behandlung nach klinischem Standard (nicht genauer spezifiziert)

**3.2 (11) Behandlung Experimental-
gruppe
(Art, Dauer, Frequenz)**

- Gruppenpsychotherapie
- zwei Therapeuten: 1 Mann, 1 Frau
- Patienten aller drei Störungsbilder gemischt in geschlossenen Gruppen von 7-10 Patienten
- klientenzentriertes Störungs- und Therapiekonzept (Rogers; Biermann-Ratjen et al., 1995)
- modifiziert durch störungsspezifische Behandlungsregeln (Eckert & Wuchner 1996), bzgl. des Settings Gruppe orientiert am Wirkfaktorenmodell von Yalom (Yalom, 1974, Eckert & Biermann-Ratjen, 1985)
- Ziel: Borderline-Patienten sollen sich mit den Erfahrungen auseinandersetzen, durch die sie sich in ihrem Selbstkonzept und damit existenziell bedroht fühlen

- 2 Sitzungen a 1,5 h pro Woche, ca. 100 Sitzungen insg., Zeitraum: gutes Jahr

**3.3 (12) Behandlung Kontroll-
gruppe
(Art, Dauer, Frequenz)**

- Behandlung nach klinischem Standard (?)

**3.4 (13) Evtl. weitere Kontroll-
gruppen
(Art, Dauer, Frequenz)**

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 2
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	3 B4
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	9 B5
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	3 B6
K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	1 C4

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantziellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 9
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	9 C6
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	9 C7a
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	9 C7b
	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	1 C7c

**3.5 (19) Beobachtungszeitraum/
Messzeitpunkte/Katamnese**

- Beobachtungszeitraum ca. 4 Jahre, 2 Messzeitpunkte, Eingangsuntersuchung, keine Post-Messung, Katamnese ca. 4 Jahre nach Eingangsuntersuchung (etwa 1,5 Jahre nach Therapieende)

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 1
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	9 10
K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	B11 1

3.6 (15) Anzahl der Zentren

Wahrscheinlich 1

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...)

Keine Randomisierung

3.8 (17) ggf. Concealment der Randomisierung

Keine Randomisierung

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 9
K27. (B.8.)	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ($n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	B8 3

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

Verbesserung der Fähigkeit der Borderline-Patienten, sich selbst und andere realistisch als kohärente und integrierte Individuen anzusehen, dadurch Abbau/ Reduktion von Selbstverletzungen, Suizidhandlungen, psychot. Zuständen, intensive + instabile Beziehungen

- Borderline Verhalten und Erleben: DIB-Gesamtwert

4.2 (21) Sekundäre Zielkriterien

- Impuls-Kontrolle
- Affekte
- psychotische Symptome
- zwischenmenschliche Beziehungen

- jeweils DIB-Wert

4.3 () Verblindung der Erhebung der Zielkriterien

Nicht bekannt

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,	A7 2
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend	1A8
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)	1C9
K31. (A.9.)	Multiple Informationsquellen	1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle	3A9
K32. (A.10.)	Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)	1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler 2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern 3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten	A10 1

4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)

- 2-faktorielle Varianzanalyse pro DIB-Wert mit den Faktoren „Diagnose“ und „Zeitpunkt“
 - die unterschiedlichen Behandlungsweisen (Gruppentherapie vs. klin. Standard) innerhalb der Gruppe der Schizophrenen und der depressiven Patienten findet statistisch keinen Eingang

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	1 A6
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung	2 A13
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	1 A14

4.5 (22) Ergebnisse

Prä-Katamnese:
 - Borderline-Patienten weisen zum Zeitpunkt der Nachuntersuchung, im Durchschnitt vier Jahre nach der Erstuntersuchung, bedeutsam weniger borderlinespezifische Symptome auf als zum Zeitpunkt der Erstuntersuchung.
 - Die schizophrenen und depressiven Patienten haben zum Zeitpunkt prä eine in etwa ähnlich ausgeprägte, im Vergleich zu den Borderline-Patienten erwartungsgemäß jedoch signifikant geringere borderlinespezifische Symptomatik, die sich zur Katamnese signifikant verringert.
 - Obwohl in allen drei untersuchten Gruppen die globale borderlinespezifische Symptomatik zurückgeht, ist die Abnahme in der Gruppe der Borderline-Patienten am größten.
 - Die Bereiche „Impulskontrolle“ und „psychotische Symptome“ tragen jeweils einzeln und quantitativ am stärksten zur Reduktion des gesamtborderlinespezifischen Erlebens und Verhaltens bei.
 - In den Bereichen „Affekte“ und „zwischenmenschliche Beziehungen“ vermindert sich das borderline-spezifische Verhalten bei allen Störungsgruppen in ähnlichem Maße.
 - Zum Erstuntersuchungszeitpunkt sind bis auf einen Patienten alle klinisch diagnostizierten Borderline-Patienten auch DIB-identifizierte Borderline-Patienten. Von den schizophrenen und depressiven Patienten erreichen und/oder überschreiten zwei bzw. ein Patient den Schwellenwert des DIB. Zur katamnestischen Untersuchung erzielt keiner der Kontrollgruppen-Patienten den Cut-off-Wert, und nur noch zwei Borderline-Patienten erfüllen die Kriterien des DIB.

	Kriterium	Operationalisierung	Rating
K36. (A.11.)	Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten	1) für alle Outcome-Variablen berichtet 2) ausschließlich für die primären Zielkriterien berichtet 3) nicht für alle primären Zielkriterien berichtet	A11 1
K37. (B.12.)	Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte)	1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe 2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig 3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG	B12 3

4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen

- keine

	Kriterium	Operationalisierung	Rating
K38. (A.12.)	Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen	1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen 3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet	A12 3

4.7 (24) Fazit/Schlussfolgerungen der Autoren

- bedeutsame Verbesserung borderlinespezifischer Symptomatik durch eine auf 100 Sitzungen beschränkte klientenzentrierte Gruppenpsychotherapie, der Großteil der Patienten erfüllt zur Katamnese nicht mehr die Kriterien einer Borderline-Störung nach DIB
 - ein Anteil der Veränderungen könnte aber auch auf spontane Remission zurückzuführen sein

	Kriterium	Operationalisierung	Rating
K39. (A.1.)	Manipulation der Daten	1) keine Hinweise auf Ergebnismanipulation 3) Hinweise auf Ergebnismanipulation	A1

5. Beurteilung der Ergebnisse der Studie		
D1	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual". (Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt) (ja = 1; nein = 3; nicht zutreffend = freilassen)	D1
D2	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D2
D3	In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D3

6. Praxistransfer

	Kriterium	Operationalisierung	Rating
K40. (C.10.)	Spezifikation und Herstellbarkeit notwendiger Settingbedingungen	1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team) 2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team) 3) Notwendige Settingbedingungen nicht herstellbar	C10
K41. (C.11.)	Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation	1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar 2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar 3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar	C3
K42. (C.12.)	Spezifikation und Erfassbarkeit relevanter Patientenmerkmale	1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar 2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar 3) Relevante Patientenmerkmale praktisch nicht erfassbar	C12 1
K43. (C.13.)	Spezifikation und Herstellbarkeit relevanter Treatmentaspekte	1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual) 2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung) 3) Relevante Treatmentmerkmale praktisch nicht herstellbar	C13 1

7. Gesamtbewertung der Studie

Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)			
Durchschnittswert für B) Interne Validität (Summe / 12)			
Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)			
Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4)			

Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

<p>Kriterien für 'A. Allgemeine methodische Qualität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> 2. wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde <p>und</p> <ol style="list-style-type: none"> 3. wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist. 	
<p>Kriterien für 'B. Interne Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> 2. der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist. 	
<p>Kriterien für 'C. Externe Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> 2. das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> 3. bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde <p>und</p> <ol style="list-style-type: none"> 4. bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde. 	
<p>Die Ergebnisse der Studie belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>	

Abschließende Bewertung

Allgemeine methodische Qualität der Studie

Die Diagnosestellung einer Borderline-Persönlichkeitsstörung erfolgte objektiv und reliabel (DIB, zwei Beurteiler).

Methodische Mängel ergeben sich daraus, dass es keine Post-Untersuchung gibt (nur Voruntersuchung und Katamnese). Das Studiendesign ist unklar: Während alle Borderline-Patienten Gruppentherapie erhalten, bekommen die Patienten der anderen Störungsgruppen zum Teil „klinische Standardintervention“. Für die Borderline-Patienten existiert somit keine echte Vergleichsgruppe.

Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

deutliche Verbesserung der borderline-spezifischen Symptomatik bei Patienten mit Borderline-Störung, Schizophrenie, Depression (Vergleich Prä vs. Katamnese)

stärkste Verbesserung bei Patienten mit Borderline-Störung

der Großteil der Patienten erfüllt zur Katamnese nicht mehr die Kriterien einer Borderline-Störung

Anteil einer Spontanremission an der Veränderung nicht ausgeschlossen

Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Die interne Validität ist Mangelhaft. Die Zuordnung der Patienten zu den Gruppen erfolgt nicht randomisiert. Alle Borderline-Patienten erhalten Gruppentherapie, während Patienten der anderen Störungsgruppen zum Teil die „klinische Standardintervention“ erhalten. Bei den Borderline Patienten gibt es keine unbehandelte Vergleichsgruppe. Die Vergleichsbedingung „klinische Standardintervention“ wird nicht beschrieben. Für die statistischen Vergleiche werden Patienten jeweils eines Störungsbildes mit und ohne Gruppentherapie gepoolt. Die Qualität der Durchführung der Therapie wird nicht überprüft; die Qualifikation der Behandler ist nicht angegeben

Externe Validität

(Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)

Die Patientenstichprobe ist nicht selektiv, die Repräsentativität der Behandler ist nicht bekannt. Die Frequenz (2x/ Woche) der Intervention ist repräsentativ und gut auf ein ambulantes Setting im dt. Versorgungssystem übertragbar.

Stärken der Studie:

Diagnose und Verlaufsmessung u.a. über DIB. Katamnesezeitraum von ca. 1,5 J..

Schwächen der Studie:

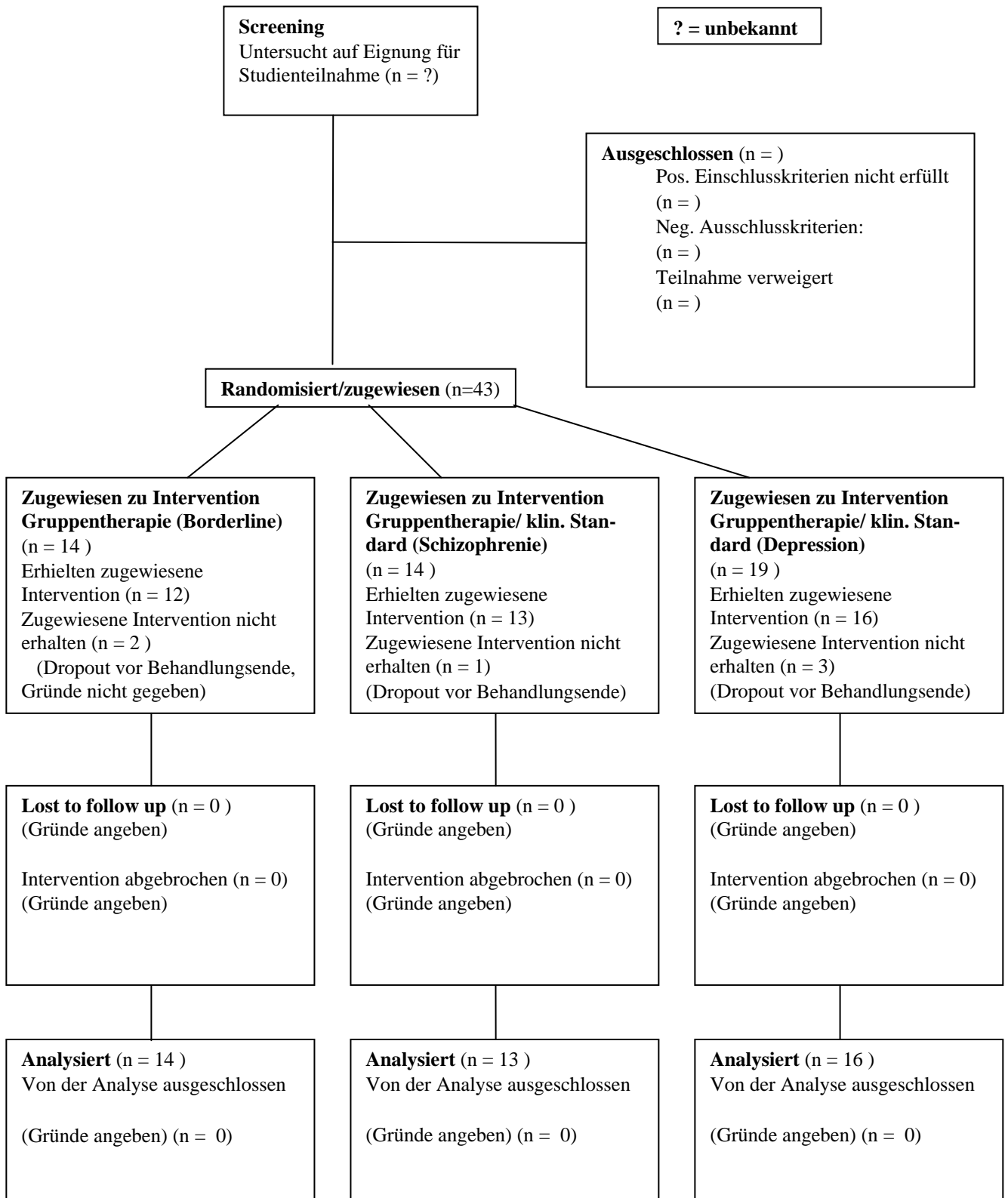
Es ist kein klares Versuchsdesign zu erkennen. Nur schizophrene oder depressive Patienten erhalten teilweise Gruppentherapie und teilweise Standardintervention. Diese beiden Interventionen werden nicht verglichen (es sind sehr kleine Gruppen). Bei den Borderline Patienten existiert keine echte Vergleichsgruppe. Beim Vergleich der Veränderungswerte der drei Störungsgruppen werden schizophrene und depressive Patienten jeweils über beide Interventionsbedingungen gepoolt. Die Kontrollbedingung „klinische Standardintervention“ wird nicht näher beschrieben. Auf eine Datenerhebung unmittelbar nach Ende der Therapie wird verzichtet.

Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie ist mit Einschränkungen geeignet für einen Nachweis der Wirksamkeit der GT bei Persönlichkeitsstörungen.

Methodische Einschränkungen ergeben sich insb. aus dem Design, das keine echte Vergleichsgruppe bereitstellt, um die Therapiewirkung von einer Remissionswirkung zu trennen. Zudem gibt es keinen Messzeitpunkt unmittelbar nach dem Therapiezeitraum, sondern nur eine ca. 1,5 Jahres-Katamnese.

8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

K5 (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

K14 (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

K33 (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

K39 (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

Anmerkungen zur durchgeführten Studienbewertungen

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

MS

Studie:

Autoren:

Ends et al.

Titel:

A Study of Three Types of Group Psychotherapy with Hospitalized Male Inebriates

Quelle:

Quarterly journal of studies on alcohol 1957; 18: 263-277

1. Basisinformationen

1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Klientenzentrierte Gruppenpsychotherapie

1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Vergleich von 3 stationären Gruppenpsychotherapieverfahren hinsichtlich ihrer Wirksamkeit auf die Verbesserung des sozialen Umgangs/der Sozialverträglichkeit bei männlichen Alkoholikern.

Hypothese: Verhaltenstherapie ist am wirksamsten, Gesprächspsychotherapie am zweitwirksamsten und psychodynamische Therapie am wenigsten wirksam.

1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Alkoholabhängigkeit, Anwendungsbereich 2.1: Abhängigkeit von Alkohol, Drogen oder Medikamenten nach vorangegangener Entgiftungsbehandlung.

Nicht formal diagnostiziert, aber jeder Patient erfüllte die Kriterien von > 1 Monat unkontrollierten Trinkens (der Median war 6 Jahre) sowie eine lange Vorgeschichte exzessiven Trinkens

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnostikstellung (mittels (teil-) standardisierter Interviews)	1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnosestellung	A2 2
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	C1 2

2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschlusskriterien:

- Inanspruchnahmepopulation von Alkoholikern eines Krankenhauses (Willmar Minnesota State Hospital)
- männliches Geschlecht
- Alter 25-45

Ausschlusskriterien:

- Ergebnisse im Army General Classification Test (Intelligenztest) < 40. Perzentil (Ausschluss von Minderbegabung)

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	B1 1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	B2 1
K5. (C.3.)	Selektivität der Stichprobe aufgrund der Ausschlusskriterien	1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten 2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen) 3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien	C3 2

2.3 () Setting/Patientenzugang

Inanspruchnahmepopulation eines Krankenhauses

	Kriterium	Operationalisierung	Rating
K6. (C.2.)	Art der Rekrutierung der Stichprobe	1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege 2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege 3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung)	C2 1

2.4 (7) Zentrale Patientenmerkmale
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

Stichprobe von n = 63
Alter 36.9 Jahre (SD 5.9, Range 25-45)
Keine weiteren Angaben.

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Keine Angabe

	Kriterium	Operationalisierung	Rating
K7. (A.16.)	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	A16 2

2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

Keine Angabe zur Gesamtzahl gescreenter Patienten
96 Patienten wurden der Behandlung zugewiesen
63 completers (nur die wurden ausgewertet)

Aufteilung in 16 Therapiegruppen mit je 6 Patienten
Gruppengrößen:
VT = 16
GPT = 15
Analytisch = 15
Kontrollgruppe = 17

Zur Katamnese:
VT = 14
GPT = 12
Analytisch = 15
Kontrollgruppe = 17

Keine Angabe zu Drop-outs.

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 2
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 3
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A17 3
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A3 2
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 1

2.7 (10) Vergleichbarkeit der Gruppen

<p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen Keine Angabe</p> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer) Alle Gruppen gingen über 15 Sitzungen (3x wöchentlich). Es gab insgesamt 4 Therapeuten, die jeweils alle 4 unterschiedlichen Behandlungsformen praktizierten. Die Therapeuten waren vergleichbar in ihrer klinischen und psychotherapeutischen Vorerfahrung (> 2 Jahre). Alle erhielten 1 Jahr intensives Training und Ausbildung für die Studie. Alle Studientherapien wurden auf Tonband aufgenommen und supervidiert (ausführlich rückgemeldet) und die Qualität der Behandlung zu sichern. Die Vergleichbarkeit ist wahrscheinlich.</p> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen) Keine Angabe</p> <p>Vergleichbarkeit der Therapieadhärenz Keine Angabe</p>
--

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	<p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanten Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p>	<p>B9</p> <p>3</p>
K14. (A.18.)	Drop-out-Analysen	<p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>	<p>A18</p> <p>3</p>

3. Intervention

**3.1 (14) Studiendesign /
Behandlungsarme**

Es wurden 4 Behandlungsgruppen miteinander verglichen, davon 3 Psychotherapien und eine nicht-therapeutische Diskussionsgruppe als Kontrollgruppe

**3.2 (11) Behandlung Experimental-
gruppe
(Art, Dauer, Frequenz)**

Klientenzentrierte Gruppenpsychotherapie über 15 Sitzungen (3x wöchentlich), Literaturangabe zum Verfahren (Rogers).

**3.3 (12) Behandlung Kontroll-
gruppe
(Art, Dauer, Frequenz)**

Lerntheoretische (= verhaltenstherapeutische) Gruppenpsychotherapie über 15 Sitzungen (3x wöchentlich), Literaturangabe zum Verfahren.

**3.4 (13) Evtl. weitere Kontroll-
gruppen
(Art, Dauer, Frequenz)**

Psychoanalytische Gruppenpsychotherapie über 15 Sitzungen (3x wöchentlich), Literaturangabe zum Verfahren.

Sowie

Diskussionsgruppen als Kontrollgruppe, auch über 15 Sitzungen (3x wöchentlich). Die Therapeuten blieben dabei neutral.

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 2
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	B4 1
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	B5 1
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	B6 1

K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	C4 1
-----------------------	--	---	---------

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 9
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	C6 9
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	C7a 1
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	C7b 9
	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	C7c 2*

* Das intensive Training war vermutlich notwendig, um die eher geringe Erfahrung der Therapeuten auszugleichen. Die Behandlung war nicht spezifisch für die Studie, sondern entsprach der üblichen Praxis.

**3.5 (19) Beobachtungszeitraum/
Messzeitpunkte/Katamnese**

Erhebungen vor Therapiebeginn, nach Therapieende, Katamnesen nach 6, 12 und 18 Monaten
--

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 2*
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	B10 1

K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	B11 1
------------------------	-------------------	---	--------------

* Es gab kontinuierliche Rückmeldung, diese war aber vermutlich notwendig, um die eher geringe Erfahrung der Therapeuten auszugleichen. Die Behandlung war nicht spezifisch für die Studie, sondern entsprach der üblichen Praxis. Das Monitoring erfolgte zur Qualitätssicherung.

3.6 (15) Anzahl der Zentren 1

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...) Keine Angabe (vermutlich nicht randomisiert).

3.8 (17) ggf. Concealment der Randomisierung Keine Angabe

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 9
K27. (B.8.)	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ($n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	B8 9

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

Q-Sorts zu Selbst- und Idealbild (100 Items, z.B. „mir ist es unangenehm, über mich zu sprechen“, „Ich bin eine verantwortungsbewusste Person“, „Ich fühle mich hoffnungslos“) und entsprechend berechnete Indices (siehe 4.4). Als externes Kriterium gab es ein von Experten übereinstimmend als am gesündesten eingeschätztes Q-Sort.

4.2 (21) Sekundäre Zielkriterien

Rückfallraten/Verbesserung der Alkoholproblematik zu den Katamnesen. Benutzt wurden standardisierte Interviews und Wiederaufnahmeraten in die Klinik, es erfolgte eine Klassifikation in:

- „deutlich gebessert“, wenn es keine Hinweise auf erneuten Alkoholkonsum gab über den gesamten Katamnesezeitraum
- „vielleicht (*possibly*) gebessert“, wenn es nur innerhalb der ersten 3 Monate nach Entlassung 1-2 kurze Rückfälle gab
- „keine Veränderung“: Fortsetzung alter Konsumgewohnheiten ohne stationäre Wiederaufnahme
- stationäre Wiederaufnahme erforderlich

4.3 () Verblindung der Erhebung der Zielkriterien

Bei den Katamnesen waren die Untersucher blind zu der vorhergehenden Art der Behandlung

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	<ol style="list-style-type: none"> 1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien, 	A7 1
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	<ol style="list-style-type: none"> 1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend 	A8 2
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	<ol style="list-style-type: none"> 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung) 	C9 2
K31. (A.9.)	Multiple Informationsquellen	<ol style="list-style-type: none"> 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle 	A9 2
K32. (A.10.)	Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)	<ol style="list-style-type: none"> 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler 2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern 3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten 	A10 2

4.4 () **Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

Q-Sort Technik. Individuelle Korrelationen wurden berechnet zwischen der Selbstdarstellung vor Therapiebeginn, Idealbild vor Therapiebeginn, Selbstdarstellung bei Therapieende, Idealbild nach Therapieende und einer von Experten objektiv übereinstimmend als am gesündesten festgelegte Darstellung (criterion-sort genannt).

8 verschiedene Indices der Veränderungen durch die Therapie wurden berechnet mittels Fisher's Z-Transformationen (z.B. ob sich das Selbstbild im Vorher-Nachher-Vergleich in eine gesunde Richtung entwickelt, ob sich das Selbst mehr entwickelt als das Idealbild, ob sich das Selbstbild dem Idealbild annähert im Laufe der Therapie etc.).

Häufigkeitsvergleiche zwischen den Gruppen zur Katamnese

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	A6 1
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung	A13 2
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	A14 1

4.5 (22) **Ergebnisse**

Verhaltenstherapie:
4 der 8 Indices verändern sich signifikant. Die Differenz zwischen Selbstbild und Idealbild wird größer, das Selbstbild ändert sich mehr als das Idealbild, das Selbstbild entwickelt sich in eine andere Richtung als das Idealbild vor Therapiebeginn und das Idealbild entwickelt sich in eine andere Richtung als das Selbstbild vor Therapiebeginn. Nur diese Behandlungsgruppe zeigte ein gesünderes Idealbild nach Therapieende. Die Veränderungen des Selbstbildes gehen nicht in eine gesunde Richtung. Interpretiert als verbesserte und damit kritischere Selbstwahrnehmung (ohne entsprechende positive Entwicklungen).

Gesprächspsychotherapie:
4 der 8 Indices verändern sich signifikant. Das Selbstbild verändert sich durch die Therapie in eine gesunde Richtung, das Selbstbild nähert sich dem Idealbild an durch die Therapie, das Selbstbild entwickelt sich in Richtung des Idealbildes nach Therapieende und das Idealbild bei Therapieende nähert sich dem entwickelten Selbstbild an. Interpretiert als integrierende, gesunde Entwicklung (sowohl Selbst- als Idealbild ändern sich).

Analytische Therapie:
4 der 8 Indices verändern sich signifikant. Das Selbstbild ändert sich stärker als das Idealbild, die Veränderungen ergeben eine größere Selbstakzeptanz, die Entwicklung des Selbstbildes geht sowohl in Richtung des Idealbildes vor Therapiebeginn als auch in Richtung des Idealbildes nach Therapieende (das Idealbild bleibt stabil). Interpretiert als einseitige Entwicklung zu einem nicht hinterfragten Idealbild.

Kontrollgruppe:

2 der 8 Indices zeigen signifikante Veränderungen. Das Selbstbild ändert sich stärker als das Idealbild (in ungesunde Richtung) und das Idealbild entwickelt sich ebenfalls in Richtung geringerer Gesundheit. Interpretiert als fehlende therapeutische Wirksamkeit in der Kontrollgruppe.

Ergebnisse der Katamnese:

Die klientenzentrierte und analytische Therapie ist der Kontrollgruppe signifikant überlegen in den Raten der Besserung der Alkoholabhängigkeit. Nur die klientenzentrierte Therapie erzielte signifikant geringere Wiederaufnahmeraten als die Kontrollgruppe.

7 der nachuntersuchten Patienten (58.3%) waren deutlich gebessert (keine Rückfälle), 1 (8%) vielleicht gebessert (nur kurze Rückfälle direkt nach Entlassung), 3 (25%) wurden stationär wiederaufgenommen (im Gegensatz zu 53% der Kontrollgruppe).

	Kriterium	Operationalisierung	Rating
K36. (A.11.)	Vollständige Darstellung der Ergebnisse zu allen Outcome-Maßen und zu allen relevanten Messzeitpunkten	1) für alle Outcome-Variablen berichtet 2) ausschließlich für die primären Zielkriterien berichtet 3) nicht für alle primären Zielkriterien berichtet	A11 1
K37. (B.12.)	Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte)	1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe 2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig 3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG	B12 1

4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen

Keine Angabe

	Kriterium	Operationalisierung	Rating
K38. (A.12.)	Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen	1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet	A12 3

4.7 (24) Fazit/Schlussfolgerungen der Autoren

Die Therapiegruppen unterscheiden sich in den erreichten Veränderungen. Die klientenzentrierte Therapie ist den anderen Verfahren überlegen, da sich hier sowohl das Selbst- als auch das Idealbild in einem konstruktiven Prozess persönlichen Wachstums entwickeln. In der Katamnese ist die klientenzentrierte Therapie den anderen Verfahren deutlich überlegen.

	Kriterium	Operationalisierung	Rating
K39. (A.1.)	Manipulation der Daten	1) keine Hinweise auf Ergebnismanipulation 3) Hinweise auf Ergebnismanipulation	A1 1

5. Beurteilung der Ergebnisse der Studie		
D1	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual". (Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt) (ja = 1; nein = 3; nicht zutreffend = freilassen)	D1
D2	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D2
D3	In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D3 1

6. Praxistransfer

	Kriterium	Operationalisierung	Rating
K40. (C.10.)	Spezifikation und Herstellbarkeit notwendiger Settingbedingungen	1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team) 2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team) 3) Notwendige Settingbedingungen nicht herstellbar	C10 1
K41. (C.11.)	Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation	1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar 2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar 3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar	C11 1
K42. (C.12.)	Spezifikation und Erfassbarkeit relevanter Patientenmerkmale	1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar 2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar 3) Relevante Patientenmerkmale praktisch nicht erfassbar	C12 9
K43. (C.13.)	Spezifikation und Herstellbarkeit relevanter Treatmentaspekte	1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual) 2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung) 3) Relevante Treatmentmerkmale praktisch nicht herstellbar	C13 1

7. Gesamtbewertung der Studie

Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)			
Durchschnittswert für B) Interne Validität (Summe / 12)			
Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)			
Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4)			

Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

<p>Kriterien für 'A. Allgemeine methodische Qualität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde <p>und</p> <ol style="list-style-type: none"> wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist. 	
<p>Kriterien für 'B. Interne Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist. 	
<p>Kriterien für 'C. Externe Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde <p>und</p> <ol style="list-style-type: none"> bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde. 	
<p>Die Ergebnisse der Studie belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>	

Abschließende Bewertung

Allgemeine methodische Qualität der Studie

Die Studie entspricht in ihrer allgemeinen Qualität dem Stand der Forschung zu Zeiten ihrer Publikation. Die Q-Sort-Technik erlaubt differenzierte inhaltliche Aussagen zu Prozessforschung und Therapieoutcome, die mittels moderner üblicher Messinstrumente und Auswertungstechniken nicht möglich sind. Die Ergebnisse der Q-Sort-Technik erlauben allerdings keine Rückschlüsse auf die spezifische (Alkohol-) Problematik der untersuchten Patienten. Da es sich um stationäre Psychotherapie handelt, wird der Alkoholkonsum vermutlich verboten gewesen sein und die Abstinenz kontrolliert worden sein, so dass diese Schwäche nicht wichtig erscheint. Das bei dieser Art von Störung wichtige Outcome sind Rückfälle nach stationärer Entlassung, diese sind erfasst worden. Zu den Katamnesezeitpunkten gibt es Daten über stationäre Rückfälle, die allein für sich genommen die Wirksamkeit der untersuchten Verfahren hinreichend belegen können.

Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Es handelt sich um stationäre Psychotherapie der Alkoholabhängigkeit. Auch wenn der Schwerpunkt in der Auswertung der Studie ein anderer war, belegen die Daten eindrücklich die durch klientenzentrierte Gruppenpsychotherapie erreichte Reduktion von Rückfällen nach stationärer Entlassung, die den Raten einer Kontrollgruppe signifikant überlegen ist (25 vs. 53%).

Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Die Qualität der Behandlung erscheint vergleichbar. Als Einschränkungen der internen Validität wurden begleitende Interventionen nicht untersucht.

Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Die Stichprobe ist mit Ausnahme der Begrenzung auf das männliche Geschlecht kaum selektiv, es handelt sich im Wesentlichen um die Inanspruchnahmepopulation eines Krankenhauses. Als Einschränkung der externen Validität wurden keine Drop-out und ITT-Analysen durchgeführt (die zum damaligen Stand der Forschung auch noch nicht vorausgesetzt wurden).

Es wurden Standardverfahren der Therapie eingesetzt, das stationäre Kurzzeit-Gruppensetting ist bei Patienten mit Alkoholabhängigkeit auch in Deutschland üblich. Alle Therapeuten praktizierten in der vorliegenden Studie alle Therapieverfahren; dies wurde gemacht, um die Vergleichbarkeit der Qualität der Behandlung zu gewährleisten. In der Versorgungspraxis würden die Therapien von Therapeuten der jeweiligen Therapierichtung durchgeführt werden, was die Wirksamkeit der Therapie vermutlich erhöhen würde, zumindest die Übertragbarkeit der Ergebnisse jedenfalls nicht einschränkt.

Stärken der Studie:

Die Vergleichbarkeit der durchgeführten Therapien, der lange Katamnesezeitraum, das Outcome-Kriterium der Rückfälle nach Entlassung.

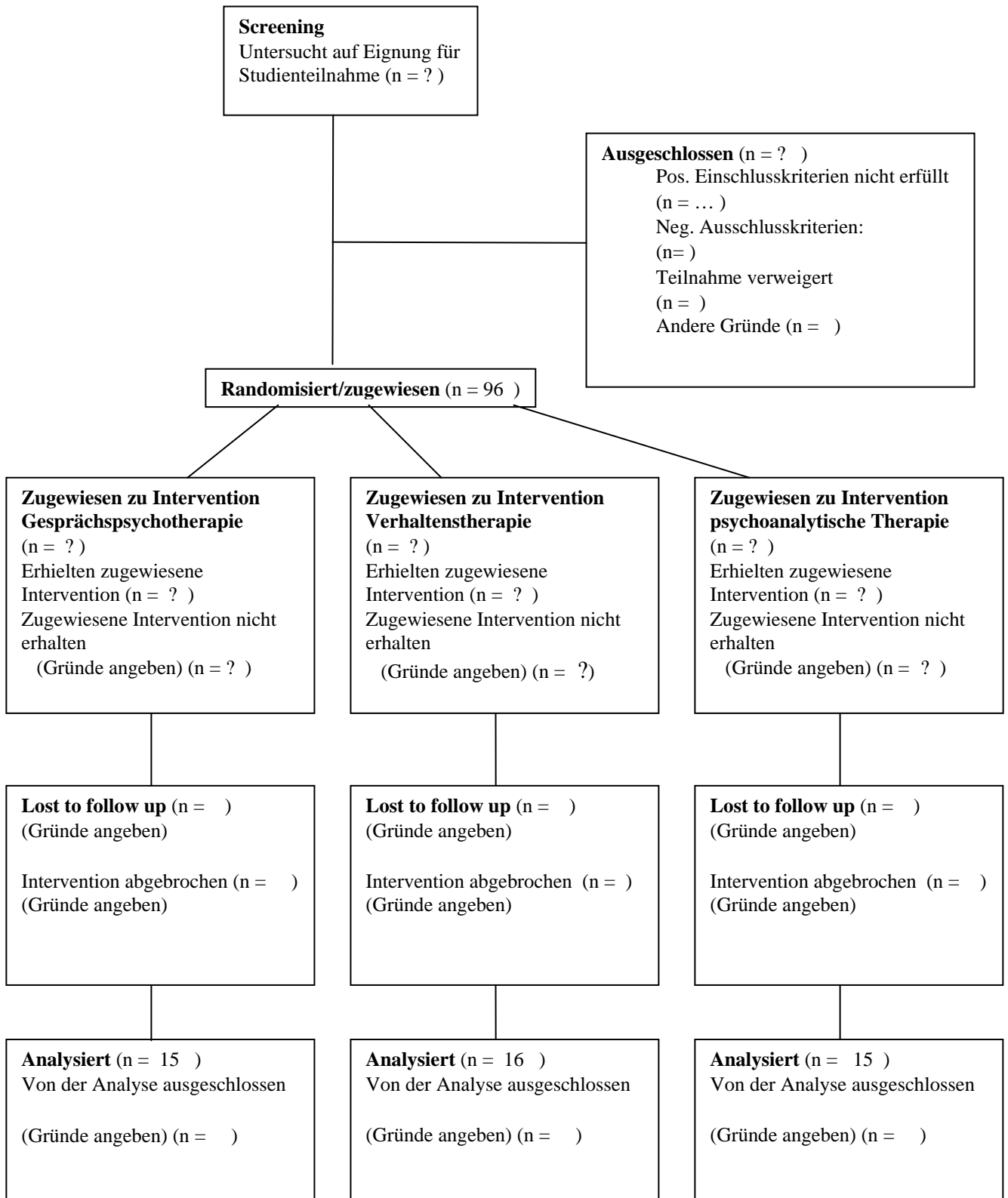
Schwächen der Studie:

Es liegen keine Diagnosen nach gängigen Klassifikationssystemen vor, der Krankheitswert der Störungen kann jedoch angenommen werden, da die Patienten stationär wegen Alkoholkonsums aufgenommen wurden. Für die Vergleiche vor und nach Therapieende liegen keine heute üblichen Messverfahren vor, das bei dieser Art von Störungen wichtigste Outcome-Kriterium der Rückfälle nach Entlassung wurde jedoch erfasst. Es fehlen Angaben zu den Drop-outs.

Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie zeigt, dass stationäre Gesprächspsychotherapie im Gruppensetting einer Kontrollgruppe signifikant überlegen ist hinsichtlich der Reduktion von Rückfällen bei Patienten mit Alkoholabhängigkeit.

8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

K5 (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

K14 (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

K33 (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

K39 (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

Anmerkungen zur durchgeführten Studienbewertungen

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

MT

Studie:

Autoren:

Goldman et al.

Titel:

The effects of adding emotion-focused interventions to the client-centered relationship conditions in the treatment of depression

Quelle:

PsychotherRes (2006), 16(5): 537-549

1. Basisinformationen

1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

- klientenzentrierte Therapie

1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

- Differentieller Vergleich der Wirksamkeit von klientenzentrierter Therapie (CCT) und „emotion-focused therapy“ (EFT) bei Depression
- Annahme der Wirksamkeit beider Therapieformen, aber einer besseren Wirksamkeit der EFT durch die emotionsfokussierten Techniken zusätzlich zu einer klientenzentrierten Beziehung, die beiden Therapieformen gemeinsam ist

1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

- Major Depression (Diagnose SKID)

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnostikstellung (mittels (teil-) standardisierter Interviews)	1) Diagnostikstellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnostikstellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnostikstellung	1 A2
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	1 C1

2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschlusskriterien:

- Major Depression

Ausschlusskriterien:

- aktuelle Behandlung der Depression (auch Medikation)
- komorbide Störung nach DSM-IV (Bipolar I, Panikstörung, Substanzabhängigkeit, Essstörung, Psychotische Störung, mind. 2 schizotypische Merkmale, paranoide, Borderline oder antisoziale Persönlichkeitsstörung)
- wenn, Behandlung mit anderem Fokus für indiziert gehalten wurde (z.B. suizidal oder kürzlicher Suizidversuch) bzw. Krisenintervention (Verlust einer nahen Person im letzten Jahr; kürzlich oder aktuell Opfer von Inzest oder sexuellem Missbrauch; aktueller körperlicher Missbrauch in der Beziehung)

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	1 B1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	1 B2
K5. (C.3.)	Selektivität der Stichprobe aufgrund der Ausschlusskriterien	1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten 2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen) 3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien	3 C3

2.3 () Setting/Patientenzugang

- Informationsverbreitung über lokale Medien (TV, Zeitungen, Radio), Kliniken, Behandlungszentren

	Kriterium	Operationalisierung	Rating
K6. (C.2.)	Art der Rekrutierung der Stichprobe	1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege 2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege 3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung)	3 C

2.4 (7) Zentrale Patientenmerkmale (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- Alter im Mittel ca. 39,5 (SD 9,7) Jahre (22-60 J.)
 - 63% Frauen
 - 34% nie verheiratet, 32% verheiratet oder zusammenlebend, 34% getrennt
 - 45% Highschool-Abschluss, 45% College-Abschluss, 10% postgraduiert

 - BDI 26,4 (SD 7,2): (8% leichte bis mittlere Depression, 60% mittlere bis schwere Depression, 32% sehr schwere Depression)
 - SCID GAF Wert 62,9 (SD 5,4)
 - komorbide Diagnosen: 13% generalisierte Angststörung, 32% Persönlichkeitsstörung (überwiegend ängstlich-vermeidend)

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

unklar

	Kriterium	Operationalisierung	Rating
K7. (A.16.)	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	3 A16

2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten (Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- 458 Personen meldeten sich via Telefon und wurden einer Screening Befragung unterzogen (Erfüllung der Ein- und Ausschlusskriterien)

 - 104 Personen davon nahmen an zwei diagnostischen Sitzungen teil (1. Sitzung SCID Modul Depressive Störung, bei Erfüllung 2. Sitzung andere SCID-Module)

 - 42 Personen randomisiert einer Therapiegruppe zugewiesen (Therapiedauer 4-5 Monate)

 - 4 Dropouts, 2 pro Bedingung (Veränderung der Lebensumstände, Krankheit, Umzug, andere Behandlung, Beendigung der Therapie vor der 8. Sitzung)

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 2
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 2
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A7 2
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	2 A3
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 9

2.7 (10) Vergleichbarkeit der Gruppen

<p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen: # ist gegeben bzgl. erhobener soziodemograph. Variablen und der Ausgangswerte</p> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer): # ist gegeben</p> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen): # ist gegeben</p> <p>Vergleichbarkeit der Therapieadhärenz: # ist gegeben</p>
--

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	<p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p>	B9 1
K14. (A.18.)	Drop-out-Analysen	<p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>	A18 3

3. Intervention

**3.1 (14) Studiendesign /
Behandlungsarme**

- randomisierte klinische Studie mit 2 Therapievergleichsgruppen

**3.2 (11) Behandlung Experimental-
gruppe
(Art, Dauer, Frequenz)**

- Klientenzentrierte Therapie (CCT)
- nach Manual für "Relational CCT" (Greenberg & Goldman, 1999; Rice, Greenberg & Watson, 1994)
- Therapeuten realisieren die Grundhaltungen Wertschätzung, Empathie, Echtheit und ermöglichen so Selbstexploration und Symbolisierung von Emotionen; dies führt zu einer Stärkung des Selbst

**3.3 (12) Behandlung Kontroll-
gruppe
(Art, Dauer, Frequenz)**

- Emotionsfokussierte Therapie (EFT)
- nach Manual von Greenberg et al.(1993)
- Rahmen „Relational CCT“ (s. CCT) + emotionsfokussierte erfahrungsbasierte und Gestalttechniken, um affektiv-kognitive Probleme zu lösen

- ambulant, Einzeltherapie 1x pro Woche, Dauer 4-5 Monate
- 14 Therapeuten unterschiedlichen Erfahrungsgrades (führen beide Therapien durch)

**3.4 (13) Evtl. weitere Kontroll-
gruppen
(Art, Dauer, Frequenz)**

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 1
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	B4
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	B5 1
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	B6 1
K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	C4

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantziellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 1
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	C6 2
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	C7a 2
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	C7b 2
	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	C7c 3

**3.5 (19) Beobachtungszeitraum/
Messzeitpunkte/Katamnese**

- Beobachtungszeitpunkte unmittelbar vor und nach der Therapie und 8 Wochen nach Therapiebeginn (d.h. Beobachtung über 4-5 Monaten)
- keine Katamnese

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 3
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	2 10
K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	3 11

3.6 (15) Anzahl der Zentren

1

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...)

Randomisierung

3.8 (17) ggf. Concealment der Randomisierung

ja

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 3
K27. (B.8.)	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ($n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	B8 2

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

Verringerung der Depressivität (SKID, BDI)
Verringerung der Symptombelastung (SCL-90-R)

4.2 (21) Sekundäre Zielkriterien

Selbstbewusstsein (Rosenberg Self-Esteem Inventory, Bachmann et al., 1977)
Interpersonelle Funktionsprobleme (IIP, Horowitz et al., 1988)

außerdem Maße zur Bewertung der in der Therapie wahrgenommenen Empathie und Therapie-„Adherence“ Bewertung

4.3 () Verblindung der Erhebung der Zielkriterien

nein

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,	A7 2
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend	1A8
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)	1C9
K31. (A.9.)	Multiple Informationsquellen	1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle	A9 2
K32. (A.10.)	Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)	1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler 2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern 3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten	A10 1

4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)

- mittels ANOVA (Gruppe x Therapeut) Analyse eines Therapeuteneffekts

- mittels ANCOVA (Gruppe x Zeitpunkt) mit Prä-Werten als Kovariate Analyse der (differenziellen) Therapiewirksamkeitseffekte

- mittels Chi-Quadrat-Tests Vergleich der klinisch signifikanten Wirksamkeit (Einteilung in drei Gruppen anhand von SCID und BDI Cut-off: Major Depression, Teilremission, Vollremission)

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	1 A6
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung	A13 2
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	A14 3

4.5 (22) Ergebnisse

der spezifische Therapeut hat keinen Einfluss

CCT und EFT führten zu deutlichen Veränderungen in allen Maßen (z.B. BDI von 26 auf 10 bei CCT und auf 5 bei EFT)

die Veränderungen sind klinisch bedeutsam; bei Therapieende erfüllt nur noch ein CCT-Klient und kein EFT-Klient die Kriterien einer Major Depression

bzgl. der klinisch signifikanten Veränderung kein Unterschied zwischen den Gruppen

in den Outcome-Maßen signif. Unterschiede zugunsten der EFT im BDI und der SCL-90-R

	Kriterium	Operationalisierung	Rating
K36. (A.11.)	Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten	1) für alle Outcome-Variablen berichtet 2) ausschließlich für die primären Zielkriterien berichtet 3) nicht für alle primären Zielkriterien berichtet	A11 1
K37. (B.12.)	Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte)	1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe 2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig 3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG	B12 1

4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen

nein

	Kriterium	Operationalisierung	Rating
K38. (A.12.)	Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen	1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen 2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen 3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet	A12 3

4.7 (24) Fazit/Schlussfolgerungen der Autoren

Beide Therapieformen sind sehr wirksam bei der Behandlung einer Major Depression. Es gibt Hinweise für zusätzliche positive Effekte von EFT-Interventionen im Vergleich zur CCT.

	Kriterium	Operationalisierung	Rating
K39. (A.1.)	Manipulation der Daten	1) keine Hinweise auf Ergebnismanipulation 3) Hinweise auf Ergebnismanipulation	A1 1

5. Beurteilung der Ergebnisse der Studie		
D1	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual". (Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt) (ja = 1; nein = 3; nicht zutreffend = freilassen)	D1
D2	In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D2
D3	In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde. (ja = 1; nein = 3; nicht zutreffend = freilassen)	D3

6. Praxistransfer

	Kriterium	Operationalisierung	Rating
K40. (C.10.)	Spezifikation und Herstellbarkeit notwendiger Settingbedingungen	1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team) 2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team) 3) Notwendige Settingbedingungen nicht herstellbar	C10
K41. (C.11.)	Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation	1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar 2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar 3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar	C1
K42. (C.12.)	Spezifikation und Erfassbarkeit relevanter Patientenmerkmale	1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar 2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar 3) Relevante Patientenmerkmale praktisch nicht erfassbar	C12 1
K43. (C.13.)	Spezifikation und Herstellbarkeit relevanter Treatmentaspekte	1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual) 2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung) 3) Relevante Treatmentmerkmale praktisch nicht herstellbar	C13 1

7. Gesamtbewertung der Studie

Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)			
Durchschnittswert für B) Interne Validität (Summe / 12)			
Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)			
Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4)			

Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

<p>Kriterien für 'A. Allgemeine methodische Qualität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> 2. wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde <p>und</p> <ol style="list-style-type: none"> 3. wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist. 	
<p>Kriterien für 'B. Interne Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> 2. der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist. 	
<p>Kriterien für 'C. Externe Validität' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> 1. das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden <p>und</p> <ol style="list-style-type: none"> 2. das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden <p>und</p> <ol style="list-style-type: none"> 3. bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde <p>und</p> <ol style="list-style-type: none"> 4. bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde. 	
<p>Die Ergebnisse der Studie belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>	

Abschließende Bewertung

Allgemeine methodische Qualität der Studie

Die allgemeine methodische Qualität der Studie ist mit wenigen Einschränkungen gut.

Diagnosestellung und Veränderungsmessung erfolgt mittels reliabler und valider Verfahren. Die Qualität der Durchführung der beiden untersuchten Verfahren wird mittels geeigneter Verfahren überprüft.

Die Stichprobe und daraus resultierende Power sind für differentielle Vergleiche etwas zu gering. Es gibt keine Katamnese-Erhebung.

Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Beide Therapien führen gleichermaßen zu klinisch bedeutsamen Verbesserungen der Symptomatik. In einzelnen Outcome-Maßen schneidet die EFT etwas besser ab (BDI, SCL-90-R)

Anzumerken ist, dass die Studie zusätzlich integrativ mit der Kohorte von Greenberg & Watson, 1998 analysiert wurde. Diese Ergebnisse wurden bei der Bewertung nicht berücksichtigt.

Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Die interne Validität ist insgesamt gut. Die Effekte zur Post-Erhebung sind deutlich und können nicht allein auf Spontanremission zurückgeführt werden. Hervorzuheben ist, dass jeder Therapeut beide Verfahren durchführt; somit werden differentielle Effekte der Person des Therapeuten vermieden.

Externe Validität

(Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)

Die externe Validität hinsichtlich der Patienten ist eingeschränkt, da die Patienten unter Annahme vieler Ausschlusskriterien über Anzeigenwerbung akquiriert wurden, was Selektionseffekte nicht ausschließen lässt. Die externe Validität hinsichtlich Dauer und Frequenz der Therapie ist gut. Einschränkungen ergeben sich aus einem umfangreichen Training und Supervision der Therapeuten. Es handelt sich um 14 unterschiedlich erfahrene Therapeuten, was zu einer guten Gesamtrepräsentativität führt.

Stärken der Studie:

- Ausschaltung von event. Therapeuteneffekten, da jeder Therapeut beide Therapien durchführt
- randomisierte Zuordnung

Schwächen der Studie:

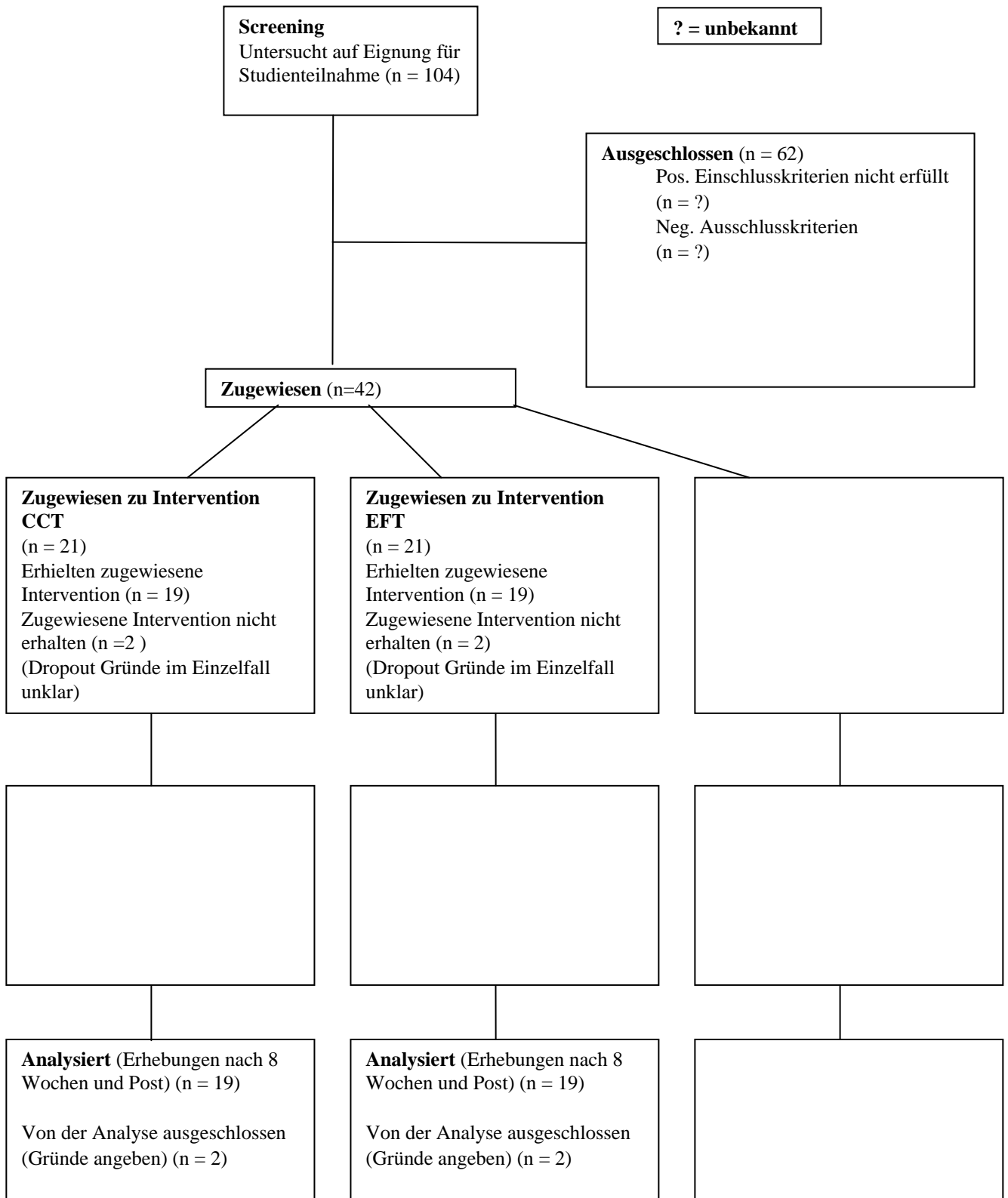
- N und resultierenden Power sind zu gering für differentielle Fragestellungen.
- keine Katamnese
- viele Ausschlusskriterien d.h. Selektivität der Stichprobe anzunehmen
- umfangreiche Trainingsmaßnahmen und Supervision der Therapeuten

Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie wird als geeignet für den Nachweis der Wirksamkeit der GT bei depressiven affektiven Störungen angesehen.

Obwohl der Vergleich der beiden Methoden einem Power-Problem unterliegt, sind die Prä-Post Effekte in beiden Gruppen überzeugend.

8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

K5 (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

K14 (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

K33 (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

K39 (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

Anmerkungen zur durchgeführten Studienbewertungen

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,
bitte in der Regel „9“ eintragen.*

Projekt: Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

Beurteiler:

MT

Studie:

Autoren:

Grawe/ Plog

Titel:

Differentielle Psychotherapie I + II
I - Indikation und spezifische Wirkung von Verhaltenstherapie und Gesprächspsychotherapie. Eine Untersuchung an phobischen Patienten.
II - Der Zusammenhang von Lebensbedingungen und spezifischen Therapieeffekten im Vergleich von GT und VT

Quelle:

Arbeiten zur Theorie und Praxis der Rehabilitation in der Medizin, Psychologie und Sonderpädagogik 11 & 12, 1976, Verlag Hans Huber: Bern, Stuttgart, Wien

1. Basisinformationen

1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

- Gesprächspsychotherapie

1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

- Effekte von GT und VT bei psychiatrischen Patienten mit schweren Phobien?: Für beide Therapieformen wird eine stärkere Reduktion der phobischen Symptomatik erwartet als in der Kontrollbedingung.

- Differentieller Vergleich der Wirksamkeit von GT und VT? Worin besteht g.f. der Unterschied in der Wirksamkeit?: Größere Verbesserung der Symptomatik für VT erwartet. Für GT stärkere Annäherung von Selbst- und Idealbild erwartet als für VT.

- Bei wem bewirken die Therapiemethoden Effekte (Vorhersagekraft der Ausgangswerte)?

- Wodurch könnten Effekte, die bei GT und bei VT auftreten, bedingt sein?

Plog:

- differentielle Langzeiteffekte (Zufriedenheit mit den Lebensbedingungen)

1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

2. Population

2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

- Phobie (als primäre Störung), am häufigsten Agoraphobie und/ oder soziale Phobie, wenige Patienten mit einer anderer komplexer phobischer Störung

	Kriterium	Operationalisierung	Rating
K1. (A.2.)	Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews)	1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS) 2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil 3) keine adäquate Diagnosestellung	2 A2
K2. (C.1.)	Stichprobe von Patienten mit Störungen mit Krankheitswert	1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen) 2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch 3) Patienten ohne identifizierbare Störung mit Krankheitswert	1 C1

2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschlusskriterien:

- psychiatrische Patienten mit komplexer phobischer Störung wie Agoraphobie oder Sozialer Phobie (z.B. nicht ausschließlich Tierphobie)
- phobische Symptome stehen im Vordergrund des Beschwerdebildes
- medikamentöse Behandlung allein, nach Eindruck des Nervenarztes nicht hilfreich

Ausschlusskriterien

- phobische Symptome sind Begleiterscheinung einer anderen psychischen Störung
- Alter < 20 oder > 40 Jahre
- psychotherapeutische Vorbehandlung (scheinbar weich gehandhabt)

	Kriterium	Operationalisierung	Rating
K3. (B.1.)	Spezifizierung der Einschluss- und Ausschlusskriterien	1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien 2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben 3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert	1 B1
K4. (B.2.)	Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden	1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.) 2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus. 3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus.	2 B2
K5. (C.3.)	Selektivität der Stichprobe aufgrund der Ausschlusskriterien	1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten 2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen) 3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien	2 C3

2.3 () Setting/Patientenzugang

- die Patienten wurden durch niedergelassene Nervenärzte überwiesen (den koop. Nervenärzten waren Ein- und Ausschlusskriterien bekannt)

	Kriterium	Operationalisierung	Rating
K6. (C.2.)	Art der Rekrutierung der Stichprobe	1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege 2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege 3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung)	2 C

2.4 (7) Zentrale Patientenmerkmale
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- Alter im Mittel ca. 29,8 Jahre (20-49 J.)
 - 75% Frauen
 - 37 verheiratet, 14 ledig, 2 geschieden
 - 31 Volkshochschule, 18 Mittelstufe, 2 höhere Schule, 2 Sonderschule
 - im Mittel bereits seit 8,6 Jahren phobische Symptome

2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

54 (3 x 18), durch Patientenregistrierungsfehler nicht ganz realisiert

	Kriterium	Operationalisierung	Rating
K7. (A.16.)	Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe	1) adäquate statistische Power der Vergleiche 2) Eingeschränkte Power der statistischen Vergleiche (.50-.80) 3) unzureichende statistische Power der Vergleiche (<.50)	3 A16

2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- 57 Patienten gescreent, 53 Patienten eingeschlossen
 - überwiegend randomisierte Zuweisung (siehe S. 51) zu GT-Gruppe, VT-Gruppe, Wartegruppe
 - 1 Behandlungsabbruch in der GT-Gruppe, 1 Patient verweigert Nachuntersuchung in der GT-Gruppe

	Kriterium	Operationalisierung	Rating
K8. (A.5.)	Stichprobengröße pro Gruppe	1) n pro Gruppe > 30 2) n pro Gruppe 10-30 3) n pro Gruppe < 10 (Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)	A5 2
K9. (A.15.)	Intention to treat – Analysen durchgeführt	1) ITT-Analysen durchgeführt 2) Keine ITT-Analysen bei geringem Risiko für attrition bias 3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias	A15 2
K10. (A.17.)	Vollständige Beschreibung der Drop-Outs	1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out 2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out 3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe	A7 1
K11. (A.3.)	Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	1 A3
K12. (A.4.)	Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt)	1) deutlich besser als in Studien mit vergleichbaren Patientengruppen 2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen 3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen	A4 1

2.7 (10) Vergleichbarkeit der Gruppen

<p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen: # ist gegeben bzgl. erhobener soziodemograph. Variablen, ist gegeben bzgl. der meisten Ausgangswerten</p> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer): # ist gegeben</p> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen): # nicht bekannt</p> <p>Vergleichbarkeit der Therapieadhärenz: # in GT-Gruppe scheiden 2 Patienten aus (1 prä-post, 1 Follow-up), in der VT-Gruppe keiner</p>

	Kriterium	Operationalisierung	Rating
K13. (B.9.)	Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale	<p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p>	B9 2
K14. (A.18.)	Drop-out-Analysen	<p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>	A18 3

3. Intervention

**3.1 (14) Studiendesign /
Behandlungsarme**

- randomisierte klinische Studie mit 2 Therapie- und einer Wartegruppe

**3.2 (11) Behandlung Experimental-
gruppe
(Art, Dauer, Frequenz)**

- Verhaltenstherapie (VT)
- „broad spectrum behavior therapy“ (Lazarus, 1966)
- Verhaltensanalyse und darauf aufbauende Therapieplanung eng an Schema zur funktionalen Verhaltensanalyse von Kanfer und Saslow (1969) angelehnt
- die spezif. anzuwendenden Techniken wurden für jeden Fall durch die ganze Therapeutengruppe während der Supervisionssitzung auf Grundlage der Verhaltensanalyse bestimmt (systemat. Desensibilisierung, Flooding, operante Verfahren)

- ambulant, Einzeltherapie 1-2x/pro Woche, Dauer max. 6 Monate
- 7 Therapeuten (fortgeschrittene eng supervidierte Studenten)

**3.3 (12) Behandlung Kontroll-
gruppe
(Art, Dauer, Frequenz)**

- Gesprächspsychotherapie (GT)
- Realisierung der drei notwendigen und hinreichenden therapeutischen Grundhaltungen (Tausch, 1970)

- ambulant, Einzeltherapie 1-2x/pro Woche, Dauer max. 6 Monate
- 7 Therapeuten (weit fortgeschrittene eng supervidierte Studenten)

**3.4 (13) Evtl. weitere Kontroll-
gruppen
(Art, Dauer, Frequenz)**

Wartegruppe (setzt langjährige medikamentöse Behandlung bei niedergelassenem Neurologen fort)

	Kriterium	Operationalisierung	Rating
K15. (B.3.)	Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)	1) ausführliches Therapiemanual 2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie 3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens	B3 2
K16. (B.4.)	Operationale Definition der Kontrollbedingungen	1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung 2) Ex post facto Beschreibung der Kontrollbedingungen 3) keine Beschreibung der Kontrollbedingung	B4
K17. (B.5.)	Strukturelle Äquivalenz bei Kontrollbedingungen	1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz 2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab 3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab	B5 1
K18. (B.6.)	Manualtreue, Treatment Integrity	1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt 2) Manualtreue/Treatmentintegrität durch Fragebögen belegt 3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen	B6 1
K19. (C.4.)	Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer	1) Intervention wie in klinischer Alltagspraxis 2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert 3) Intervention gegenüber klinischer Alltagspraxis stark verändert	C4

	Kriterium	Operationalisierung	Rating
K20. (B.7.)	Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie)	1) Ausschluss begleitender nicht-randomisierter Interventionen 2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantziellen, differenziellen Einfluss der begleitenden Interventionen hin 3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin	B7 9
K21. (C.6.)	Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)	1) keine Einschränkungen 2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen 3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen	C6
K22. (C.7.)	Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode) a) Klinische Tätigkeit der Therapeuten	1) Therapeuten sind praktizierende Kliniker 2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten) 3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)	C7a 2
	b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)	1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie 2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie 3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).	C7b 2
	c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie	1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an) 2) kurzes Training für die Studie / intensives Training einiger Therapeuten 3) intensives Training vor der Studie	C7c 2

**3.5 (19) Beobachtungszeitraum/
Messzeitpunkte/Katamnese**

- Akquisebeginn 1970, Dauer Akquisezeitraum ca. 2 Jahre
 - Beobachtungszeitpunkte unmittelbar vor und nach der Therapie, Untersuchung der beiden Therapiegruppen auch 4 Monate nach Therapieende, keine Nachuntersuchung der Wartegruppe
 - 2 bis 2,5 Jahre nach Therapieende Nachbefragung zum Thema Zufriedenheit mit den Lebensumständen

	Kriterium	Operationalisierung	Rating
K23. (C.5.)	Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)	1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt) 2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten) 3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen)	C5 3
K24. (B.10.)	Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung)	1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up 2) ausschließlich prä-post-Messung 3) ausschließlich Post-Messung	B10
K25. (B.11.)	Follow-up-Messung	1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate) 2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe 3) keine Katamnese	B11 2

3.6 (15) Anzahl der Zentren

1

3.7 (16) Allokation zu Interventionsgruppen (Randomisierung, self-selection, ...)

Randomisierung (mit vereinzelt Abweichungen am Anfang, bei denen nur über die beiden Therapiegruppen randomisiert wurde, um die Therapeuten optimal auszulasten)

3.8 (17) ggf. Concealment der Randomisierung

Zuordnung dem Diagnostiker zum Zeitpunkt prä nicht bekannt

	Kriterium	Operationalisierung	Rating
K26. (C.8.)	Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention	1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen 2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen 3) alle Patienten werden der Therapie (zufällig) zugewiesen	C8 3
K27. (B.8.)	Gruppenzuweisung	1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt 2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe 3) keine randomisierte oder parallelisierte Zuweisung	B8 2

4. Ergebnisse

4.1 (20) Primäre Zielkriterien

Symptomverringern (Selbst- und Fremdeinschätzung (Angehöriger, Psychiater) auf Skalen von Gelder und Marks, 1966; Intensität und Häufigkeit klinischer Symptome, u.a. Angst und Depressivität, subjektive Bedeutsamkeit bestimmter phobischer Situationen)

4.2 (21) Sekundäre Zielkriterien

Persönlichkeitsveränderungen/ Selbstkonzeptänderungen (FPI; 16-PF von Cattell; Selbst- vs. Idealbild (Q-Sort nach Mannesmann, 1972)

Verbesserung der Eigenbefindlichkeit (Eigenschaftswörterliste von Janke, 1961; Therapieveränderungsfragebogen von Mehnert, 1969)

Reduktion des Medikamentenverbrauchs (4-stufige Skala)

Verringerung von Störungen des Arbeits-, Freizeit-, Beziehungs- und Sexualverhaltens (Teil der Skalen von Gelder und Marks, 1966)

Zufriedenheit mit den Lebensbedingungen

4.3 () Verblindung der Erhebung der Zielkriterien

z.T.

	Kriterium	Operationalisierung	Rating
K28. (A.7.)	a priori Definition der primären und sekundären Zielkriterien	<ol style="list-style-type: none"> 1) a priori definiert unter Berücksichtigung der Wertigkeit 2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien 3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien, 	A7 2
K29. (A.8.)	Reliable und valide Messung zumindest der primären Zielkriterien	<ol style="list-style-type: none"> 1) reliable und valide Outcome-Verfahren 2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren 3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend 	2A8
K30. (C.9.)	Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität)	<ol style="list-style-type: none"> 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität 2) Zielkriterien beziehen sich nur auf eine Dimension 3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung) 	1C9
K31. (A.9.)	Multiple Informationsquellen	<ol style="list-style-type: none"> 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen 2) zwei Informationsquellen 3) eine Informationsquelle 	1A9
K32. (A.10.)	Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)	<ol style="list-style-type: none"> 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler 2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern 3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten 	A10 1

4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)

- Vergleich von prä-post Veränderungswerten in Form 1. einfacher Differenzwerte, 2. Regressionsabweichungswerte (Argumentation gegen einen reinen Vergleich von Mittelwerten nach Bergin, 1971)

- Vergleich der Werte innerhalb eines Zeitpunktes zwischen den Gruppen und der genannten Veränderungswerte jeweils mit T-Test und U-Test, Vergleich der Varianzen mit Fmax-Test, Bildung der möglichen Korrelationen

	Kriterium	Operationalisierung	Rating
K33. (A.6.)	Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert	1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt 2) teilweise post-hoc Definition der Vergleiche 3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen	1 A6
K34. (A.13.)	Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt	1) Anwendungsvoraussetzungen geprüft und erfüllt 2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren 3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung	A13 1
K35. (A.14.)	Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests)	1) adäquate und umfassende statistische Analysen 2) weitgehend adäquate statistische Analysen 3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)	A14 3

4.5 (22) Ergebnisse

Vergleich Prä-Post:
 ~~~~~Vergleich der Mittelwerte gegen Wartegruppe und differentiell

~~~Symptomverbesserung  
 # VT und GT führen jeweils zu hochsignifikanter Verbesserung der phobischen Symptomatik im Vergleich zu der Wartegruppe (in Selbst- und Fremdeinschätzungsmaßen)
 # VT und GT unterscheiden sich nicht in ihrer Wirksamkeit

~~~Persönlichkeitsänderung  
 # VT und GT führen zu einer signifikant größeren Übereinstimmung von Selbst- und Idealbild  
 # im FPI führt nur VT zu Verbesserungen (weniger nervös, depressiv, emotional labil)

~~~Eigenbefindlichkeit  
 # verbessert sich bei VT und GT hochsignifikant und im gleichen Maße

~~~Medikamentenverbrauch  
 # nimmt nur bei VT signifikant ab, auch im Vergleich zur GT signifikant

~~~Freizeitverhalten  
 # scheint sich in beiden Gruppen leicht zu bessern

~~~~~Vergleich der Veränderungsvarianz gegen Wartegruppe und differentiell

~~~Symptomverbesserung  
 # VT und GT führen jeweils zu signifikanter Verbesserung der phobischen Symptomatik im Vergleich zu der Wartegruppe (in Selbst- und Fremdeinschätzungsmaßen)
 # VT besser als GT, aber nicht in allen Maßen signifikant (d.h. GT führt bei nicht so vielen Patienten zu starken Effekten, wie VT bei einigen, aber auch nicht zu sehr geringen Effekten, wie VT bei einigen anderen)

~~~Persönlichkeitsänderung  
 # bei VT erhöhte Veränderungsvarianzen auf 8 Skalen des 16-PF, bei

#### GT auf 4 Skalen

# Interkorrelation der Veränderungswerte ergibt, dass VT und GT unterschiedliche Veränderungsmuster (Verbesserungsmuster) erzeugen; bei VT hängen die erzielten Veränderungswerte enger miteinander zusammen als bei GT und Wartegruppe (d.h. z.B. bei VT stärkster Zusammenhang zwischen Besserung der Symptomatik und des Befindens)

#### ~~~~~Vergleich Post vs. 4-Monats-Katamnese

# positive Veränderungen bleiben bei beiden Therapien im wesentlichen bestehen  
# wieder leichter Anstieg von phobischem Vermeidungsverhalten, aber nicht auf das Level von vor der Therapie  
# bei VT-Patienten wieder ein leichter Anstieg von FPI-Erregbarkeit  
# bei GT-Patienten auf einigen FPI-Skalen Verbesserungen  
# subjektives Befinden wird bei VT und bei GT eher noch besser

#### ~~~~~Differentielle Indikation

# VT führt zu besonders deutlichen Verbesserungen bei Patienten mit starken Symptomen/ hohem phobischen Leidensdruck  
# für GT Wirksamkeit findet sich kein spezifischer Indikator, sie wirkt also auch bei schwächerem phobischen Leidensdruck

#### ~~~~~Zufriedenheit mit den Lebensbedingungen

# vor Therapie selbes Maß an (Un)zufriedenheit  
# GT Patienten sehen in deutlich stärkerem Maß Zusammenhänge Beeinträchtigungen und Zufriedenheit mit Lebensbedingungen als VT Patienten  
(wird nicht detaillierter ausgeführt)

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |



**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**



|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br>9 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

VT und GT sind wirksam bei Phobien. VT wirkt besonders gut bei Patienten mit starkem phobischen Leidensdruck, der auch sehr auf die Phobie konzentriert ist. Die Wirksamkeit der GT ist von der Stärke des Leidensdruckes unabhängig. Spezifisch bei GT sind die Erfolgskriterien relativ unabhängig (d.h. z.B. Befindlichkeit kann deutlich gebessert sein, obwohl Symptomatik nicht).

|                       | Kriterium              | Operationalisierung                                                                    | Rating |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1     |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |         |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1<br>1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2      |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3      |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11      |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>1 |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Die allgemeine methodische Qualität der Studie ist mit Einschränkungen gut.

Problematisch ist, dass die Diagnosestellung basierend auf dem klinischen Urteil erfolgt. Die zur Veränderungsmessung eingesetzten Verfahren sind z.T. veraltet. Die Stichprobe und daraus resultierende Power sind für differentielle Vergleiche zu gering. Z.T. wäre eine Korrektur für multiples Testen nötig gewesen und nicht durchgeführt; die Hauptergebnisse würden allerdings bei einer Korrektur bestehen bleiben.

Als positiv ist zu bewerten, dass die Wartegruppe in der schon länger erfolglosen medikamentösen Behandlung blieb, da das völlige Absetzen der Medikation sicherlich nicht eingehalten worden wäre.

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Verglichen mit der Wartegruppe führen sowohl VT als auch GT zu einer signif. Verbesserung der phobischen Symptomatik, die zur Katamnese anhält, weiterhin zu positiven Persönlichkeitsveränderungen und Verbesserungen der Eigenbefindlichkeit. Bei den gängigen Mittelwertvergleichen unterscheiden sich die Verfahren nicht in ihrer Wirksamkeit. Bei Betrachtung der Veränderungsvarianz schneidet die VT besser ab als die GT. Der Medikamentenverbrauch verringert sich nur bei VT signifikant (Mittelwertvergleich). Die Ergebnisse bleiben zur Katamnese im Wesentlichen stabil.

Bei VT und GT ergeben sich unterschiedliche Veränderungsmuster; bei VT hängen die erzielten Veränderungswerte enger miteinander zusammen als bei GT und Wartegruppe.

Hinsichtlich einer differentiellen Indikation ergibt sich, dass die VT bei Patienten mit starken Symptomen/ hohem phobischen Leidensdruck zu besonders deutlichen Verbesserungen führt. Für GT Wirksamkeit findet sich kein spezifischer Indikator, sie wirkt also auch bei schwächerem phobischen Leidensdruck.

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Die interne Validität ist insgesamt gut.

Problematisch ist, dass die Supervision für beide Gruppe unterschiedlich abläuft (Einzel für GT, Gruppe und Trainingstherapien vor der Studie für VT). In einigen wenigen der zahlreichen Maße unterscheiden sich die Gruppen zum Zeitpunkt Prä, was aber nicht so problematisch ist, da mit Veränderungswerten gerechnet wird.

### Externe Validität

(Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem)

Die externe Validität hinsichtlich der Patienten ist gut, der Intervention, der Therapiefrequenz und -länge ist gut.

Die Behandler (insb. VT) sind recht unerfahren (fortgeschrittene Studenten). Dem Ungleichgewicht versucht man durch eine stärkere Supervision/ Training bei VT entgegenzuwirken.

### Stärken der Studie:

- Zwei Therapiegruppen + Wartegruppe. Beide Therapieformen wurden als „best-practise“ realisiert.
- randomisierte Zuordnung

### Schwächen der Studie:

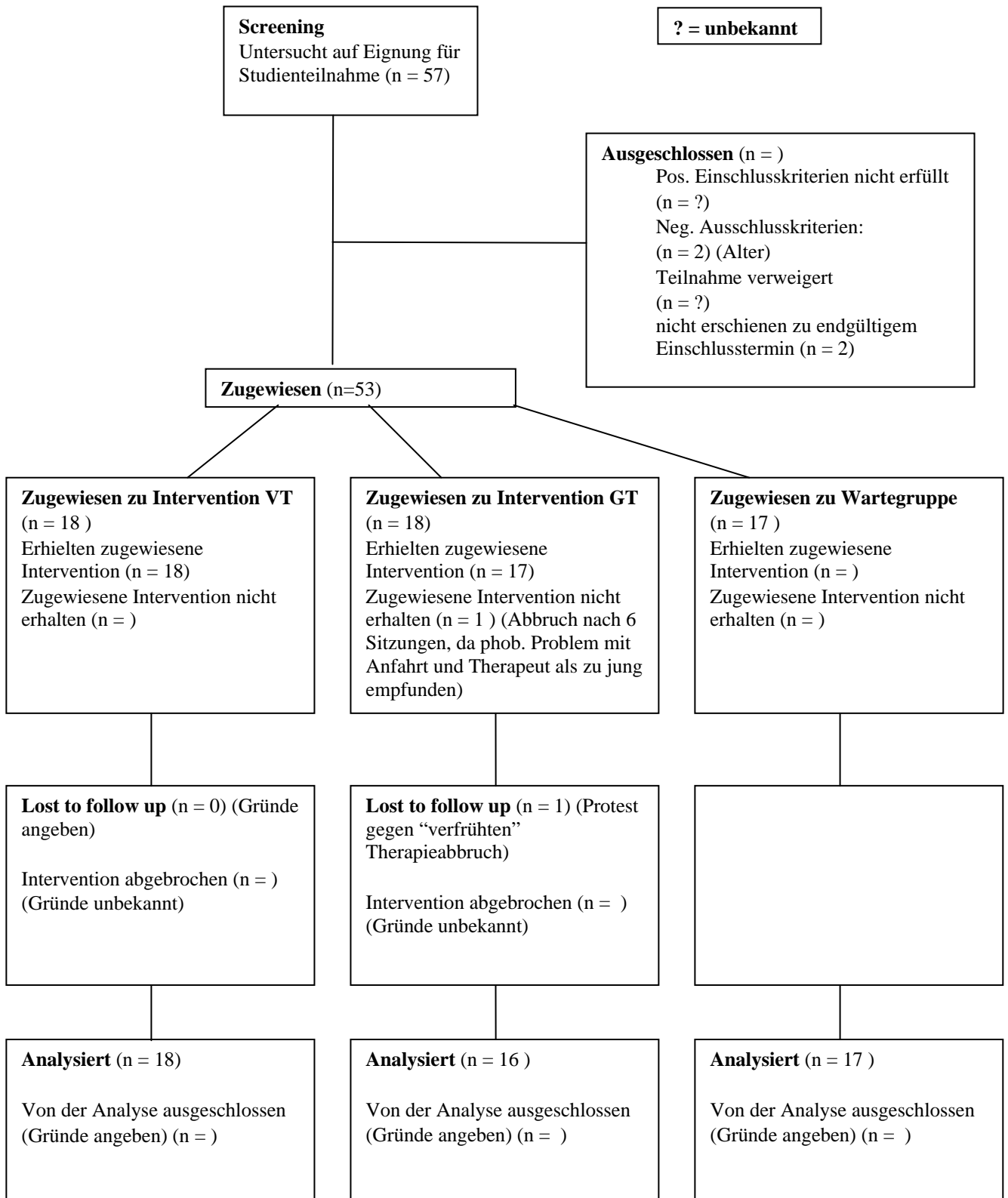
- N und resultierenden Power sind zu gering für differentielle Fragestellungen. Die externe Validität der Behandler ist gering.

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie wird als geeignet für den Nachweis der Wirksamkeit der GT bei Angststörungen angesehen.

Obwohl der Vergleich mit der VT einem Power-Problem unterliegt, und d.h. nur als deutlicher Hinweis interpretiert werden sollte, wird ein Wirksamkeitsnachweis durch den Vergleich mit der Wartegruppe erbracht. Diese Studie ist eine von wenigen zum Thema Angststörungen, bei der GT nicht nur eine unspezif. Vergleichsbedingung darstellt, sondern die Realisierung einer guten GT angestrebt wird.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



Nachuntersuchung (Studie Plog) nach 2 - 2,5 Jahren: VT (n = 16), GT (n = 14)



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.



## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

MS

**Studie:**

Autoren:

King et al.

Titel:

Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy and usual general practitioner care in the management of depression as well as mixed anxiety and depression in primary care

Quelle:

Health Technology Assessment 2000, 4: No. 19

## 1. Basisinformationen

### 1.1 ( ) Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Nicht-direktive Beratung (Kurzzeitpsychotherapie)

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Vergleichende Evaluierung der klinischen und Kosteneffektivität von allgemeinärztlich üblicher Behandlung sowie 2 Kurzzeitpsychotherapieverfahren (nicht-direktive Beratung und kognitive Verhaltenstherapie) bei Depressionen und bei Depressionen mit Angst gemischt.

Untersuchung des Einflusses der Präferenz für ein bestimmtes Behandlungsverfahren auf das Outcome.

### 1.3 (2) Design

- x Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- x Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Depression und Depression mit Angst gemischt (Anwendungsbereich 1.1: Affektive Störungen)

Allgemeinärzte wählten Patienten aus, für die eine psychotherapeutische Intervention indiziert war. Weiteres Kriterium war ein Wert im Beck Depressions Inventar > 14 (der tatsächliche Mittelwert war 26.5 (SD 8.6).

Computerisierte Version des Clinical Interview Schedule-Revised zur Überprüfung der Diagnosekriterien nach ICD-10 (nur für 93.8% der Stichprobe, 19.5% erfüllen nicht die Kriterien einer psychischen Störung)

|               | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|---------------|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K1.<br>(A.2.) | Objektive und reliable Diagnostikstellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>1 |
| K2.<br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                               | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>2 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschlusskriterien:

- Mindestalter von 18 Jahren
- BDI > 14

Ausschlusskriterien:

- ernsthafte Suizidabsichten (erhoben mit spezifischem Protokoll)
- vorhergehende Psychotherapie in den letzten 6 Monaten
- Einnahme von Antidepressiva
- Eingeschränkte Mobilität
- Hirnorganische Störungen
- Unfähigkeit, Fragebögen zu beantworten

|               | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|---------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K3.<br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>1 |
| K4.<br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>1 |

|                      |                                                              |                                                                                                                                                                                                                                                                                                         |             |
|----------------------|--------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien | C3<br><br>1 |
|----------------------|--------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|

### 2.3 () Setting/Patientenzugang

Überweisung durch den Allgemeinarzt (insgesamt 73 beteiligte Ärzte).

|                      | Kriterium                           | Operationalisierung                                                                                                                                                                                                                                                                                                                                            | Rating      |
|----------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K6.</b><br>(C.2.) | Art der Rekrutierung der Stichprobe | 1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege<br>2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege<br>3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung) | C2<br><br>1 |

### 2.4 (7) Zentrale Patientenmerkmale (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

Alter 36.8 Jahre (SD 12.2, Range 18-79)  
25.2% männlich  
62.1% gehören der sozialen Schicht II oder III an (Mittelschicht)  
89.9% europäischer Herkunft  
64.9% erwerbstätig  
46.3% verheiratet/zusammenlebend  
32.1% mit Abiturvergleichbarem Bildungsabschluss

### 2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Poweranalyse a priori durchgeführt  
Pro Subgruppe sollten 65 Patienten untersucht werden, um eine Power von .90 für die gewünschten Effekte zu erreichen.

|                       | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating       |
|-----------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K7.</b><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br><br>1 |

### 2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten (Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

627 Patienten gescreent  
163 wurden ausgeschlossen, da sie  
- < 14 Punkte im BDI hatten (38.0%)  
- Antidepressiva nahmen (16.0%)  
- nicht kontaktiert werden konnten (13.5%)  
- doch nicht teilnehmen wollten (13.5%)  
- vorhergehende Psychotherapien hatten (11.7%)  
- die überweisenden Praxis aus praktischen Gründen nicht konnte (2.5%)  
- ungeeignet (ohne weitere Angabe) (2.5%)  
- Suizidalität (1.2%)  
- Substanzmissbrauch (0.6%), war vorher nicht als Ausschlusskriterium definiert  
- mangelnde Sprachkenntnisse ((0.6%)

464 wurden Behandlung zugewiesen: davon 197 (42.5%) randomisiert zu 1 von 3 Behandlungen, 130 (28.0%) randomisiert zu einer der Psychotherapien und 137 (29.5%) wählten selbst ihre Behandlung.

Zur 4-Monats erheben liegen von 88.9% der Patienten Daten vor, bei

12 Monaten 80.5%

Zum Flow-chart siehe Originalpublikation

Gruppengrößen bei den vollständig randomisierten Patienten:  
 CBT n = 63 (kognitive Verhaltenstherapie)  
 NDC n = 67 (nicht-direktive Beratung)  
 GP n = 67 (Allgemeinärztliche Versorgung)

Patienten mit nur 1 Therapiesitzung wurden nicht als drop-outs gezählt. Alle Analysen wurden sorgfältig jeweils mit ITT- und PP-Stichproben durchgeführt.

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>1  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>1 |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br>2 |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br>2  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>2  |

**2.7 (10) Vergleichbarkeit der Gruppen**

**Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen**

Keine relevanten Gruppenunterschiede in der randomisierten Patientengruppe zwischen den verschiedenen Behandlungsformen bei Baseline.

Bei Baseline unterschieden sich die Gruppen mit Präferenz vs. Randomisierungswillige geringfügig (Patienten mit Behandlungspräferenz waren seltener verheiratet, hatten eine höhere Gesamtsymptombelastung im CIS-R (geringe Effektstärke des Unterschieds < .30), ein niedrigeres soziales Funktionsniveau (Effektstärke <.20), niedrigere Lebensqualität (Effektstärke <.20) sowie eine höhere Anzahl selbstgenannter Probleme (Effektstärke <.30) als Patienten, die randomisiert zugewiesen wurden.

**Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)**

Die durchschnittliche Anzahl der Sitzungen war 6.3 (SD 3.5) für CBT und geringfügig höher mit 7.4 (4.1) für NDC. Ansonsten sind die Gruppen vergleichbar (hinsichtlich Qualifikation und Erfahrung der Behandler).

**Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)**

Die Gruppen unterscheiden sich in der Begleitmedikation:  
 Antidepressiva: GP 49.3%, CBT 27.0%, NDC 29.9%  
 Tranquilizer: GP 17.9%, CBT 6.3%, NDC 14.9%

**Vergleichbarkeit der Therapieadhärenz**

Patientenpräferenz bzw. Randomisierung hatte keinen Einfluss auf die Follow-up Raten.

Bei den voll-randomisierten: 14.3% bei CBT vs. 10.4% bei NDC nahmen die Therapie nicht wahr, weitere 4.8% bei CBT vs. 4.5% bei NDC hatten nur 1 Sitzung.

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating       |
|-----------------|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | 1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen<br>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert<br>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt | B9<br><br>1  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | 1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen<br>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt<br>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage                                                                                                                       | A18<br><br>9 |

### 3. Intervention

|                                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>3.1 (14) Studiendesign /<br/>Behandlungsarme</b>                            | Es wurden 3 Behandlungsgruppen miteinander verglichen, 2 Psychotherapien und eine Kontrollgruppe mit allgemeinärztlichem Treatment as usual                                                                                                                                                                                                                                                                                                              |
| <b>3.2 (11) Behandlung Experimental-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b> | Nichtdirektive Beratung nach Rogers (Kurzzeitpsychotherapie) mit 6 bis max. 12 Therapiestunden (i.d.R. 1x wöchentlich).<br>Alle 13 Therapeutinnen waren spezifisch professionell ausgebildet und anerkannt (zusätzlich gab es ein Manual für die Studie).<br>Supervision von 1 Stunde pro 6 Stunden Patientenkontakt wurde vorausgesetzt. 1 Patient wurde von einem Berater behandelt, der nicht dem Projekt angehörte, aber die Anforderungen erfüllte. |
| <b>3.3 (12) Behandlung Kontroll-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b>     | Kognitive Verhaltenstherapie (Kurzzeitpsychotherapie) mit 6 bis max. 12 Therapiestunden (i.d.R. 1x wöchentlich).<br>Alle Therapeuten waren spezifisch professionell ausgebildet und anerkannt. Es gab ein Manual für die Studie und ein kurzes Training. Insgesamt 11 Therapeuten, davon 8 Frauen. Supervision von 1 Stunde pro 6 Stunden Patientenkontakt wurde vorausgesetzt.                                                                          |
| <b>3.4 (13) Evtl. weitere Kontroll-<br/>gruppen<br/>(Art, Dauer, Frequenz)</b> | Allgemeinärztlich übliche Standardbehandlung (GP care), in den ersten 4 Monaten sollten die Patienten nicht zu Psychotherapeuten/psychotherapeutischen Interventionen überwiesen werden.                                                                                                                                                                                                                                                                 |

|                | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                            | Rating      |
|----------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| K15.<br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | 1) ausführliches Therapiemanual<br>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie<br>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens                                                                                                                                            | B3<br><br>1 |
| K16.<br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | 1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung<br>2) Ex post facto Beschreibung der Kontrollbedingungen<br>3) keine Beschreibung der Kontrollbedingung                                                                                                                                                                                            | B4<br><br>1 |
| K17.<br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | 1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz<br>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab<br>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab | B5<br><br>1 |
| K18.<br>(B.6.) | Manualtreue, Treatment Integrity                                                     | 1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt<br>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt<br>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen                                                                                                                       | B6<br><br>1 |
| K19.<br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | 1) Intervention wie in klinischer Alltagspraxis<br>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert<br>3) Intervention gegenüber klinischer Alltagspraxis stark verändert                                                                                                                                                                                | C4<br><br>1 |



|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating       |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br><br>2  |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br><br>1  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br><br>1 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br><br>1 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br><br>1 |

**3.5 (19) Beobachtungszeitraum/ Messzeitpunkte/Katamnese**

Untersuchungen vor Behandlungsbeginn, nach weiteren 4 und 12 Monaten. Es gab keine Untersuchung exakt zu Behandlungsende. Durchschnittlich wurden die Therapien 1-2 Wochen nach der Aufnahmeuntersuchung begonnen und dauerten 6-7 Wochen, die erste Nachuntersuchung ist also etwa 3-5 Wochen nach Therapieende, die Katamnese etwa 9 Monate nach Therapieende.

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating       |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br><br>1  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br><br>1 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br><br>1 |

**3.6 (15) Anzahl der Zentren**

2: Department of Psychiatry and Behavioural Sciences, Royal Free

Campus of the Royal Free and University College Medical School, London, und das National Primary Care Research and Development Centre (NPCRDC) der Universität Manchester.

Insgesamt waren 24 allgemeinärztliche Praxen beteiligt

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

3 getrennte Gruppen wurden untersucht:  
 - entweder erfolgte eine randomisierte Zuweisung,  
 - eine randomisierte Zuweisung nur zwischen den Psychotherapien  
 - oder freie Wahl bei starker Präferenz für eine bestimmte Behandlung.  
  
 Die Randomisierung erfolgte blockweise nach hohen vs. niedrigen BDI-Werten (Kriterium war 23 Punkte).

**3.8 (17) ggf. Concealment der Randomisierung**

Nach eigener Angabe akzeptierte Methode des Concealments: Die Randomisierung erfolgte durch verschlossene, undurchsichtige Umschläge

|                | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating           |
|----------------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| K26.<br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                         | C8<br>3<br>bwz.1 |
| K27.<br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br>1<br>bwz.3 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

Depressive Symptome: Beck Depressions Inventar

### 4.2 (21) Sekundäre Zielkriterien

- Allg. Psychopathologie: Brief Symptom Inventory
- Soziales Funktionsniveau: Social Adjustment Scale SAS (Selbstrating)
- Patientenzufriedenheit: der Fragebogen basierte auf Elliott and Shapiro's impact of events scale
- Lebensqualität: EuroQoL

Direkte und indirekte Kosten  
 Selbsteinschätzung der individuellen Problemverbesserung durch die Therapie  
 Qualitative Interviews zu den Behandlungs- und Studiererfahrungen

Die Therapeuten füllten bei Therapieende einen Fragebogen zu Länge, Verlauf und Ergebnissen der Therapie aus.

### 4.3 () Verblindung der Erhebung der Zielkriterien

Die Zielkriterien werden als Selbsteinschätzungen der Patienten erhoben.

|                       | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                 | Rating  |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K28.</b><br>(A.7.) | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                   | A7<br>1 |
| <b>K29.</b><br>(A.8.) | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                       | A8<br>1 |
| <b>K30.</b><br>(C.9.) | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung) | C9<br>1 |
| <b>K31.</b><br>(A.9.) | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                  | A9<br>3 |

**4.4 () Statistische Methodik  
(Subgruppenanalysen,  
Sensibilitätsanalysen)**

ITT-Analyse, fehlende Daten mittels LOCF-Methode fortgeführt (im Vergleich dazu auch PP-Analyse durchgeführt).

Messwiederholungs-Varianzanalysen mit den Faktoren randomisiert zugewiesenen Gruppe (3gestuft) und Studienzentrum (2gestuft).

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>1  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>1 |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>1 |

**4.5 (22) Ergebnisse**

Die Messwiederholungsvarianzanalyse ergibt signifikante Effekte für den Zeitfaktor und die Interaktion Zeit x Behandlungsgruppe (wie gewünscht wird der der Faktor Behandlungsgruppe allein nicht signifikant, da die Gruppen sich bei Baseline noch nicht unterscheiden).

Beide Psychotherapien waren der GP care nach 4 Monaten in der Reduktion depressiver Symptomatik signifikant überlegen (der Unterschied im BDI von 4-5 Punkten ist auch klinisch relevant). Für die anderen Outcome-Maße galt das nicht (für die BSI-Werte gab es einen Trend in Richtung derselben Ergebnisse).

Die beiden Psychotherapien unterschieden sich nicht signifikant in ihrer Wirksamkeit, wenn alle randomisiert zugewiesenen Patienten analysiert wurden.

Die Patienten in Psychotherapie waren signifikant zufriedener mit ihrer Behandlung als die in GP care.

Nach 12 Monaten unterschieden sich die 3 Behandlungen nicht mehr, da die GP care sich deutlich verbessert hatte.

Die Patienten in nicht-direktiver Psychotherapie waren zufriedener als die der anderen beiden Behandlungen.

Das Outcome aller CBT vs. NDC Patienten, die randomisiert zugewiesen wurden, unterschied sich nicht signifikant.

Das Outcome von Patienten, die randomisiert einer bestimmten Behandlungsgruppe zugewiesen wurden, unterschied sich nicht von denen, die diese Behandlung frei gewählt hatten.

Die 3 Behandlungen unterschieden sich nicht in den Kosten, daher sind die Psychotherapien kurzfristig kosteneffektiver.

|                 | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K36.<br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-Maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| K37.<br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |

X

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

Keine Angabe

|                 | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                    | Rating   |
|-----------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K38.<br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br>3 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Nicht-direktive Psychotherapie und kognitive Verhaltenstherapie sind GP care in der kurzfristigen Reduktion depressiver Symptomatik überlegen und verursachen keine höhere Kosten als diese. Diese Überlegenheit ist nach 12 Monaten jedoch nicht mehr feststellbar, da sich die GP-Gruppe in ihren Verbesserungen angleicht.

CBT und NDC unterscheiden sich in ihrer Wirksamkeit nicht voneinander.

|                | Kriterium              | Operationalisierung                                                                | Rating  |
|----------------|------------------------|------------------------------------------------------------------------------------|---------|
| K39.<br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br>3) Hinweise auf Ergebnismanipulation | A1<br>1 |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |             |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1<br><br>1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2<br><br>1 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3          |

## 6. Praxistransfer

|                 | Kriterium                                                                | Operationalisierung                                                                                                                                                                                                                                                                                   | Rating       |
|-----------------|--------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen         | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br><br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br><br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br><br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlerqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br><br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br><br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br><br>1 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale             | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br><br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br><br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br><br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte            | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br><br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br><br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br><br>1 |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Die allgemeine Qualität ist ausreichend. Die großen Stärken der Studie gleichen die geringfügigen Schwächen aus.

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Nicht-direktive Kurzzeitpsychotherapie ist der allgemein-ärztlich üblichen Behandlung von Patienten mit Depressionen und mit Depression und Angst gemischt signifikant in der Reduktion der depressiven Symptomatik überlegen. Diese Wirksamkeit ist vergleichbar mit der Wirksamkeit von kognitiv-behaviouraler Kurzzeittherapie. Die Randomisierung bzw. Wahl einer präferierten Behandlung hatte keinen Einfluss auf das Behandlungsergebnis. Die Katamnese belegt, dass die erzielten Verbesserungen stabil bleiben.

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Der Einfluss der Randomisierung bzw. Präferenz wurde untersucht. Die Manualtreue/Einhaltung der Therapieform war gewährleistet. Als geringfügige Einschränkungen wurde der Einfluss der Begleitmedikation nicht untersucht und nicht alle Patienten erfüllten die ICD-Kriterien der Einschlussdiagnosen.

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Es handelt sich um eine Studie unter Feldbedingungen, es gibt daher keine (oder kaum) Einschränkungen in der Übertragbarkeit. Die Möglichkeit, bei starker Präferenz eine bestimmte Behandlung zu wählen, ist ebenfalls direkt übertragbar in die Praxis. Da sich die Patienten mit Präferenz nicht unterschieden von denen, die randomisiert zugewiesen wurden (unwesentliche Unterschiede bei Baseline und keine Unterschiede im Outcome), können die Ergebnisse der Studie mit hoher Wahrscheinlichkeit generalisiert werden auf die Gesamtgruppe depressiver Patienten. Der Zugangsweg in die Studie ist ebenfalls übertragbar auf die deutsche Praxis (dass nicht alle Patienten die exakten ICD-Kriterien erfüllen, entspricht auch der üblichen Praxis der psychotherapeutischen Versorgung in Deutschland). Die Begleitmedikation entspricht ebenfalls der üblichen Praxis in Deutschland.

### Stärken der Studie:

Das Design, das sowohl Randomisierung als auch freie Wahl der Behandlungsform ermöglichte und die Durchführung der Studie unter Feldbedingungen. Die Externe Validität ist kaum eingeschränkt ohne substanzielle Einbußen in der internen Validität.

### Schwächen der Studie:

Die Zielkriterien wurden ausschließlich über Selbsteinschätzungen der Patienten erhoben, allerdings stellt das Beck Depressions Inventar als primäres Zielkriterium ein etabliertes Messinstrument mit gesicherten psychometrischen Eigenschaften dar. Der Anteil fehlender Daten ist teilweise hoch, alle Analysen wurden jedoch sorgfältig sowohl für ITT- (mit forgeführten Daten) als auch PP-Stichproben berechnet, um dieses Defizit auszugleichen.

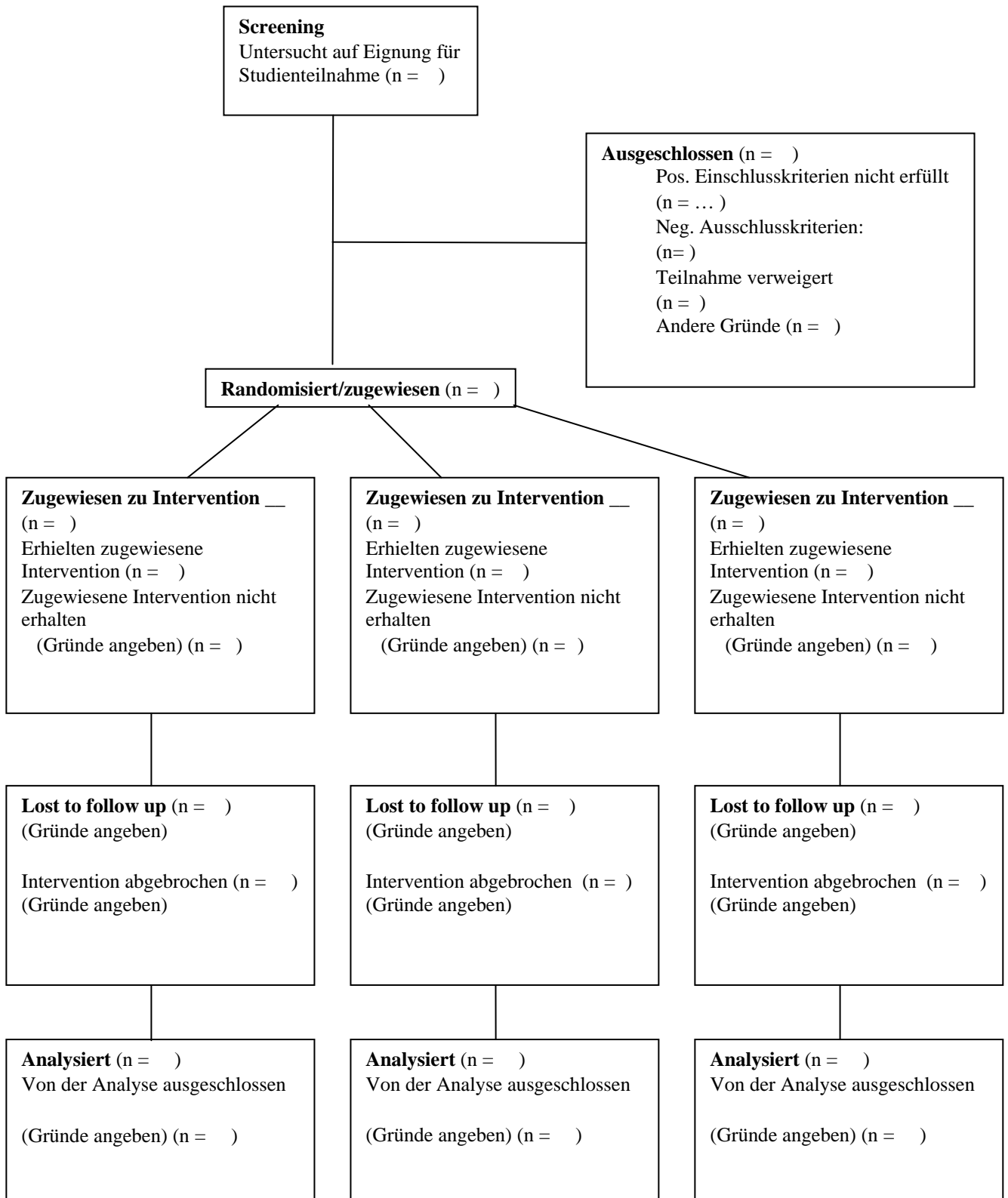
### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie belegt unter Feldbedingungen, die auf Deutschland übertragbar sind, die Wirksamkeit von nicht-direktiver Kurzzeitpsychotherapie auf die Reduktion depressiver Symptomatik bei depressiven Störungen. Die erreichten Verbesserungen sind klinisch relevant, der allgemeinärztlichen Versorgung signifikant überlegen, unterscheiden sich nicht signifikant von der Wirksamkeit kognitiv-behaviouraler Kurzzeitpsychotherapie und sind bis zur Katamnese zeitlich stabil.

Die Ergebnisse sind umso überzeugender, da es sich um durchschnittlich sehr kurze Therapien mit 7 Sitzungen handelt.



**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

HK

**Studie:**

Autoren:

Machado, L. A. C., Azevedo, D. C., Capanema, M. B., Neto, T. N. & Cerceau, D. M.

Titel:

Client-centered therapy vs exercise therapy for chronic low back pain: A pilot randomized controlled trial in Brazil.

Quelle:

Pain Medicine (2007), 3, 251-258.

## 1. Basisinformationen

1.1 () **Benennung der Vorgehensweise**  
(Verfahren/Methode/Interventionstechnik)

Clienten-centered therapy (Bezug auf Rogers, 1951)

1.2 (6) **Fragestellung / Zielsetzung der Studie / Forschungshypothese**

Gibt es bei Rückenschmerzpatienten (low back pain) Unterschiede im Behandlungserfolg zwischen einer klienten-zentrierten und einer Sport-/Bewegungsgruppe ?

1.3 (2) **Design**

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) **Evidenzstufe**

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Rückenschmerzen im unteren Rückenabschnitt („low back pain“)

|                      | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|----------------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K1.</b><br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>2 |
| <b>K2.</b><br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>1 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschluss: jünger als 65 Jahre; mindestens drei Monate anhaltende Schmerzen im Bereich zwischen 12. Rippe und Gefäßfalte;  
Ausschluss: mögliche oder nachgewiesene ernsthafte Wirbelsäulenerkrankung (Fraktur, Tumor, Infektion), frühere Wirbelsäulen-OP, Spondylolisthese, Schwangerschaft und andere Erkrankungen, die eine begleitende Intervention erfordern

|                      | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|----------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K3.</b><br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>2 |
| <b>K4.</b><br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>1 |
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>3 |

### 2.3 () Setting/Patientenzugang

- Über einen triage-center (wahrscheinlich ein übliche Art Verteilungsinstanz im brasilianischen Gesundheitssystem)

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

|  |  |  |   |
|--|--|--|---|
|  |  |  | 1 |
|--|--|--|---|

|                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                              |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>2.4 (7) Zentrale Patientenmerkmale</b><br/>(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)</p> | <p>Da getrennt für die beiden Behandlungsgruppen aufgeführt, hier nur Näherungswerte nach Augenschein</p> <p>Alter: M=43, SD=12,5; Geschlecht: 70% weiblich; bisherige Erkrankungsdauer in Monaten: M=68, SD=90; Anteil der patienten, die aufgrund ihrer Rückenschmerzen nicht arbeitstätig sind: 40%; Einnahme von Schmerzmitteln: 40%</p> |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                                 |               |
|-----------------------------------------------------------------|---------------|
| <p><b>2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl</b></p> | Keine Angaben |
|-----------------------------------------------------------------|---------------|

|                               | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating       |
|-------------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <p><b>K7.</b><br/>(A.16.)</p> | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br><br>3 |

|                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten</b><br/>(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)</p> | <p>Gescreent: 172; eingeschlossen: 39; randomisiert zugewiesen: 33 (klienten-zentriert: 16; Sport/Bewegung: 17; Posttreatment-erhebung: 21 ((klienten-zentriert: 11; Sport/Bewegung: 10); Katamnese: 27 (klienten-zentriert: 13; Sport/Bewegung: 14)</p> <p>Unklare Angaben zu Therapieabbruch: 28 Patienten (85% der Stichprobe!) gingen in die Intention to Treat-Analysen ein, dabei keine Aufschlüsselung nach Behandlungsgruppen</p> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating       |
|--------------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <p><b>K8.</b><br/>(A.5.)</p>   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br><br>2  |
| <p><b>K9.</b><br/>(A.15.)</p>  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br><br>1 |
| <p><b>K10.</b><br/>(A.17.)</p> | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br><br>2 |
| <p><b>K11.</b><br/>(A.3.)</p>  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br><br>1  |
| <p><b>K12.</b><br/>(A.4.)</p>  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br><br>1  |



## 2.7 (10) Vergleichbarkeit der Gruppen

|                                                                                                                                                                                |                   |                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------------|
| Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen                                                                                                                |                   |                |
| Zwar randomisierte Zuweisung und fast keine statistischen Unterschiede, doch einige Auffälligkeiten im Gruppenvergleich (nicht signifikant aufgrund der kleinen Gruppengrößen) |                   |                |
|                                                                                                                                                                                | Klientenzentriert | Sport/Bewegung |
| Dauer der Symptome (Monate)                                                                                                                                                    | 87,5              | 50,3           |
| Anteil Frauen                                                                                                                                                                  | 81,3              | 58,8           |
| Anteil aufgrund Rückenschmerz nicht arbeitender Patienten                                                                                                                      | 56,3              | 23,5           |
| nach BDI-Cut-Off klinisch depressiv                                                                                                                                            | 69%               | 82%            |
| visuelle Analog-Skala Schmerz                                                                                                                                                  | 6,0               | 7,0            |
| Einnahme von Schmerzmitteln (%)                                                                                                                                                | 31,3              | 47,1           |
| <ul style="list-style-type: none"> <li>➔ klientenzentrierte Gruppe: eher chronifiziert</li> <li>➔ Sport/Bewegung: eher akut</li> </ul>                                         |                   |                |
| Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)                                                                                                          |                   |                |
| Nicht gegeben: Kontrollgruppe „Sport-/Bewegungstherapie erhielt nur die Hälfte der Stunden der Klienten-zentrierten Therapie                                                   |                   |                |
| Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)                                                               |                   |                |
| Keine Angaben zu begleitenden Behandlungen                                                                                                                                     |                   |                |
| Vergleichbarkeit der Therapieadhärenz                                                                                                                                          |                   |                |
| Keine Angaben zur Überprüfung der Adhärenz                                                                                                                                     |                   |                |

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating       |
|-----------------|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | 1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen<br>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert<br>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt | B9<br><br>2  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | 1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen<br>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt<br>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage                                                                                                                       | A18<br><br>2 |

### 3. Intervention

|                                                                                |                                                                                                                                                                                                                                                                                                |
|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>3.1 (14) Studiendesign /<br/>Behandlungsarme</b>                            | <ul style="list-style-type: none"> <li>- Klienten-zentrierte Therapie</li> <li>- „Exercise Therapy“ (Sport-/ Bewegungstherapie)</li> </ul>                                                                                                                                                     |
| <b>3.2 (11) Behandlung Experimental-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Klienten-zentrierte Gruppentherapie („basierend auf den Prinzipien von Rogers, 1951“); Gruppengröße: 10 Patienten; Sitzungsdauer ja 80 Min.; Häufigkeit: 2*wöchentlich über 9 Wochen (insgesamt: 1440 Min)</li> </ul>                                 |
| <b>3.3 (12) Behandlung Kontroll-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b>     | <ul style="list-style-type: none"> <li>- Exercise Therapie (Sport-/ Bewegungstherapie): Übungsprogramm mit 20 Minuten Walking, Dehnungs- und Kräftigungsübungen; Gruppengröße: 10 Patienten; Sitzungsdauer ja 40 Min.; Häufigkeit: 2*wöchentlich über 9 Wochen (insgesamt: 720 Min)</li> </ul> |
| <b>3.4 (13) Evtl. weitere Kontroll-<br/>gruppen<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>-</li> </ul>                                                                                                                                                                                                                                            |

|                       | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                            | Rating      |
|-----------------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K15.</b><br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | 1) ausführliches Therapiemanual<br>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie<br>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens                                                                                                                                            | B3<br><br>2 |
| <b>K16.</b><br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | 1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung<br>2) Ex post facto Beschreibung der Kontrollbedingungen<br>3) keine Beschreibung der Kontrollbedingung                                                                                                                                                                                            | B4<br><br>1 |
| <b>K17.</b><br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | 1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz<br>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab<br>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab | B5<br><br>3 |
| <b>K18.</b><br>(B.6.) | Manualtreue, Treatment Integrity                                                     | 1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt<br>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt<br>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substanzielle Abweichungen                                                                                                                       | B6<br><br>3 |
| <b>K19.</b><br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | 1) Intervention wie in klinischer Alltagspraxis<br>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert<br>3) Intervention gegenüber klinischer Alltagspraxis stark verändert                                                                                                                                                                                | C4<br><br>2 |

|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7       |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6       |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br>1 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br>1 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br>1 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

- Pre-; Post-; Follow-Up-Erhebung (6 Monate nach Behandlungsbeginn)

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating   |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br>1  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br>1 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br>2 |

**3.6 (15) Anzahl der Zentren**

1 Einrichtung

**3.7 (16) Allokation zu Interventionsgruppen (Randomisierung,**

randomisierte Zuteilung

self-selection, ...)



**3.8 (17) ggf. Concealment der Randomisierung**

Gegeben; Fremdeinschätzer waren blind für die Gruppenzugehörigkeit der Patienten

|                       | <b>Kriterium</b>                                                                        | <b>Operationalisierung</b>                                                                                                                                                                                                                                                                                                                                                                                                                       | <b>Rating</b>  |
|-----------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br><b>3</b> |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br><b>2</b> |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

Fremdeinschätzungsverfahren

- Brazil Roland-Morris Questionnaire (BRM, Nusbaum et al., 2001) Skala zur Einschätzung von Beeinträchtigungen durch Rückenschmerzen

Selbsteinschätzungsverfahren

- Visuelle Analog Skala für Schmerzstärke (10 cm Skala: "kein Schmerz" --- "Schmerz wie er schlimmer nicht mehr möglich ist")

### 4.2 (21) Sekundäre Zielkriterien

Selbsteinschätzungsverfahren

- Beck-Depressions-Inventar (BDI)
- Ad hoc Einschätzung der Behandlung („excellent“ --- „schlecht“)

### 4.3 () Verblindung der Erhebung der Zielkriterien

- gegeben für Fremdeinschätzung durch Geheimhaltung der Treatmentzuweisung

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating       |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | <ol style="list-style-type: none"> <li>1) a priori definiert unter Berücksichtigung der Wertigkeit</li> <li>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien</li> <li>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,</li> </ol>                                                                                                                                                                         | A7<br><br>1  |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | <ol style="list-style-type: none"> <li>1) reliable und valide Outcome-Verfahren</li> <li>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren</li> <li>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend</li> </ol>                                                                                                                                                                             | A8<br><br>1  |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | <ol style="list-style-type: none"> <li>1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität</li> <li>2) Zielkriterien beziehen sich nur auf eine Dimension</li> <li>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)</li> </ol>                                                                                                                       | C9<br><br>1  |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | <ol style="list-style-type: none"> <li>1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen</li> <li>2) zwei Informationsquellen</li> <li>3) eine Informationsquelle</li> </ol>                                                                                                                                                                                                                                                                        | A9<br><br>2  |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | <ol style="list-style-type: none"> <li>1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler</li> <li>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern</li> <li>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten</li> </ol> | A10<br><br>1 |

**4.4 ( ) Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

Multiple Regressionsanalysen (simultan alle primären Outcomvariablen als Regressoren aufgenommen)  
 - Allerdings unklar, wie Gruppenzugehörigkeit statistisch modelliert wurde (durch Indikatorvariablen? Welche Art der Kodierung?)

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>1  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>3 |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>3 |

**4.5 (22) Ergebnisse**

Hinsichtlich der drei Outcome-Kriterien (Schmerz-Beeinträchtigung, Visuell-Analog Schmerzstärke, BDI) unterscheiden sich die Behandlungsgruppen nicht signifikant voneinander. Allerdings schneidet die Sport-/Bewegungsgruppe nach deskriptiven Parametern jeweils besser ab als die klienten-zentrierte Gruppe. Dies gilt auch hinsichtlich der subjektiven Behandlungseinschätzung (Sport: MD=exzellente vs Klienten-zentriert: Md=befriedigend) und Therapieanspruchnahme (Sport: M=8 Sitzungen; Klienten-zentriert: M=6 Sitzungen).

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcomemaßen und zu allen relevanten Messzeitpunkten                                              | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>2 |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

- keine Angaben/Analysen

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating              |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br><b>3</b> |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Kurzfristig (Prä-Post-Unterschiede) erwies sich die Sport-/Bewegungstherapie gegenüber dem klienten-zentrierten Ansatz als überlegen. Langfristig (6-Monats-Follow-up) zeigten sich aber keine Unterschiede mehr zwischen den Behandlungsgruppen.

|                       | Kriterium              | Operationalisierung                                                                    | Rating             |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br><b>1</b> |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>1 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>1 |



## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

#### **schwach**

- zu kleine Gruppengrößen (unzureichende statistische Power; offenbar substantielle wenn auch nicht statistisch signifikante Baseline-Unterschiede zwischen den Gruppen)
- hoher Anteil an Patienten mit fehlenden Daten zu den verschiedenen Messzeitpunkten. Es bleibt unklar auf welche Stichprobe sich die Intention to treat-Analyse bezieht, bzw. welche Patienten aus welchen Gründen hier nicht berücksichtigt wurden. Der Anteil der Patienten, die die therapeutischen Angebote vollständig durchgeführt haben ist extrem gering. Es stellt sich die Frage, ob bei der Durchführung der Studien bzw. der Behandlungsangeboten nicht grundsätzlich etwas schief gelaufen ist.

#### **Zusammenfassung der Ergebnisse**

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Hinsichtlich der Outcome-Variablen wurden keine signifikanten Unterschiede gefunden. Allerdings besteht konsistent der Trend, dass die Sport-/Bewegungsgruppe einen höheren Therapie-Benefit erzielte. Weiterhin (allerdings wiederum nur deskriptiv, da keine Inferenzstatistik präsentiert wurde) wurde die Sport-/Bewegungstherapie positiver beurteilt (Median: exzellent) im Vergleich zur klienten-zentrierten Therapie (Median=bedrohend) und auch länger besucht (8 Sitzungen vs. 6 Sitzungen)

#### **Interne Validität** (Sicherheit kausaler Schlussfolgerungen)

##### **nicht gegeben**

- zu deutliche Unterschiede zwischen den Behandlungsgruppen zur Baseline (die nur wegen der kleinen Gruppengröße nicht signifikant wurden): klienten-zentriert: eher chronische Patienten; Sport/bewegung: eher akute Patienten.
- Fragwürdige Datenanalysen: wiederum zu kleine Stichprobe (N=27) für Regressionsanalysen mit mindestens 4 Prädiktoren, zumal Kollinearitätsproblematik eine Rolle spielen dürfte (wahrscheinlich hohe Korrelationen zwischen Schmerzbeeinträchtigung, Schmerzstärke und Depressivität)
- (ferner: keinerlei Angaben zu begleitenden Interventionen.)

#### **Externe Validität**

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

#### **Stark eingeschränkt**

Bei der klienten-zentrierten Gruppe kann kaum noch von einer Behandlung gesprochen werden. Da scheint aus nicht bekannten Gründen die Umsetzung der Behandlungen gescheitert zu sein. Die Ergebnisse sind nicht übertragbar auch die Versorgungspraxis.

Weiterhin wurden nur nicht-komorbide Erkrankungen eingeschlossen, dadurch Verlust bei den gescreenten Patienten von etwa ca. 75%;

#### **Stärken der Studie:**

Überzeugende Auswahl/Zusammenstellung von Outcome-Kriterien; randomisierte Zuteilung, Concealment;

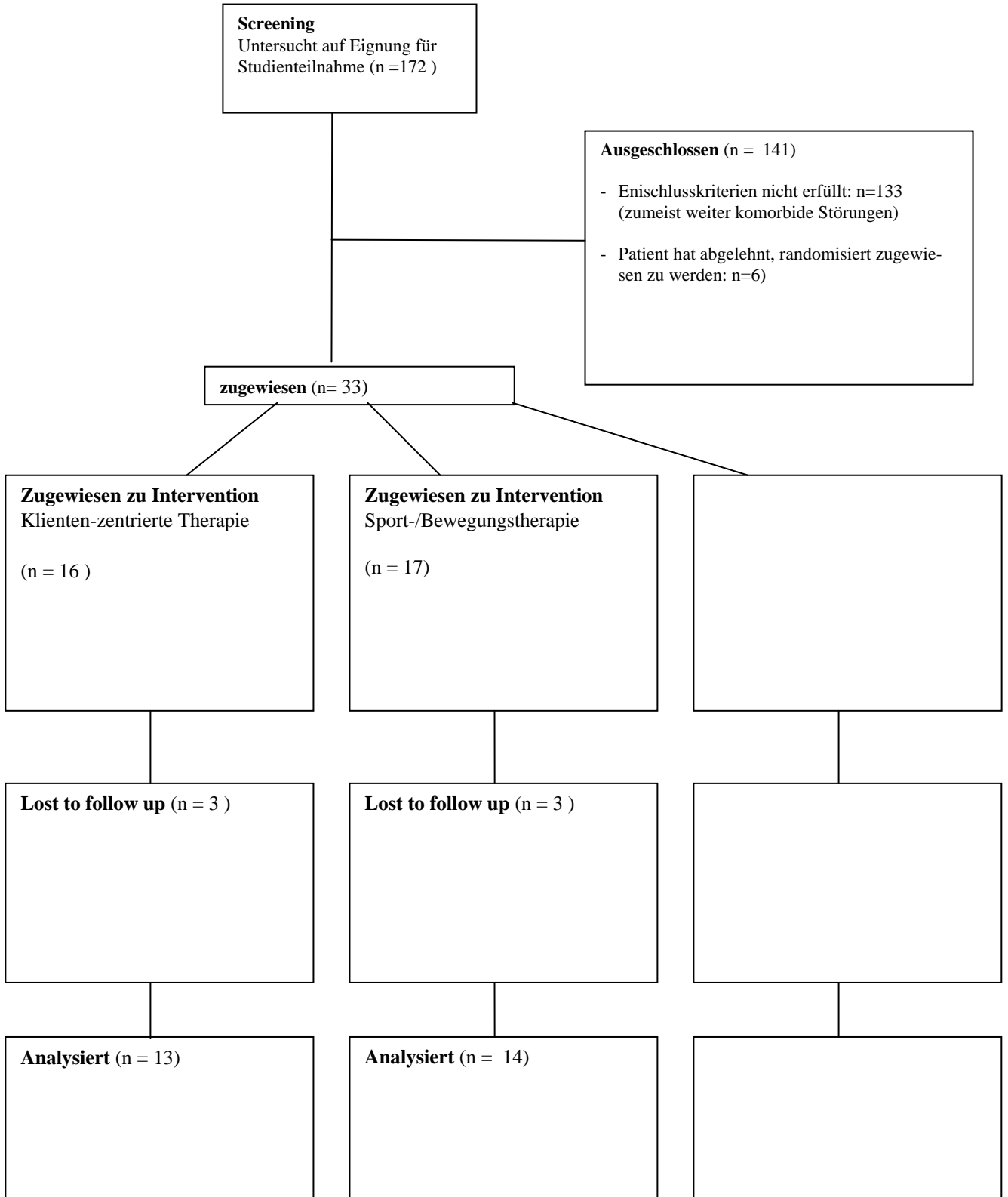
#### **Schwächen der Studie:**

Grundsätzlich zu geringe Datenbasis; zu hoher Anteil an vorzeitig abgebrochenen Behandlungen

#### **Bedeutung der Studie für die Bewertung der Psychotherapiemethode**

**irrelevant:** keine Überlegenheit des gesprächspsychotherapeutischen Ansatzes gegenüber Vergleichsbedingung nachweisbar; GT-Behandlung wurde allein von der Therapieadhärenz völlig unzureichend realisiert. zu kleine Datenbasis; zu hohe Abbruchquote.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

JS

**Studie:**

Autoren:

Markowitz, James et al.

Titel:

Treatment of Depression Symptoms in Human Immunodeficiency Virus-Positive Patients

Quelle:

Arch Gen Psychiatry, Vol.55; 1993

## 1. Basisinformationen

### 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

- supportive therapy (ST) – ähnelt der klientenzentrierten Therapie nach Rogers; Psychoedukation über Depression wird mit eingeschlossen
- legt keinen festen Rahmen oder spezifische Themen vor

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

- vergleichende Wirksamkeit von Interpersonaler Psychotherapie (IPT), Kognitiv-behavioraler Therapie (CBT) und Supportiver Therapie mit (SWE) und ohne Imipramin (ST)

### 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.



## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

- Depressive Symptome bei HIV-infizierten Patienten
- diagnostiziert anhand DSM-III-R

|               | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating   |
|---------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K1.<br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>1  |
| K2.<br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>3* |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

- IN:
- bekannter HIV-Positiv-Status seit mind. 6 Monaten
  - mind. 15 Punkte auf der 24-Items-Hamilton-Depression Rating Scale (HDRS)
  - klinische Feststellung bedeutsamer depressiver Symptome
  - körperliche Gesundheit ausreichend für ambulante therapeutische Behandlung
- OUT:
- bedeutsame medizinische Krankheit außer HIV
  - Schizophrenie, bipolare Störung
  - Kontraindikationen für Imipraminbehandlung
  - aktueller Substanzmissbrauch
  - kogn. Beeinträchtigung (MMST-Score <25)
  - Unfähigkeit, Englisch zu sprechen
  - andere psychiatrische Behandlung (außer HIV-Selbsthilfegruppe)

|               | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|---------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K3.<br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>1 |
| K4.<br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>1 |
| K5.<br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>2 |

**2.3 (7) Setting/Patientenzugang**

Anzeige und Überweisung

|                          | Kriterium                           | Operationalisierung                                                                                                                                                                                                                                                                                                                                            | Rating      |
|--------------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K6.</b><br><br>(C.2.) | Art der Rekrutierung der Stichprobe | 1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege<br>2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege<br>3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung) | C2<br><br>3 |

**2.4 (7) Zentrale Patientenmerkmale**

(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- 36,9 Jahre (SD = 6,9; Range: 24 - 59 Jahre)
- 85% Männlich
- 83% mind. College-Abschluss
- knapp 60% Weiß, ca.18% Schwarz, ca.19% Hispanic
- ca. 50% haben Achse-II-Diagnose und Substanzmissbrauch in der Geschichte
- 84% haben mind. 1 Jahr lang die Diagnose
- Schweregrad am Anfang (erste Probanden) eher leicht, später zunehmend mittlerer Schweregrad (spätere Probanden)

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

keine Angaben

|                           | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating       |
|---------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K7.</b><br><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br><br>2 |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- keine Angaben über Anzahl gescreenter im Gegensatz zu aufgenommenen Patienten
- insg. 101 Patienten zu den Gruppen randomisiert zugewiesen
- n = 69 beendeten die Therapie, d.h. 32 Dropout-Fälle
- Analyse sowohl für die Completer als auch ITT-Analyse durchgeführt

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>2  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>1 |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br>3 |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br>3  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>9  |

**2.7 (10) Vergleichbarkeit der Gruppen**

Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen

- keine sign. Unterschiede zwischen den Gruppen
- Schweregrad der körperlichen Erkrankung ist aber sehr verschieden (z.B. Schweregrad III und IV ca.51% bei CBT vs. 80% bei ST)
- mehr Persönlichkeitsstörungen bei IPT (n.s.)

Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)

- IPT und CBT-Probanden erhielten mehr Sitzungen (durchschn. 11,5); Unterschied zu ST/SWI (durchschn. 9,5) ist aber nicht sign.
- Sitzungen von IPT und CBT waren länger in der Dauer

Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)

- nur eine Gruppe erhielt Medikament (kein Placebo in den anderen Gruppen)

Vergleichbarkeit der Therapieadhärenz

- wurde kontrolliert und für alle Therapien festgestellt, dass Adhärenz da ist

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating       |
|-----------------|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | 1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen<br>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert<br>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanten Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt | B9<br><br>2  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | 1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen<br>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt<br>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage                                                                                                                       | A18<br><br>3 |

### 3. Intervention

|                                                                                |                                                                                                                          |
|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| <b>3.1 (14) Studiendesign /<br/>Behandlungsarme</b>                            | <ol style="list-style-type: none"> <li>1. IPT</li> <li>2. CBT</li> <li>3. ST</li> <li>4. ST + Imipramin (SWI)</li> </ol> |
| <b>3.2 (11) Behandlung Experimental-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b> | ST/SWI<br>zwischen 8 und 16 Sitzungen (je nach Bedarf), mit einer Dauer von 30 – 50 Minuten                              |
| <b>3.3 (12) Behandlung Kontroll-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b>     | IPT<br>16 Sitzungen à 50 Minuten innerhalb von 17 Wochen                                                                 |
| <b>3.4 (13) Evtl. weitere Kontroll-<br/>gruppen<br/>(Art, Dauer, Frequenz)</b> | CBT<br>16 Sitzungen à 50 Minuten innerhalb von 17 Wochen                                                                 |

|                | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                          | Rating                                                                                                  |
|----------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| K15.<br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | <ol style="list-style-type: none"> <li>1) ausführliches Therapiemanual</li> <li>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie</li> <li>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens</li> </ol>                                                                                                                                            | <span style="font-size: 1.5em; color: #cccccc;">B3</span><br><br><span style="font-size: 2em;">1</span> |
| K16.<br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung</li> <li>2) Ex post facto Beschreibung der Kontrollbedingungen</li> <li>3) keine Beschreibung der Kontrollbedingung</li> </ol>                                                                                                                                                                                            | <span style="font-size: 1.5em; color: #cccccc;">B4</span><br><br><span style="font-size: 2em;">1</span> |
| K17.<br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz</li> <li>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab</li> <li>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab</li> </ol> | <span style="font-size: 1.5em; color: #cccccc;">B5</span><br><br><span style="font-size: 2em;">2</span> |
| K18.<br>(B.6.) | Manualtreue, Treatment Integrity                                                     | <ol style="list-style-type: none"> <li>1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt</li> <li>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt</li> <li>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen</li> </ol>                                                                                                                       | <span style="font-size: 1.5em; color: #cccccc;">B6</span><br><br><span style="font-size: 2em;">1</span> |
| K19.<br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | <ol style="list-style-type: none"> <li>1) Intervention wie in klinischer Alltagspraxis</li> <li>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert</li> <li>3) Intervention gegenüber klinischer Alltagspraxis stark verändert</li> </ol>                                                                                                                                                                                | <span style="font-size: 1.5em; color: #cccccc;">C4</span><br><br><span style="font-size: 2em;">1</span> |

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

|                       |                                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                     |          |
|-----------------------|---------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
|                       |                                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 3        |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                   | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                               | C6<br>2  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler<br>(insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                  | C7a<br>1 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                           | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln). | C7b<br>1 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                           | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                 | C7c<br>3 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

- keine Angabe über Beobachtungszeitraum (anzunehmen, dass es mehrere Jahre waren)
- MZPe: Woche vor Behandlungsbeginn; Woche 8; Woche 16

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating   |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br>2  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br>2 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br>3 |

**3.6 (15) Anzahl der Zentren**

keine Angaben, wahrscheinlich mehrere

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

balancierte Randomisierung mittels Computerverfahren

**3.8 (17) ggf. Concealment der Randomisierung**

blinde Zuteilung der Probanden durch PC

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                              | Rating             |
|-----------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br><br><b>3</b> |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br><br><b>2</b> |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

Hamilton Depression Rating Scale (HDRS oder HAMD) (24 Items) und Beck Depression Inventory (BDI)

### 4.2 (21) Sekundäre Zielkriterien

CD4 cell count und 100-Punkte-Kalinowsky-Skala für Einschätzung physischen Zustandes

### 4.3 () Verblindung der Erhebung der Zielkriterien

- ja, unabhängige trainierte Bewerter, die blind für die Bedingung waren
- Ausnahme: bei ITT-Analyse wurde als letzter Wert die Beurteilung des Therapeuten auf der HAMD genommen (Therapeuten bewerteten Patienten nach jeder Sitzung)

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>1  |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | A8<br>1  |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | C9<br>2  |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | A9<br>2  |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10<br>1 |



4.4 () **Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

ANCOVA, kontrolliert für Ausgangswerte  
ANOVA für wiederholte Messungen

|                 | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|-----------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K33.<br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>1  |
| K34.<br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| K35.<br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>1 |

4.5 (22) **Ergebnisse**

ITT:  
 ○ HAMD nach 8 und 16 Wochen in allen Gruppen sign. gesunken; bei IPT und SWI sign. stärker als bei CBT und ST  
 ○ BDI nach 8 Wochen bei IPT und SWI, nach 16 Wochen bei allen Gruppen sign. geringer; bei SWI sign. geringer als bei CBT und ST; bei IPT sign. geringer als bei ST  
 Completer:  
 ○ im HAMD ist IPT sign. besser als CBT und ST  
 ○ im BDI ist IPT sign. besser als CBT und ST; SWI sign. besser als ST  
 ○ keine sign. Unterschiede zwischen IPT und SWI bzw. CBT und ST  
 Remissionsraten spiegeln die gleichen Ergebnisse wider: IPT 46%, CBT 30%, ST 21% UND SWI 50%

|                 | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K36.<br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| K37.<br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>2 |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

keine Angaben

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating              |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br><b>3</b> |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

IPT und SWI für diese Subpopulation anscheinend am besten  
 kritisch: geringe Anzahl von Diagnosen (53%), weniger Behandlungsin-  
 tensität bei ST

|                       | Kriterium              | Operationalisierung                                                                    | Rating             |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br><b>1</b> |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>2 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>1 |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

- insgesamt mittelmäßig
- gut: ITT-Analyse; verschiedene Therapeuten, Monitoring
- Nachteil:
  - Manches ist nicht beschrieben, was methodische Mängel beinhalten könnte, z.B. Zeitraum (hat den Anschein, als könnten es mehrere Jahre sein → Verzerrungen möglich, da z.B. anderer Umgang mit HIV) und Rekrutierung und Dropouts
  - verschiedene Dauer der Behandlungen, dadurch uneinheitliche Zeitabstände der Messzeitpunkte zur Beendigung der Therapie

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

- IPT und Supportive Therapie mit Imipramin sind sign. erfolgreicher bei HIV-Patienten mit depressiven Symptomen als CBT und Supportive Therapie ohne Imipramin; aber: kein Follow-Up!
- klinische Relevanz noch fraglich, da nur 53% mit Diagnose affektiver Störung

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

- insgesamt mittel: zwar gutes Monitoring und valide Messinstrumente, aber Gruppen sind nicht vergleichbar hinsichtlich der Dauer und Intensität der Behandlung
- teilweise große Unterschiede in der Verteilung der Merkmale zwischen den Gruppen (Schweregrad, Persönlichkeitsstörungen, Diagnosen) – Analyse: Persönlichkeitsstörungen wirken sich nicht aus; jedoch bei höherem Schweregrad mehr Verbesserung

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

++ viele Behandler, gute Übertragbarkeit aufs Versorgungssystem  
-- einjähriges Training

### Stärken der Studie:

- verschiedene Therapeuten
- ITT-Analyse

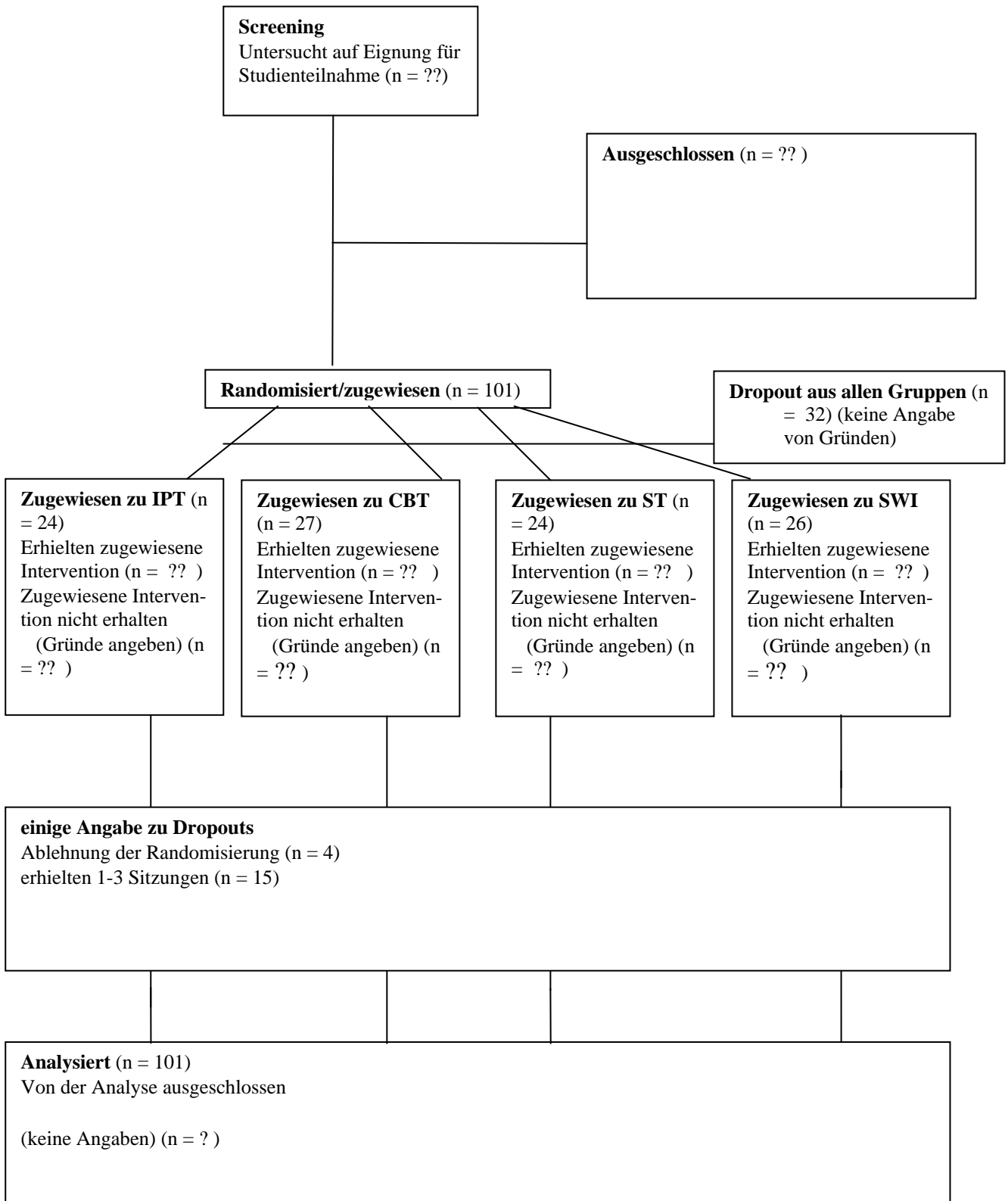
### Schwächen der Studie:

- Rekrutierung und Dropout-Fälle nicht beschrieben
- verschiedene Intensitäten der Gruppen; Umgang mit Medikamentenverschreibung nicht beschrieben (zusätzlicher Kontakt?)

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

- Die Studie zeigt eine hohe Wirksamkeit von ST in Kombination mit Imipramin, aber nicht ohne → spricht hier eher für die medikamentöse Wirkung und wenig Effekte der ST allein. Jedoch ist die Vergleichbarkeit der Gruppen dadurch eingeschränkt, dass die Intensität der Behandlungen stark unterschiedlich ist. Die Intensität der ST geringer als bei den anderen.
- Die methodischen Voraussetzungen für einen Wirksamkeitsnachweis sind dahingehend nicht gegeben, als dass keine unbehandelte Kontrollgruppe existiert und die anderen Behandlungen intensiver waren.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

zu K2: nur 53% der Probanden erfüllten die Diagnose einer aktuellen affektiven Störung, noch weniger die Diagnose einer aktuellen oder life-time-Depression



# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

JS

**Studie:**

Autoren:

Meyer

Titel:

The Hamburg Short Psychotherapy Comparison Experiment

Quelle:

Karger, 1981

## 1. Basisinformationen

### 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Klientenzentrierte Psychotherapie (Kurzzeittherapie)  
○ folgt den freien Assoziationen des Patienten, akzeptiert und unterstützt ihn, fördert die Selbstexploration und den Ausdruck von Emotionen

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Vergleich von klientenzentrierter Kurzzeittherapie mit psychodynamischer Kurzzeittherapie und Aussagen über ihre Effektivität

### 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

indikationsübergreifend – keine Angaben über Diagnosen oder Häufigkeit bestimmter psychischer Störungen

|                      | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|----------------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K1.</b><br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>3 |
| <b>K2.</b><br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>2 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

IN:  
 Alter zwischen 20 und 40  
 keine vorherige Psychotherapie

|                      | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|----------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K3.</b><br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>2* |
| <b>K4.</b><br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>3  |
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>3  |

### 2.3 () Setting/Patientenzugang

Überweisung durch Arzt oder andere Kliniken, Anmeldung zum Gespräch in Uniambulanz

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

**2.4 (7) Zentrale Patientenmerkmale**  
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

Angaben werden nur für alle n = 214 Patienten gemacht, für die die ersten Testdaten erhoben wurden

- 59% Frauen
- 28,1 Jahre (SD = 5,1)
- 47% single, 43% verheiratet; 11% arbeitsunfähig; Überrepräsentation des mittleren und oberen sozialen Status
- keine Angaben
- keine Angaben
- 31% länger als 10 Jahre an Symptomatik leidend, 27% weniger als 2 Jahre
- keine Angaben; Patienten in allen Skalen des FPI signifikant auffällig

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

keine Angaben

|                | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating   |
|----------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K7.<br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br>2 |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- 286 wurden eingeladen, 214 kamen zur ersten Testung; dann zwei Aufnahmeinterviews – 14 beendeten diese nicht vollständig, 18 wurden ausgeschlossen, 5 wollten keine Therapie
- für restliche 177 Patienten wurde Fallvorstellung im „focus formulation seminar“ durchgeführt, dort wurde Prognose formuliert
- in erster Hälfte der Untersuchung wurden nur Patienten mit „guter“ oder „mittlerer“ Prognose zur Therapie zugeteilt, in der zweiten Hälfte auch noch Patienten mit „negativer“ Prognose
- ausgeschlossen wurden alle Patienten mit „kontraindikativer“ Prognose
- 88 Patienten wurde Therapie angeboten, 19 davon wollten nicht, eine Person trat aus der Untersuchung aus
- → N = 68 Patienten, die zugeteilt wurden

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>2  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>3 |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br>2 |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br>1* |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>2  |

## 2.7 (10) Vergleichbarkeit der Gruppen

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen</p> <ul style="list-style-type: none"> <li>○ keine sign. Unterschiede zwischen den Gruppen (weniger als statistisch zu erwarten wären)</li> <li>○ unmittelbare Behandlungsgruppen etwas höhere Symptomschwere als Wartegruppen (Tendenz)</li> </ul> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)</p> <ul style="list-style-type: none"> <li>○ PDT in selber Zeit mehr Sitzungen als KZT</li> </ul> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)</p> <ul style="list-style-type: none"> <li>○ keine Angaben, wurde nicht kontrolliert</li> </ul> <p>Vergleichbarkeit der Therapieadhärenz</p> <ul style="list-style-type: none"> <li>○ Setting und Konzeption der Therapie war für die KZT „normal“, für die PDT ungewöhnlich und für diese Therapeuten schwieriger umzusetzen</li> </ul> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating        |
|-----------------|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | 1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen<br>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert<br>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanten Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt | B9<br><br>1   |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | 1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen<br>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt<br>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage                                                                                                                       | A18<br><br>2* |

### 3. Intervention

#### 3.1 (14) Studiendesign / Behandlungsarme

- 4 Gruppen:
- Klientenzentrierte Therapie (KZT)
  - Psychodynamische Therapie (PDT)
  - 2 Wartegruppen

#### 3.2 (11) Behandlung Experimental- gruppe (Art, Dauer, Frequenz)

- Klientenzentrierte Therapie (KZT)
- maximal 30 Sitzungen
  - 2 Sitzungen pro Woche für 15 Wochen
  - durchschn. 18,83 Sitzungen in 19,67 Wochen

#### 3.3 (12) Behandlung Kontroll- gruppe (Art, Dauer, Frequenz)

- Psychodynamische Therapie (PDT)
- maximal 30 Sitzungen
  - 2 Sitzungen pro Woche für 15 Wochen
  - durchschn. 26,03 Sitzungen in 19,02 Wochen

#### 3.4 (13) Evtl. weitere Kontroll- gruppen (Art, Dauer, Frequenz)

- Wartegruppen
- warteten auf Behandlung
  - Wartezeit durchschn. 19,1 Wochen für KZT, durchschn. 17,4 Wochen für PDT

|                | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                            | Rating       |
|----------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K15.<br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | 1) ausführliches Therapiemanual<br>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie<br>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens                                                                                                                                            | B3<br><br>2* |
| K16.<br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | 1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung<br>2) Ex post facto Beschreibung der Kontrollbedingungen<br>3) keine Beschreibung der Kontrollbedingung                                                                                                                                                                                            | B4<br><br>1  |
| K17.<br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | 1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz<br>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab<br>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab | B5<br><br>2  |
| K18.<br>(B.6.) | Manualtreue, Treatment Integrity                                                     | 1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt<br>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt<br>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen                                                                                                                       | B6<br><br>1  |
| K19.<br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | 1) Intervention wie in klinischer Alltagspraxis<br>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert<br>3) Intervention gegenüber klinischer Alltagspraxis stark verändert                                                                                                                                                                                | C4<br><br>2  |

|                | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating       |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K20.<br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br><br>3  |
| K21.<br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br><br>1  |
| K22.<br>(C.7.) | Qualifikation der Behandler<br>(insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                     | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br><br>2 |
|                | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br><br>1 |
|                | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br><br>1 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

|                                                                                                                                                                                                                                                                                                              |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1972, für 1,5 Jahre<br>MZPe:<br>o Eingangstestung, 3-4 Aufnahme-Interviews mit Therapeut; dann entweder:<br>- Behandlungsgruppe mit Post-, Follow-Up-Messung nach 3 und nach 9 Monaten oder<br>- Wartegruppe mit nochmaliger Testung nach Wartezeit, dann Post-, Follow-Up-Messung nach 3 und nach 9 Monaten |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                 | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating       |
|-----------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K23.<br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br><br>1  |
| K24.<br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br><br>1 |
| K25.<br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br><br>2 |



|                                                                                          |                |
|------------------------------------------------------------------------------------------|----------------|
| <b>3.6 (15) Anzahl der Zentren</b>                                                       | Keine Angaben  |
| <b>3.7 (16) Allokation zu Interventionsgruppen</b> (Randomisierung, self-selection, ...) | Randomisierung |
| <b>3.8 (17) ggf. Concealment der Randomisierung</b>                                      | ja             |

|                       | <b>Kriterium</b>                                                                        | <b>Operationalisierung</b>                                                                                                                                                                                                                                                                                                                                                                                                                       | <b>Rating</b>      |
|-----------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br><br><b>3</b> |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br><br><b>2</b> |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

- Freiburger Persönlichkeitsinventar (FPI)
- Gießen-Test (GT)
- Adaptation der Taylor Manifest Anxiety Scale (SAL)
- Leistungs-Prüf-System (LPS)
- Symptom-Liste mit 152 Items neurotischer oder psychosomatischer Beschwerden, nach Häufigkeit und Intensität geratet
- Therapeuten-Sitzungs-Fragebogen nach jeder Sitzung
- Therapeuten-Fragebogen zum Therapieerfolg, der Nützlichkeit der Fokusformulierung und der Passung des Patienten
- Patienteneinschätzung des Erfolges zu Follow-Up-Messungen

### 4.2 (21) Sekundäre Zielkriterien

/

### 4.3 () Verblindung der Erhebung der Zielkriterien

ja

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>2  |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | A8<br>1  |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | C9<br>1  |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | A9<br>1  |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10<br>2 |

**4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

- einfaktorische Anova für Vergleich von Wartegruppe einer Therapie- richtung mit sich selbst nach Therapie
- zweifaktorielle Varianzanalyse für Vergleich aller Warte-Patienten mit den zwei unmittelbaren Behandlungsgruppen

|                        | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                      | Rating          |
|------------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungs- gruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                         | A6<br><b>1</b>  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                   | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der An- wendungsvoraussetzungen bei gleichzeitiger Robustheit der an- gewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziel- lem Risiko für deren Verletzung | A13<br><b>2</b> |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statisti- schen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                                     | A14<br><b>2</b> |

**4.5 (22) Ergebnisse**

Insgesamt nur in sieben von 314 Variablen sign. Unterschiede zwi- schen den Behandlungs-Gruppen, diese jedoch zugunsten KZT

FPI:

- in Skala „Psychosomatische Beschwerden“ wird „normales“ Level von KZT signifikant erreicht, von PDT nur in Tendenz
- „Aggression“ erhöht sich bei PDT, nicht bei KZT
- in Skala „Depression“ sign. Verbesserung bei KZT, nicht bei PDT
- in Skala „Soziabilität“ und in Maskulinitäts-Skala für beide Gruppen sign. Annäherung an Normalität, die im Follow-Up erhalten bleibt

SAL:

- Angstwerte gehen signifikant zurück in KZT-Gruppe, in geringerem Ausmaß in PDT

GT:

- Testwerte inkonsistent und teilweise widersprüchlich

Symptom-Liste:

- Keine signifikanten Unterschiede zwischen den Gruppen, KZT ist in 12 Symptomen, PDT in 3 Symptomen besser als die Warte- gruppe

|             | Kriterium                                                 | Operationalisierung                     | Rating |
|-------------|-----------------------------------------------------------|-----------------------------------------|--------|
| <b>K36.</b> | Vollständige Darstellung der Ergebnisse zu allen Outcome- | 1) für alle Outcome-Variablen berichtet | A11    |

|                            |                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |              |
|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K37.</b><br><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br><br>1 |
|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

Werden unsystematisch berichtet

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating              |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br><b>2</b> |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Abschließende Diskussion wird nicht präsentiert

|                       | Kriterium              | Operationalisierung                                                                    | Rating             |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br><b>1</b> |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>3 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>1 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>3 |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

- Sehr breiter Ansatz – Design gut (Prä, Post, Follow-Up), aber Methodik teilweise ungewöhnlich
- Keine Angaben über Diagnosen oder Diagnostik und einzelne Symptome
- Statistische Auswertung ist der Datenlage angemessen

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

- sehr wenig signifikante Unterschiede zwischen den Treatment-Gruppen – die, die signifikant sind, sind zugunsten der KZT
- KZT und PDT signifikant besser als Wartegruppe – KZT weist in mehr Variablen sign. Unterschiede zur Wartegruppe auf

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

- Einschlusskriterien für die Untersuchung schon sehr auf positive Prognose ausgerichtet (Eignung für PDT bzw. KZT als Voraussetzung zur Teilnahme)
- Ausschluss von Patienten mit negativer Prognose – dadurch positive Verzerrung der Ergebnisse
- Therapeuten kannten teilweise die Prognose der Patienten – Einfluss darauf auf die Therapie wird nicht untersucht
- KZT-Therapeuten waren bevorteilt, da Setting und Art der Therapie ungewohnt für PDT-Therapeuten; außerdem waren KZT-Therapeuten erfahrener
- Begleitende andere Interventionen wurden nicht erfasst oder kontrolliert

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

- Intervention (vor allem das „Fokus-Seminar“ und seine Bedeutung für den Therapieverlauf) ist unüblich und nicht in der Praxis realisierbar
- Großer Teil der Therapeuten sind unerfahren
- Übertragbarkeit ist nicht gegeben
- Keine Angaben über Profil der Patienten

### Stärken der Studie:

- Prä-, Post- und Follow-Up-Messung
- Verschiedene Quellen
- Hoher methodischer Aufwand

### Schwächen der Studie:

- „für viel Aufwand kommt nicht viel bei rum“ – Ergebnisse sind nicht sehr aufschlussreich im Vergleich zum großen finanziellen, zeitlichen und personellen Aufwand für die Untersuchung

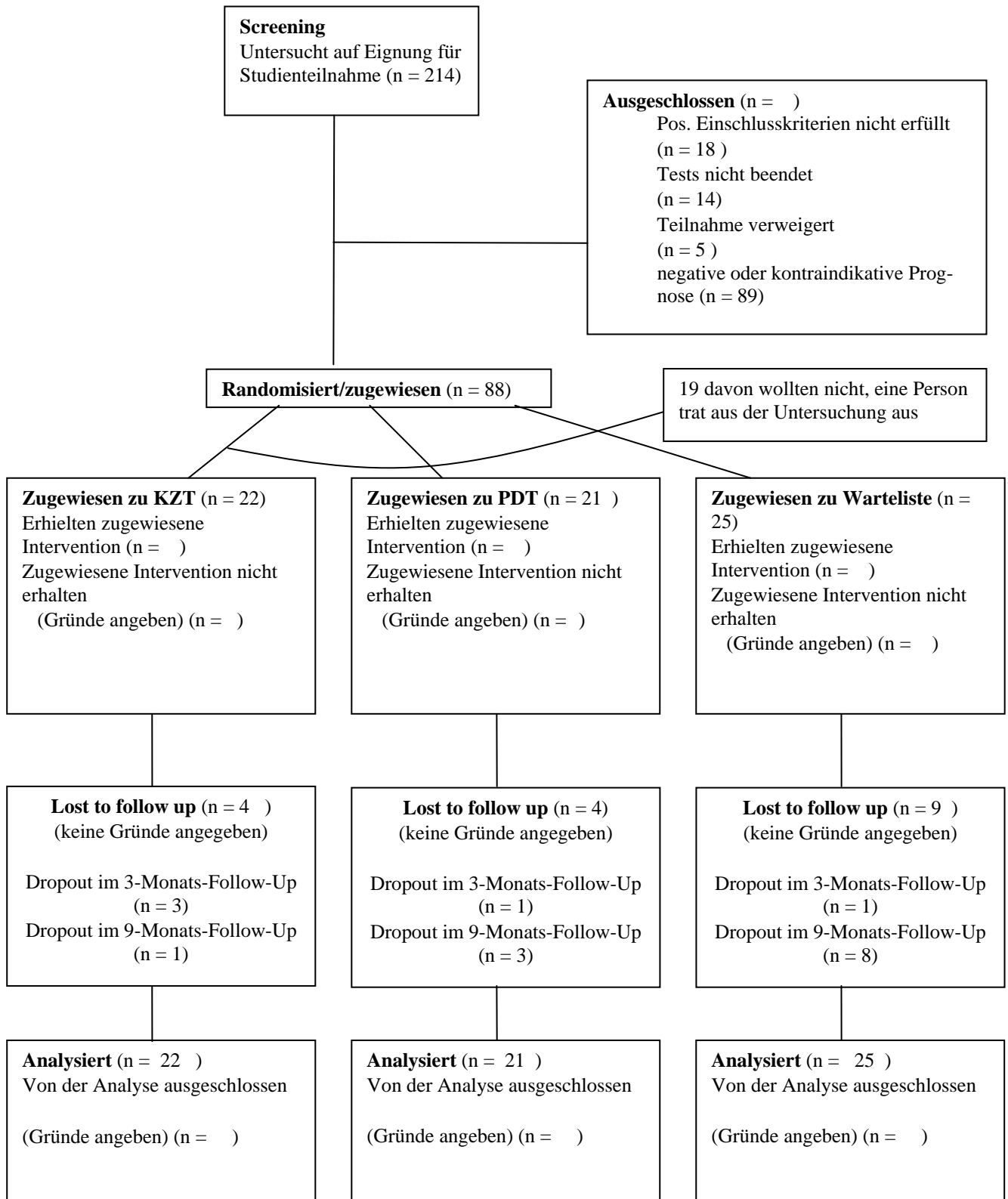
### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

- Beide Therapieverfahren sind im Vergleich zur Wartegruppe wirksam
- Zeigt keine überzeugenden Hinweise auf eine Überlegenheit der KZT gegenüber PDT, nur Tendenzen dazu
- Methodisch lässt die Studie Aussagen über die generelle Wirksamkeit von KZT im Gegensatz zu einer unbehandelten Stichprobe beschränkt zu. Sie zeigt außerdem, dass KZT genauso wirksam wie PDT ist und in einigen wenigen Bereichen sogar überlegen.





**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

zu K3: der Ausschluss wird hier zum großen Teil durch das Fokus-Formulierungs-Seminar bestimmt – Richtlinien zur Bestimmung der Prognose sind die eigentlichen Ausschlusskriterien:

Patient muss...

- fähig und willig sein, Gefühle zu explorieren
- konstruktiv mit Interpretationen umgehen
- eine Arbeitsbeziehung mit dem Therapeuten herstellen
- in einer therapeutischen Beziehung basierend auf Interpretationen arbeiten
- seine Symptome als ich-dyston erleben
- mit der Zeitbegrenzung umgehen können

das Seminar muss...

- das Patientenproblem psychodynamisch verstanden haben
- einen Therapieplan erstellen können

zu K11: durch starkes Ausschließen unpassender Patienten entstand auch kein hoher Dropout – ist aber sicherlich nicht repräsentativ

zu K14: es wurden keine klassischen Dropout-Analysen, aber ein Vergleich zwischen Therapieablehnern und –befürwortern durchgeführt

zu K15: für KZT nur Beschreibung der Therapiemerkmale; für PDT ausführlicher Therapieplan

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

JS

**Studie:**

Autoren:

Paivio & Nieuwenhuis

Titel:

Efficacy of emotion focused therapy for adult survivors of child abuse

Quelle:

Journal of Traumatic Stress 2001

# 1. Basisinformationen

## 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

emotion focused therapy bei Erwachsenen mit Missbrauch im KA (EFT-AS):

- Weiterentwicklung der GT nach Greenberg
- zentrale Rolle der Emotionen – Schwerpunkt auf der Rolle adaptiver Emotionen und auf der therapeutischen Beziehung
- Methoden: Traumaerinnerungen wachrufen, damit konfrontieren und modifizieren; verhinderte Emotionen zulassen und deren Information für die Bewertung nutzen; korrigierende interpersonale Erfahrung mit dem Therapeuten erleben

## 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

- Effektivität der EFT-AS bewerten
- Anzeichen für die Veränderung erhalten um emotionale Prozesse, die für Veränderung notwendig sind, weiter erforschen zu können

## 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

## 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Erwachsene mit Missbrauchserfahrungen im Kindesalter erfasst durch Screening, dann

- PTSD Symptom Severity-Interview (PSS-I)
- Symptoms Checklist (SCL)
- Impact of Event Scale (IES)
- Childhood Trauma Questionnaire (CTQ)

|               | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|---------------|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K1.<br>(A.2.) | Objektive und reliable Diagnostikstellung (mittels (teil-) standardisierter Interviews) | 1) Diagnostikstellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnostikstellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnostikstellung                                                                                                           | A2<br>2 |
| K2.<br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                               | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>2 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

OUT:

- <18 Jahre
- begleitende andere Therapie
- Einnahme psychoaktiver Medikation
- Krisensituation, die unmittelbarer Aufmerksamkeit bedarf
- Keine bewussten Erinnerungen an Kindesmissbrauch
- Primäres psychisches Problem beinhaltet extreme emotionale Dysregulation mit einem Risiko der Selbstschädigung
- In einer gewaltsamen Partnerschaft sein
- Aktuelle Drogen- oder Alkoholprobleme
- <.60 beim Global Severity Index (GSI) der SCL
- <15 bei der IES

IN:

- Kriterien für einsichtsorientierte Kurzzeittherapie wie Motivation, Fähigkeit zur therapeutischen Beziehungsherstellung, Fähigkeit sich auf den Kindesmissbrauch als Thema zu fokussieren, Zustimmung zum Behandlungsfokus des Wiedererlebens des Traumas, emotionaler Ausdruck

|               | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|---------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K3.<br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>1 |
| K4.<br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>2 |

|                      |                                                              |                                                                                                                                                                                                                                                                                                         |         |
|----------------------|--------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien | C3<br>2 |
|----------------------|--------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|

### 2.3 () Setting/Patientenzugang

Durch Überweisung und Anzeigenwerbung

|                      | Kriterium                           | Operationalisierung                                                                                                                                                                                                                                                                                                                                            | Rating  |
|----------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K6.</b><br>(C.2.) | Art der Rekrutierung der Stichprobe | 1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege<br>2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege<br>3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung) | C2<br>3 |

### 2.4 (7) Zentrale Patientenmerkmale (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

|                                    | EFT                                | Warteliste     |
|------------------------------------|------------------------------------|----------------|
| Alter                              | 36,9                               | 36,3           |
| Geschlecht                         | 81,8% weiblich                     | 87,5% weiblich |
| Familieneinkommen/Jahr < 40.000 \$ | 58,2%                              | 75%            |
| Ethnizität                         | 42 Weiß, 3 Aboriginal, 1 Asiatisch |                |
| Achse II Diagnose                  | 27,3%                              | 33,3%          |

### 2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Keine Angaben

|                       | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating   |
|-----------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K7.</b><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br>2 |

### 2.6 (9) Anzahl gescreener, eingeschlossener, zugewiesener und ausgewerteter Patienten (Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- 110 Probanden wurden gescreent (Info, dass abhängig von Kapazitäten, sie bis zu 20 Wochen warten müssten), davon 63 anschließend interviewt
- 46 Patienten wurden in Studie eingeschlossen (22 EFT, 24 Warteliste)

|                       | Kriterium                                  | Operationalisierung                                                                                                                                                                                 | Rating   |
|-----------------------|--------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)  | Stichprobengröße pro Gruppe                | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe) | A5<br>2  |
| <b>K9.</b><br>(A.15.) | Intention to treat – Analysen durchgeführt | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                | A15<br>2 |



|                       |                                                                                 |                                                                                                                                                                                                                                  |         |
|-----------------------|---------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
|                       |                                                                                 |                                                                                                                                                                                                                                  | 2       |
| <b>K11.</b><br>(A.3.) | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen | A3<br>2 |
| <b>K12.</b><br>(A.4.) | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen | A4<br>2 |

**2.7 (10) Vergleichbarkeit der Gruppen**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen</p> <ul style="list-style-type: none"> <li>Keine signifikanten Unterschiede bezüglich der erhobenen demographischen und klinischen Ausgangsvariablen</li> </ul> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)</p> <ul style="list-style-type: none"> <li>Wartelistegruppe hatte keine Behandlung im selben Zeitraum; nachfolgende Therapie hatte gleiche Intensität</li> </ul> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)</p> <ul style="list-style-type: none"> <li>vergleichbar</li> </ul> <p>Vergleichbarkeit der Therapieadhärenz</p> <ul style="list-style-type: none"> <li>Keine sign. Unterschiede, insgesamt hoch</li> </ul> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Rating       |
|-----------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | <p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p> | B9<br><br>1  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | <p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>                                                                                                                       | A18<br><br>1 |

### 3. Intervention

|                                                                                |                                                                                                                                                                                     |
|--------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>3.1 (14) Studiendesign /<br/>Behandlungsarme</b>                            | <ol style="list-style-type: none"> <li>1. EFT</li> <li>2. Warteliste, die nach der Wartezeit ebenfalls EFT bekommt</li> </ol>                                                       |
| <b>3.2 (11) Behandlung Experimental-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b> | <p>EFT:<br/>20 Stunden einmal in der Woche nach Behandlungsmanual</p>                                                                                                               |
| <b>3.3 (12) Behandlung Kontroll-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b>     | <p>Warteliste:<br/>Ca. 20 Wochen gewartet; in dieser Zeit einmal im Monat für &lt;5 min telefonisch kontaktiert, um Kontakt zu bewahren<br/>Dann gleiche Behandlung wie bei EFT</p> |
| <b>3.4 (13) Evtl. weitere Kontroll-<br/>gruppen<br/>(Art, Dauer, Frequenz)</b> | /                                                                                                                                                                                   |

|                | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                          | Rating                                                                              |
|----------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| K15.<br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | <ol style="list-style-type: none"> <li>1) ausführliches Therapiemanual</li> <li>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie</li> <li>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens</li> </ol>                                                                                                                                            | <p style="font-size: 1.2em; color: gray;">B3</p> <p style="font-size: 1.5em;">1</p> |
| K16.<br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung</li> <li>2) Ex post facto Beschreibung der Kontrollbedingungen</li> <li>3) keine Beschreibung der Kontrollbedingung</li> </ol>                                                                                                                                                                                            | <p style="font-size: 1.2em; color: gray;">B4</p> <p style="font-size: 1.5em;">1</p> |
| K17.<br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz</li> <li>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab</li> <li>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab</li> </ol> | <p style="font-size: 1.2em; color: gray;">B5</p> <p style="font-size: 1.5em;">1</p> |
| K18.<br>(B.6.) | Manualtreue, Treatment Integrity                                                     | <ol style="list-style-type: none"> <li>1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt</li> <li>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt</li> <li>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substanzielle Abweichungen</li> </ol>                                                                                                                       | <p style="font-size: 1.2em; color: gray;">B6</p> <p style="font-size: 1.5em;">1</p> |
| K19.<br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | <ol style="list-style-type: none"> <li>1) Intervention wie in klinischer Alltagspraxis</li> <li>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert</li> <li>3) Intervention gegenüber klinischer Alltagspraxis stark verändert</li> </ol>                                                                                                                                                                                | <p style="font-size: 1.2em; color: gray;">C4</p> <p style="font-size: 1.5em;">1</p> |

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

|                       |                                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                     |          |
|-----------------------|---------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
|                       |                                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 1        |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                   | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                               | C6<br>3  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler<br>(insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                  | C7a<br>2 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                           | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln). | C7b<br>2 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                           | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                 | C7c<br>3 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

|                                                                                                                                                                                                                                                                                                                                                                            |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Keine Angaben über Beobachtungszeitraum                                                                                                                                                                                                                                                                                                                                    |
| Messzeitpunkte:                                                                                                                                                                                                                                                                                                                                                            |
| <ul style="list-style-type: none"> <li>• Max. 1 Woche vor Behandlungsbeginn Prämessung, nach Beendigung Postmessung</li> <li>• Warteliste: Prä-wait-Messung am Anfang, und Post-wait-Messung am Ende der Wartezeit (ist gleichzeitig Prä-Behandlung; Post-Messung nach Beendigung der Behandlung)</li> <li>• Follow-up neun Monate nach Beendigung der Therapie</li> </ul> |

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating   |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br>3  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br>1 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br>2 |

**3.6 (15) Anzahl der Zentren**

|   |
|---|
| 1 |
|---|

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

Blinde Randomisierung

**3.8 (17) ggf. Concealment der Randomisierung**

Ja, blind für die Behandler

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating         |
|-----------------------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                         | C8<br><b>3</b> |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br><b>2</b> |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

- IES
- SCL-90-R
- Target Complaints Discomfort Questionnaire (TC)
- Inventory of Interpersonal Problems (IIP)
- Resolution Scale (RS)
- Structural Analysis of Social Behavior, Intrex Questionnaire (SASB)

### 4.2 (21) Sekundäre Zielkriterien

- Prozess-Maße
- Working Alliance Inventory (WAI)
  - EFT-Checklist (Maß zur Überprüfung der Therapie-Adhärenz der Therapeuten in Videoaufzeichnungen)

### 4.3 () Verblindung der Erhebung der Zielkriterien

Ja; Therapeuten bekamen keine Informationen darüber

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | <ol style="list-style-type: none"> <li>1) a priori definiert unter Berücksichtigung der Wertigkeit</li> <li>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien</li> <li>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,</li> </ol>                                                                                                                                                                         | A7<br>2  |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | <ol style="list-style-type: none"> <li>1) reliable und valide Outcome-Verfahren</li> <li>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren</li> <li>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend</li> </ol>                                                                                                                                                                             | A8<br>1  |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | <ol style="list-style-type: none"> <li>1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität</li> <li>2) Zielkriterien beziehen sich nur auf eine Dimension</li> <li>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)</li> </ol>                                                                                                                       | C9<br>1  |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | <ol style="list-style-type: none"> <li>1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen</li> <li>2) zwei Informationsquellen</li> <li>3) eine Informationsquelle</li> </ol>                                                                                                                                                                                                                                                                        | A9<br>3  |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | <ol style="list-style-type: none"> <li>1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler</li> <li>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern</li> <li>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten</li> </ol> | A10<br>9 |

**4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

3 Analysen mit Manova für wiederholte Messungen:  
 1. Vergleich EFT-Gruppe mit Warteliste-Gruppe  
 2. Vergleich EFT-Gruppe mit EFT-Warteliste-Gruppe  
 3. Vergleich alle Probanden (beide Gruppen) im Follow-Up mit Prä-/Postmessung

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>1  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>1 |

**4.5 (22) Ergebnisse**

1. signifikante Unterschieden in allen Zielvariablen bis auf SASB Control Dimension – EFT-Gruppe hat sign. mehr Verbesserungen als Warteliste-Gruppe; mittlerer Unterschied lag bei 1,53 Standardabweichungen
2. sign. Verbesserungen der Warteliste-Gruppe zwischen post-wait- und post-Behandlungs-Messzeitpunkt (außer bei SASB Control Dimension); keine sign. Unterschiede zwischen den beiden Behandlungsgruppen
3. über Prä-, Post- und Follow-up-Messung sign. Unterschiede gehen nur auf Prä-Post-Veränderung zurück – Effekt bleibt nach Follow-up erhalten

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

Keine Angaben

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating              |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br><b>3</b> |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

- Hat für diese heterogene Stichprobe Nachweis der Wirksamkeit von EFT-AS im Gegensatz zur unbehandelten Gruppe gezeigt – hohe externe Validität
- Aber: Unterschied zwischen den beiden Behandlungsgruppen könnten trotzdem vorgelegen haben (z.B. höhere Motivation der ersten EFT-Gruppe)
- Kritik: nur Selbsteinschätzung

|                       | Kriterium              | Operationalisierung                                                                    | Rating             |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br><b>1</b> |



| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>1 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>1 |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

- Design stellt die Vergleichbarkeit der Gruppen nur eingeschränkt sicher
- Verwendung reliabler Messinstrumente

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

- EFT-AS signifikant wirksamer als unbehandelte Gruppe
- auch im 9-Monats-Follow-up bleiben Effekte erhalten

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

- sehr gut
- Überprüfung der internen Reliabilität durch Monitoring der Intervention

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

- Insgesamt ausreichend
- Ein- und Ausschlusskriterien teilweise uneindeutig, dadurch schwer abschätzbar, wie stark die Selektivität der Patienten ist
- Starkes Monitoring

### Stärken der Studie:

- Mehrere Behandler, dadurch kein Personen-Effekt (wurde auch statistisch belegt)
- Randomisierte Zuteilung
- Manualisierte Behandlung

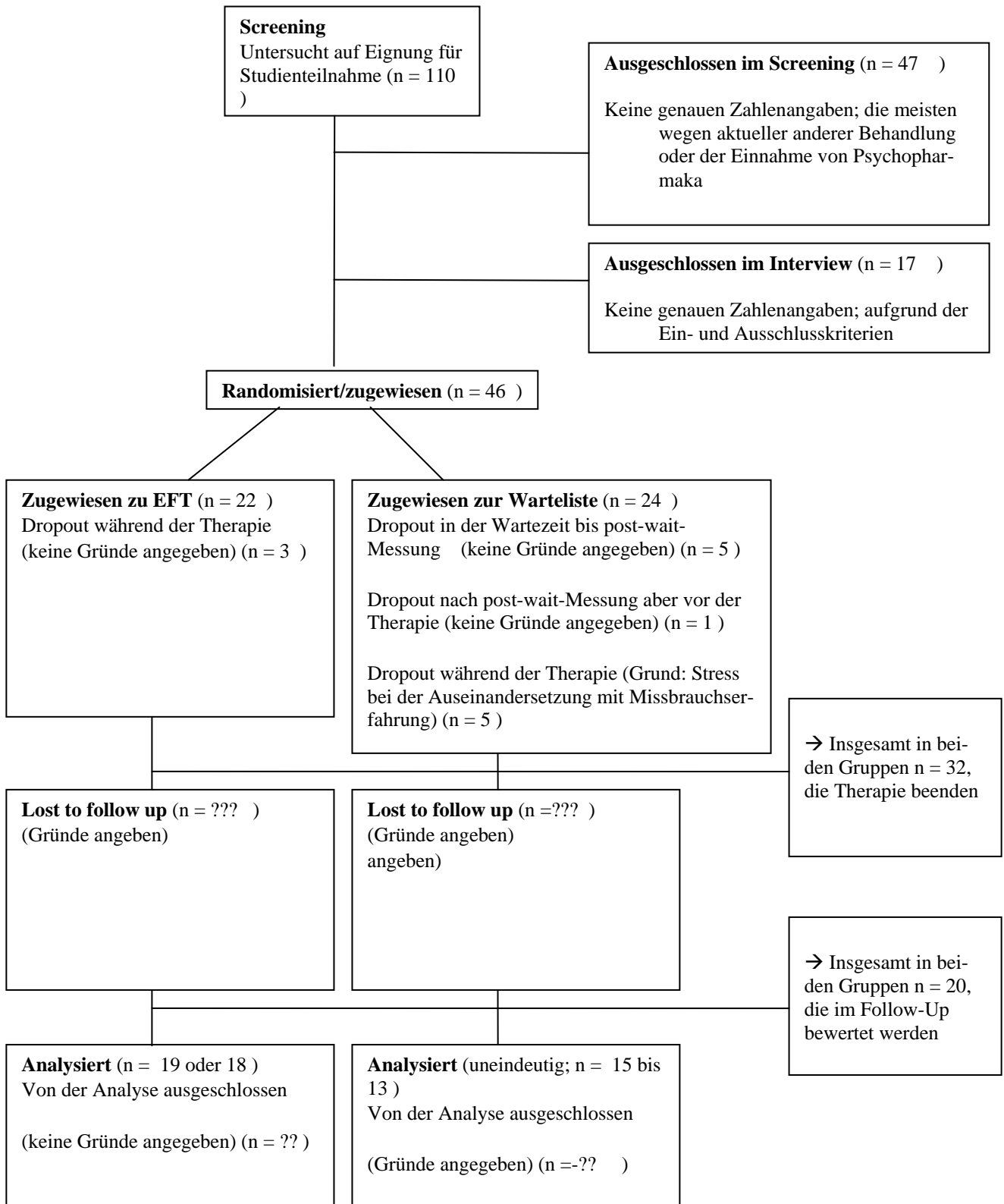
### Schwächen der Studie:

- Nur Selbsteinschätzung
- Am Ende relativ kleine Gruppen (im Follow-up)

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Zeigt, dass für diese Population EFT-AS eine wirksame Behandlungsmethode ist und signifikante Effekte im Vergleich zur unbehandelten Warte-Kontroll-Gruppe bewirkt, welche auch zur Katamnese erhalten bleiben.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

JS

**Studie:**

Autoren:

Sachse

Titel:

Das Bochumer Psychosomatik-Projekt

Quelle:

Gesprächspsychotherapie und Personzentrierte Beratung 2/97

## 1. Basisinformationen

### 1.1 ( ) Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Klärungsorientierte Psychotherapie für psychosomatische Erkrankungen, insbesondere chronisch entzündliche Darm-Erkrankungen (CEDE)

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Überprüfung der Wirksamkeit der Behandlung

### 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.



## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

CEDE, d.h. Ulcus Duodeni, Colitis Ulcerosa und Morbus Crohn

|               | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|---------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K1.<br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>2 |
| K2.<br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>1 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

Keine Angaben

|               | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|---------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K3.<br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>3 |
| K4.<br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>9 |
| K5.<br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>9 |

### 2.3 () Setting/Patientenzugang

Überweisung von Internisten, Anzeigen und Werbung im Radio

|               | Kriterium                           | Operationalisierung                                                                                                                                                                                                                                                                                                                                            | Rating  |
|---------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K6.<br>(C.2.) | Art der Rekrutierung der Stichprobe | 1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege<br>2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege<br>3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung) | C2<br>3 |

**2.4 (7) Zentrale Patientenmerkmale**

(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

1. 18 weibliche, 11 männliche; Erkrankungsdauer im Mittel 4,8 Jahre (Range 10 Monate bis 18 Jahre); sonst keine weiteren Angaben zur Stichprobe
  2. *Patienten der Behandlungsgruppe:*
    - 42 männliche, 46 weibliche Patienten; Alter zwischen 20 und 70 in etwa gleich verteilt über die Jahrgänge
    - 27% haben weitere psychosomatische Beschwerden über die Hauptdiagnose hinaus
    - Chronizität: 38% länger als 10Jahre, 26% seit 5-10 Jahren, 25% 3-5 Jahre und 11% weniger als 3 Jahre
- Gruppe Autogenes Training:*
- 12 weibliche, 8 männliche Patienten; sonst keine Angaben
- Selbsthilfegruppe*
- 14 weibliche, 11 männliche Patienten; sonst keine Angaben

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

Keine Angaben

|                | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating     |
|----------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| K7.<br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br>9/1 |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- Zwei Projekte: 1. Phase: 29 Patienten; 2.Phase: 88 Patienten
- Insgesamt bei 1. 845 Personen, bei 2. 736 Personen, die sich meldeten – es wurden 145 Personen eingeladen, davon erschienen zum Erstgespräch 136
- keine Angaben zu Dropouts

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                | Rating    |
|-----------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| K8.<br>(A.5.)   | Stichprobengröße pro Gruppe                                                | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>2/1 |
| K9.<br>(A.15.)  | Intention to treat – Analysen durchgeführt                                 | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>3  |
| K10.<br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                    | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br>3  |
| K11.            | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Pati-                                                                                                                           | A3        |

|                           |                                                                                 |                                                                                                                                                                                                                                  |             |
|---------------------------|---------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K12.</b><br><br>(A.4.) | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen | A4<br><br>9 |
|---------------------------|---------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|

**2.7 (10) Vergleichbarkeit der Gruppen**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen</p> <p>1. /<br/>2. kann nicht beurteilt werden, da nicht berichtet</p> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)</p> <p>1./<br/>2. kann nicht beurteilt werden, da nicht berichtet</p> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)</p> <p>1./<br/>2. kann nicht beurteilt werden, da nicht berichtet</p> <p>Vergleichbarkeit der Therapieadhärenz</p> <p>1./<br/>2. kann nicht beurteilt werden, da nicht berichtet</p> |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                        | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Rating              |
|------------------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>K13.</b><br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | <p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p> | <p>B9</p> <p>9</p>  |
| <b>K14.</b><br>(A.18.) | Drop-out-Analysen                                                                         | <p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>                                                                                                                       | <p>A18</p> <p>3</p> |

### 3. Intervention

**3.1 (14) Studiendesign /  
Behandlungsarme**

1. eine Behandlungsgruppe
2. drei Gruppen:
  - GT-Gruppe
  - AT-Gruppe
  - Selbsthilfe-Gruppe

**3.2 (11) Behandlung Experimental-  
gruppe  
(Art, Dauer, Frequenz)**

1. erhielten im Schnitt 33,1 Stunden (Range 28-45); keine weiteren Angaben
2. GT-Gruppe: keine Angaben über Dauer und Häufigkeit der Sitzungen

**3.3 (12) Behandlung Kontroll-  
gruppe  
(Art, Dauer, Frequenz)**

1. /
2. AT-Gruppe: keine Angaben zur Behandlung

**3.4 (13) Evtl. weitere Kontroll-  
gruppen  
(Art, Dauer, Frequenz)**

1. /
2. Selbsthilfegruppe: keine Angaben zur Behandlung

|                       | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                          | Rating        |
|-----------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|
| <b>K15.</b><br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | <ol style="list-style-type: none"> <li>1) ausführliches Therapiemanual</li> <li>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie</li> <li>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens</li> </ol>                                                                                                                                            | B3<br><br>1/3 |
| <b>K16.</b><br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung</li> <li>2) Ex post facto Beschreibung der Kontrollbedingungen</li> <li>3) keine Beschreibung der Kontrollbedingung</li> </ol>                                                                                                                                                                                            | B4<br><br>9/3 |
| <b>K17.</b><br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz</li> <li>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab</li> <li>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab</li> </ol> | B5<br><br>9/3 |
| <b>K18.</b><br>(B.6.) | Manualtreue, Treatment Integrity                                                     | <ol style="list-style-type: none"> <li>1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt</li> <li>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt</li> <li>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substanzielle Abweichungen</li> </ol>                                                                                                                       | B6<br><br>3/3 |
| <b>K19.</b><br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | <ol style="list-style-type: none"> <li>1) Intervention wie in klinischer Alltagspraxis</li> <li>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert</li> <li>3) Intervention gegenüber klinischer Alltagspraxis stark verändert</li> </ol>                                                                                                                                                                                | C4<br><br>1   |

|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating         |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br><br>3/3  |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br><br>1/1  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br><br>1/1 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br><br>1/1 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br><br>3/2 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>1.<br/>Keine Angaben über Beobachtungszeitraum<br/>MZPe:<br/> <ul style="list-style-type: none"> <li>• Prä: nach dem Vorgespräch</li> <li>• Zwischenmessung: nach der 10. Sitzung, nach der 20. Sitzung, (nach der 30. Sitzung wenn die Therapien länger dauerten)</li> <li>• Post: nach der letzten Sitzung</li> <li>• Follow-Up: nach 6 Monaten</li> </ul> Für die 2. Untersuchungsphase werden die Messzeitpunkte nicht explizit berichtet, anzunehmen ist, dass es sich um dieselben Festlegungen wie in der ersten Phase handelt</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating         |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br><br>1    |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br><br>1/9 |

|  |  |  |     |
|--|--|--|-----|
|  |  |  | 3/2 |
|--|--|--|-----|

**3.6 (15) Anzahl der Zentren** Keine Angaben

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

1. /  
2. keine Angaben, wie die Patienten in die einzelnen Gruppen kamen

**3.8 (17) ggf. Concealment der Randomisierung**

1. /  
2. /

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating      |
|-----------------------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                         | C8<br><br>9 |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br><br>9 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

In beiden Phasen werden bestimmte Skalen berichtet:

- Selbstakzeptierung – FB von Sorembé und Westhoff
- Internale Kontrolle – IPC von Krampen
- Externale Kontrolle – IPC
- Handlungsorientierung nach Misserfolg und bei der Planung – Hakemp-Fragebogen zur Handlungs-Lage Orientierung)
- Fordern können – FB zur sozialen Unsicherheit
- Nicht-Nein-Sagen-Können – U-FB
- Gedankliche Weiterbeschäftigung, Herunterspielen, Fluchttendenz, Situationskontrolle – Stress-Verarbeitungs-Fragebogen
- Symptomausprägung und Symptombelastung – FB zur Symptombelastung

### 4.2 (21) Sekundäre Zielkriterien

/

### 4.3 () Verblindung der Erhebung der Zielkriterien

/

|                       | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                               | Rating  |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K28.</b><br>(A.7.) | a priori Definition der primären und sekundären Zielkriterien                                                                                | <ol style="list-style-type: none"> <li>1) a priori definiert unter Berücksichtigung der Wertigkeit</li> <li>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien</li> <li>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,</li> </ol>                                                   | A7<br>2 |
| <b>K29.</b><br>(A.8.) | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | <ol style="list-style-type: none"> <li>1) reliable und valide Outcome-Verfahren</li> <li>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren</li> <li>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend</li> </ol>                                                       | A8<br>2 |
| <b>K30.</b><br>(C.9.) | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | <ol style="list-style-type: none"> <li>1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität</li> <li>2) Zielkriterien beziehen sich nur auf eine Dimension</li> <li>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)</li> </ol> | C9<br>1 |
| <b>K31.</b><br>(A.9.) | Multiple Informationsquellen                                                                                                                 | <ol style="list-style-type: none"> <li>1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen</li> <li>2) zwei Informationsquellen</li> <li>3) eine Informationsquelle</li> </ol>                                                                                                                                                  | A9<br>3 |



|  |  |  |   |
|--|--|--|---|
|  |  |  | 9 |
|--|--|--|---|

**4.4 ( ) Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

1. Prä-Post-Vergleich (T-Test) der behandelten Gruppe und Berechnung der Effektstärken – keine Bonferroni-Korrektur berichtet
2. Vergleiche zwischen den Gruppen zu verschiedenen Zeitpunkten

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating          |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br><b>1</b>  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br><b>2</b> |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br><b>2</b> |

**4.5 (22) Ergebnisse**

1. signifikante Verbesserungen für alle berichteten Skalen mit mittleren bis hohen Effektstärken
2.
  - signifikante Verbesserungen für alle berichteten Skalen in der GT-Gruppe mit mittleren bis hohen Effektstärken
  - im Vergleich zu den anderen beiden Kontrollgruppen ist die GT-Gruppe in fast allen Skalen signifikant überlegen
  - die beiden Kontrollgruppen weisen in der Prä-Post-Veränderung nur schwache bis mittlere Effektstärken auf
  - auch Gesundheitskosten-relevante Ergebnisse sprechen für GT – GT-Gruppe weist signifikant weniger Arztbesuche, Medikamenteneinnahme, Krankenhausaufenthalte und berufliche Ausfalltage auf

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating            |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br><b>3*</b>  |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br><b>9/3</b> |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

Keine Angaben

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                    | Rating              |
|------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br><b>3</b> |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Behandlung ist wirksam und zeigt signifikant positive Effekte

|                       | Kriterium              | Operationalisierung                                                                    | Rating              |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|---------------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br><b>3*</b> |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>1 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>1 |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Design wäre prinzipiell für Wirksamkeitsnachweis geeignet. Einschränkungen ergeben sich jedoch daraus dass die Bildung bzw. Rekrutierung der verschiedenen Behandlungsgruppen, die miteinander verglichen werden nicht ausreichend beschrieben wird und die Vergleichbarkeit somit nur eingeschränkt beurteilt werden kann. Ferner werden einige wesentliche methodische Aspekte nicht beleuchtet; die Ergebnisse nicht vollständig präsentiert.

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

es werden nur die Effektstärken und Signifikanzen berichtet – diese zeigen hohe Wirksamkeit von GT im Gegensatz zu den anderen Kontrollgruppen – die Interpretation dieser dargestellten Zusammenhänge ist jedoch vorsichtig zu gestalten, da wesentliche Informationen nicht gegeben werden

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

schwer zu beurteilen, da relevante Aspekte nicht berichtet werden – Vergleichbarkeit der Gruppen kann nicht sicher bewertet werden.

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Selektivität der Patienten ist da, wird aber nicht dokumentiert; Intervention an sich scheint übertragbar zu sein aufs deutsche Versorgungssystem; Behandlerrepräsentativität gewährleistet, da mehrere verschiedene Behandler tätig waren

### Stärken der Studie:

speziell auf CEDE ausgerichtete GT wird auf Effektivität hin untersucht  
verschieden Behandlungsgruppen werden auf Effektivität hin miteinander verglichen

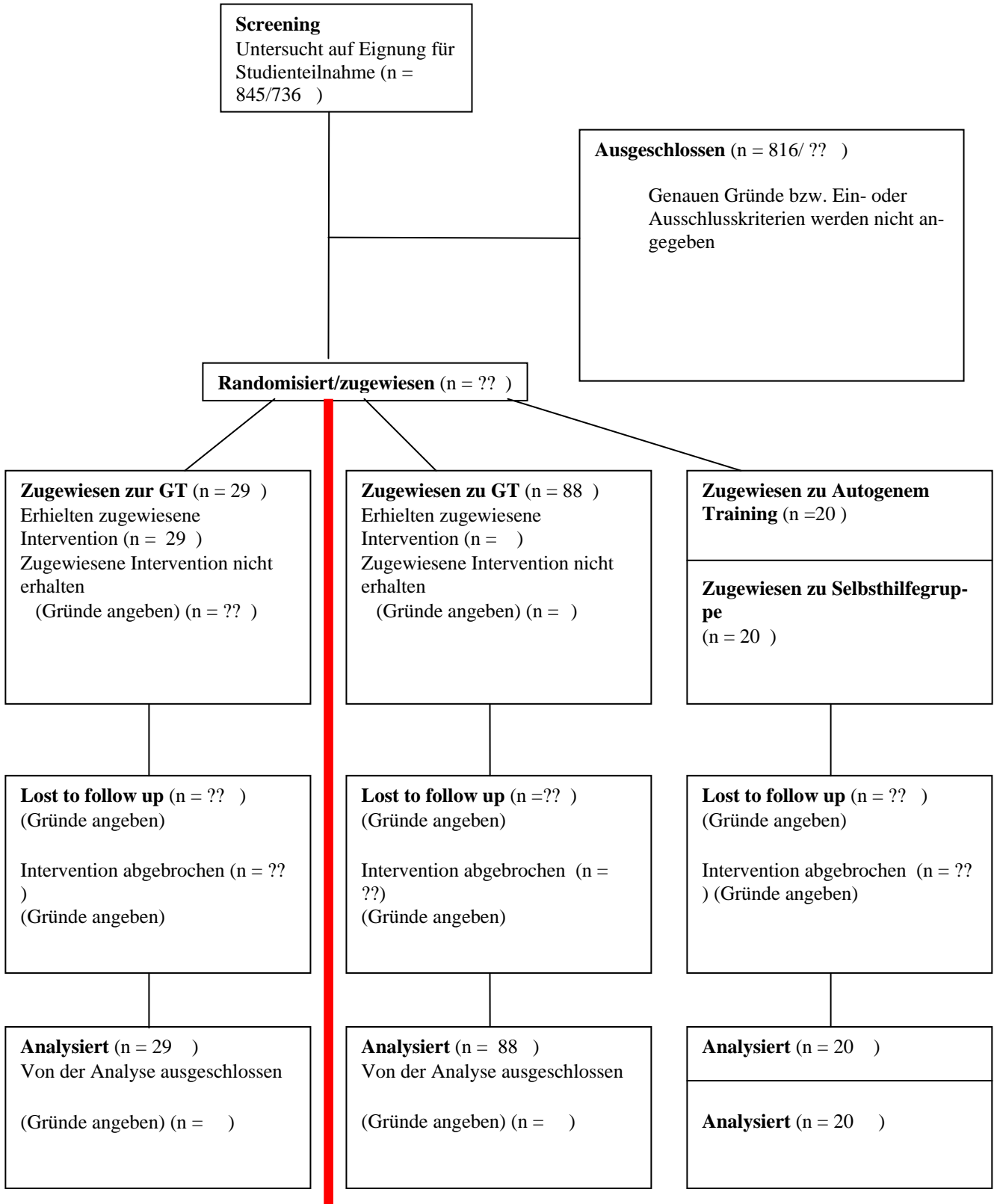
### Schwächen der Studie:

ausgewählte Berichterstattung, Auslassung relevanter Ergebnisse

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Den Ergebnissen dieser Studie zufolge zeigt die Untersuchung eine deutliche Wirksamkeit von GT bei dieser Stichprobe – ggfs. könnte man die Ergebnisse als Nachweis gelten lassen – notwendig dafür wäre jedoch, dass man die im Text nicht berichteten methodischen Aspekte noch mal nachprüfen könnte (wie kommen Gruppen genau zustande? Was ist mit den nicht berichteten Skalen? Wie vergleichbar sind die Gruppen) – somit können die vorliegenden Veröffentlichungen nur als Hinweis auf eine Wirksamkeit verwendet werden.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.



## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

- da hier zwei Phasen von Untersuchungen berichtet werden, ist die Bewertung wie folgt vorgenommen worden:
- in den Bewertungskästchen wurde, wenn verschieden, die Bewertung erst für die erste Phase, dann für die zweite Phase gegeben (z.B. 1/3)
- wenn für beide Phasen das gleiche Urteil gegeben wurde, steht nur eine Zahl da (z.B. 1)
  
- in den grauen Kästchen mit Freitext wird mit „1.“ Immer auf die erste Phase und mit „2.“ Immer auf die zweite Phase verwiesen
- wenn hier keine Unterscheidung vorgenommen wurde, gilt die Aussage für beide Phasen

Zu K36: da hier nur ausgewählte Skalen berichtet werden, wird dieses Kriterium mit „3“ bewertet – anscheinend werden die nicht signifikanten (oder widersprüchlichen?) Ergebnisse nicht präsentiert

Zu K39: Ergebnismanipulation ist dahingehend festzustellen, als dass alle berichteten Zusammenhänge signifikant sind, jedoch nicht berichtet wird, welche Zusammenhänge nicht signifikant sind

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

Kirchmann

**Studie:**

Autoren:

Tarrier, N., Lewis, S., Haddock, G. Bentall, R. et al.

Titel:

Cognitive-behavioural therapy in first-episode and early schizophrenia

Quelle:

British Journal of Psychiatry (2004), 184, 231-239

## 1. Basisinformationen

1.1 () **Benennung der Vorgehensweise**  
(Verfahren/Methode/Interventionstechnik)

Supportive Counseling (zusätzlich zur Routinebehandlung)

1.2 (6) **Fragestellung / Zielsetzung der Studie / Forschungshypothese**

Erweist sich kognitiv-behaviorale Therapie (plus Routinebehandlung) überlegen gegenüber Supportiver Beratung (plus Routinebehandlung) und Routinebehandlung allein bei Patienten mit erst- oder zweimalig auftretender akuter schizophrener Symptomatik?

1.3 (2) **Design**

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) **Evidenzstufe**

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Nicht-Chronifizierte (erst- oder zweimalig auftretende ) akute schizophrene Episode

|                      | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|----------------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K1.</b><br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>2 |
| <b>K2.</b><br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>1 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

Klinische Diagnose einer Schizophrenie, Schizo-affektiven Erkrankung oder Wahnhafte Störung; erst- oder zweimaliges Auftreten/Behandlungersuchen; psychotische Positivsymptome seit mind. 4 Wochen; Wert von mind. 4 auf der Positive and Negative Syndrome Scala (PANSS); Weder Substanzmissbrauch noch organische Faktoren als Hauptursache für psychotische Symptomatik feststellbar

|                      | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|----------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K3.</b><br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>2 |
| <b>K4.</b><br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>2 |
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>2 |

### 2.3 () Setting/Patientenzugang

- Patienten, die in einem der teilnehmenden psychiatrischen Einrichtungen vorstellig wurden.

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

|  |  |  |   |
|--|--|--|---|
|  |  |  | 1 |
|--|--|--|---|

|                                                                                                                                                |                                                                                                                                                                                                                                   |
|------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>2.4 (7) Zentrale Patientenmerkmale</b><br>(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad) | Für die drei Regionen (Liverpool, Manchester, North Nottinghamshire) separat angegeben, daher nachfolgend nur Schätzungen nach Augenschein)<br>Alter: M=27<br>Männlich=70%<br>Bildungsjahre =11<br>Erste psychotische Episode=85% |
|------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                          |                          |
|----------------------------------------------------------|--------------------------|
| <b>2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl</b> | 118 Patienten pro Gruppe |
|----------------------------------------------------------|--------------------------|

|                       | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating   |
|-----------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K7.</b><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br>1 |

|                                                                                                                                                                                               |                                                                                                                                                    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten</b><br>(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm) | Gescreent: 433 Patienten; Einschlußkriterien erfüllt: 364; randomisiert zu Treatmentbedingungen zuordenbar: 309; follow-up Interview-Erhebung: 225 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>1  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>1 |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br>2 |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br>2  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>2  |

**2.7 (10) Vergleichbarkeit der Gruppen**

Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen

Durch Randomisierung gegeben

Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)

Gegeben für die beiden Gruppen mit zusätzlicher psychotherapeutischer Intervention

Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)

Durch Weiterführung der medizinisch-psychiatrischen Routinebehandlung gegeben

Vergleichbarkeit der Therapieadhärenz

Gegeben: angeblich manualisierte Behandlungsformen (keine Angabe spezifischer Literatur, nur Hinweis auf andere Studien: z.B. Tarrier et al., 1998); fortlaufende Supervision; empirisch überprüft durch korrekte Zuordnung zufällig erhobener Audiotapes sowie Einschätzungen auf der Cognitive Therapy Scale of Psychosis

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating       |
|-----------------|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | 1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen<br>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert<br>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt | B9<br><br>1  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | 1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen<br>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt<br>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage                                                                                                                       | A18<br><br>2 |

### 3. Intervention

|                                                                                |                                                                                                                                                                                                                                                                                                                                                  |
|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>3.1 (14) Studiendesign /<br/>Behandlungsarme</b>                            | <ul style="list-style-type: none"> <li>- Supportive Beratung + medizinisch-psychiatrische Routinebehandlung</li> <li>- Kognitiv-behaviorale Therapie + m.-p. Routinebehandlung</li> <li>- medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                                                                |
| <b>3.2 (11) Behandlung Experimental-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Supportive Beratung (emotionale Unterstützung durch eine vertrauliche, wertschätzende Beziehung: „supportive relationship fostering rapport and unconditional regard for the patient“); 15-20 Stunden (tatsächlich aber nur M=7.1h</li> <li>- + medizinisch-psychiatrische Routinebehandlung</li> </ul> |
| <b>3.3 (12) Behandlung Kontroll-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b>     | <ul style="list-style-type: none"> <li>- Kognitiv-behaviorale Therapie (Verbesserung der Coping-Strategien, Problemlösetraining, Erlernen von Strategien zur Rückfallprophylaxe); 15-20 Stunden (tatsächlich aber nur M=8.6h</li> <li>- + medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                |
| <b>3.4 (13) Evtl. weitere Kontroll-<br/>gruppen<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                                                                                                                                                                                                                 |

|                       | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                          | Rating      |
|-----------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K15.</b><br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | <ol style="list-style-type: none"> <li>1) ausführliches Therapiemanual</li> <li>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie</li> <li>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens</li> </ol>                                                                                                                                            | B3<br><br>2 |
| <b>K16.</b><br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung</li> <li>2) Ex post facto Beschreibung der Kontrollbedingungen</li> <li>3) keine Beschreibung der Kontrollbedingung</li> </ol>                                                                                                                                                                                            | B4<br><br>1 |
| <b>K17.</b><br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz</li> <li>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab</li> <li>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab</li> </ol> | B5<br><br>1 |
| <b>K18.</b><br>(B.6.) | Manualtreue, Treatment Integrity                                                     | <ol style="list-style-type: none"> <li>1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt</li> <li>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt</li> <li>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substanzielle Abweichungen</li> </ol>                                                                                                                       | B6<br><br>1 |
| <b>K19.</b><br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | <ol style="list-style-type: none"> <li>1) Intervention wie in klinischer Alltagspraxis</li> <li>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert</li> <li>3) Intervention gegenüber klinischer Alltagspraxis stark verändert</li> </ol>                                                                                                                                                                                | C4<br><br>2 |

|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating       |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br><br>2  |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br><br>1  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br><br>1 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br><br>1 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br><br>1 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

- Pre-, Follow-Up (18 Monate nach Behandlungsbeginn)

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating       |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br><br>2  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br><br>1 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br><br>1 |

**3.6 (15) Anzahl der Zentren**

11 unabhängige Einrichtungen in 3 Regionen

**3.7 (16) Allokation zu Interventionsgruppen (Randomisierung,**

randomisierte Zuteilung



self-selection, ...)

**3.8 (17) ggf. Concealment der Randomisierung**

Gegeben; Fremdeinschätzer waren blind für die Gruppenzugehörigkeit der Patienten; Überprüfung durch Einschätzungen der Fremdeinschätzer bzgl. Der Treatmentzugehörigkeit der Patienten (keine signifikanten Übereinstimmungen mit tatsächlichen Zugehörigkeiten)

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                              | Rating  |
|-----------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br>3 |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br>1 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

Fremdeinschätzungsverfahren

- Positive and Negative Symptom Scale (PANSS);  
Skalen: Positivsymptome; Negativsymptome; Allgemeine Symptome
- Psychotic Symptom Rating Scales (PSYRATS)  
Skalen: Wahnvorstellungen; Auditorische Halluzinationen

### 4.2 (21) Sekundäre Zielkriterien

Analyse der medizinische Dokumentation/Patientenakten

- Häufigkeit eines Relapse (Verschlimmerung psychotischer Symptome, die länger als 1 Woche andauert und zu einer Behandlungsintensivierung führt)
- Häufigkeit einer Re-Hospitalisierung

### 4.3 () Verblindung der Erhebung der Zielkriterien

- gegeben durch Geheimhaltung der Treatmentzuweisung vor den Fremdeinschätzern

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>2  |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | A8<br>1  |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | C9<br>1  |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | A9<br>2  |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10<br>1 |

**4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

Kovarianzanalyse mit a priori Kontrasten; (logistische Regression; Survivalanalyse werden erwähnt, entsprechende Parameter aber nicht dargestellt)

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>1  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>2 |

**4.5 (22) Ergebnisse**

Die Treatmentgruppen (SC, CBT) zeigten hinsichtlich aller Skalen der PANSS signifikant stärkere Verbesserungen im Vergleich zur Routinebehandlungsgruppe, nicht aber hinsichtlich der PSYRATS-Skalen sowie der Re-Hospitalisierungs- und Relapserate.

Analysen für die einzelnen Regionen ergaben signifikant stärkere Verbesserungen für die Treatmentgruppen lediglich für eine (Liverpool) der drei Regionen

In keiner der Outcome-Variablen unterschieden sich die beiden Treatmentgruppen (SC, CBT) signifikant.

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>2 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

- keine Angaben/Analysen

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                    | Rating       |
|------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br>2 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Kognitiv-behaviorale Therapie ist ein effektiver Zusatz zur routinemäßigen psychiatrischen Behandlung bei schizophrener Positivsymptomatik

|                       | Kriterium              | Operationalisierung                                                                    | Rating      |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|-------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br>1 |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |             |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1          |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2          |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3<br><br>1 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating       |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br><br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br><br>1 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br><br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br><br>1 |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Befriedigend: umfangreiches Screening und Auswahl von Patienten; klar definierte Einschlusskriterien; Intention-to-treat-Analysen durchgeführt; aber zu knappe Darstellung insbesondere des Ergebnisteils, um Adäquatheit des statistischen Prozederes nachvollziehen zu können

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Die vereinigten Behandlungsgruppen (Supportive Beratung kombiniert mit psychopharmakol. Behandl. Und Kognitiv-behaviorale Therapie kombiniert mit psychopharmakol. Behandlung) zeigten sich gegenüber der Treatment-As-Usual-Gruppe (nur psychopharmakolog. Behandlung) zur 18-Monats-Katamnese in allen drei Skalen der Positive and Negative Syndrome Scale (PANSS) signifikant stärker verbessert, nicht aber hinsichtlich der beiden Skalen Wahnhafte Überzeugungen und Auditive Haluzinationen der Psychotic Symptom Rating Scales (PSYRATS) sowie der Relaps- und Re-Hospitalisierungsrate.

Die beiden Behandlungsgruppen unterschieden sich nicht signifikant.

Die Analyse von Regionunterschieden erbrachte, dass signifikante Unterschiede im Behandlungserfolg nur für eine (Liverpool) der drei Regionen bestanden.

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Insgesamt gut.

Einschränkend wirkt, dass die Interpretation der unterschiedlichen Ergebnisse zwischen den verschiedenen Regionen unklar ist: Weshalb zeigten sich nur bei einer der drei Regionen die Behandlungsgruppen der Routinebehandlung überlegen? Die verschiedenen psychologischen Behandlungen wurden in den einzelnen Regionen von denselben Therapeuten (über die drei Regionen insgesamt nur 5) durchgeführt; in den einzelnen Regionen waren jeweils unterschiedliche Fremdeinschätzer tätig; die Ausgangssymptomatik zwischen den Regionen war unterschiedlich; die Behandlungsdauer war zwischen den Regionen unterschiedlich.

Sind die „regionalen“ Unterschiede auf Therapeutenspezifika, unterschiedliche Tendenzen in den Fremdeinschätzungen, Unterschiede in der tatsächlichen Symptomschwere (Ausgangsniveau) oder Unterschieden in der Behandlungsdauer zurück zu führen?

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

### Weitgehend gegeben.

Leichte Einschränkungen ergeben sich daraus, dass in dieser Studie nur Patienten mit erst- bzw. zweimalig auftretender schizophrener Episode eingeschlossen wurden und keine Patienten mit chronifizierten Störungen. Ferner wurde die Supportive Beratung von ausgebildeten Verhaltenstherapeuten durchgeführt. Durchschnittlich lediglich 8 Stunden Therapiedauer (7.1 Support. Beratung; 8.6 CBT)

### Stärken der Studie:

Große Stichprobe, randomisierte Zuteilung, Concealment

### Schwächen der Studie:

Nur 5 Therapeuten, die beide Psychotherapiearme abdecken; lediglich regional arbeitende Fremdeinschätzer, regional unterschiedliche Behandlungsdauer.

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

#### Hinweis auf Wirksamkeit der klientenzentrierten Intervention (GPZ II).

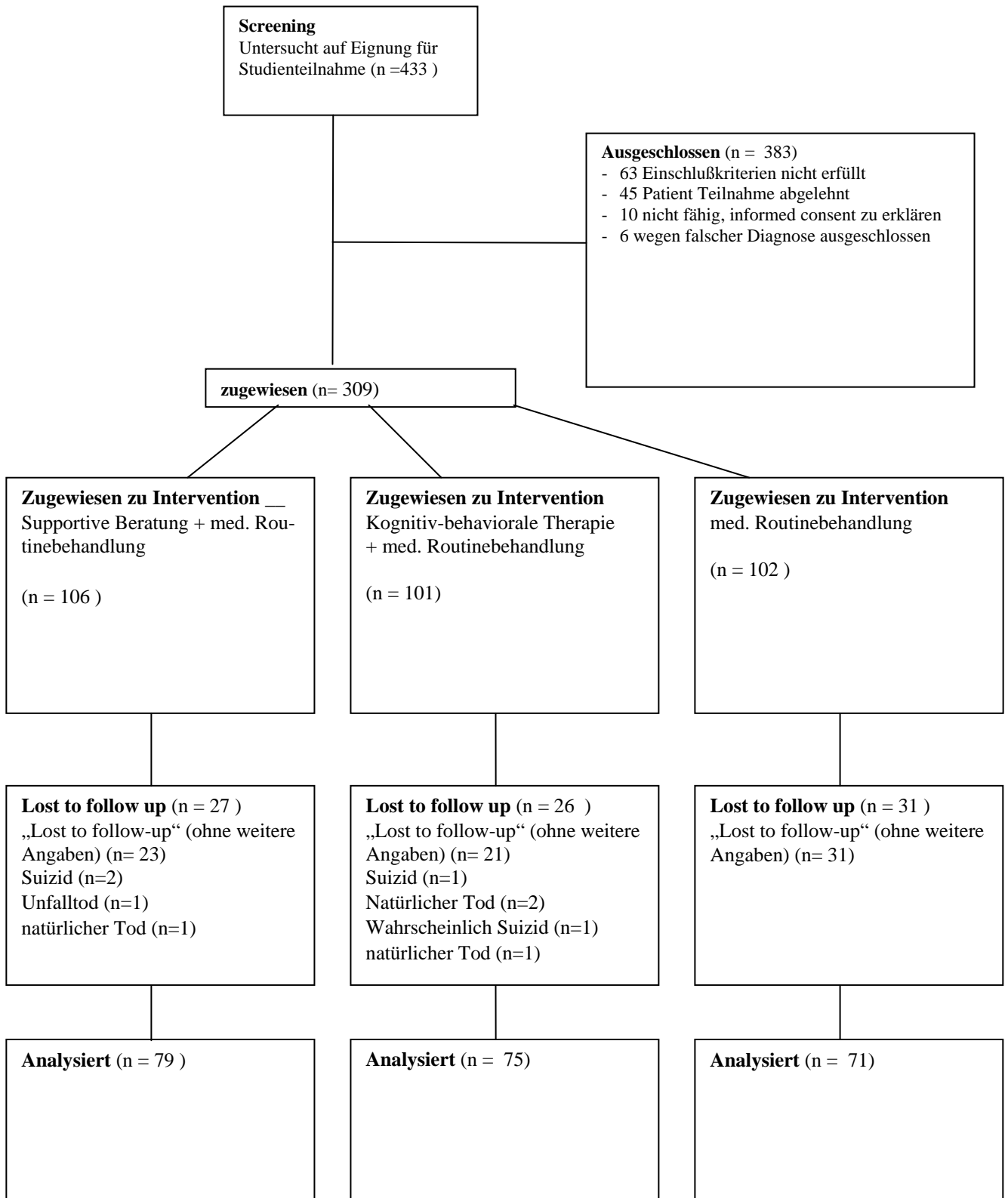
Leichte Einschränkungen ergeben sich aus den regionalen Unterschiede (nur in einer der drei Regionen waren die psycholog. Behandlungen effektiver als die Routinebehandlung), welches die Interpretation der Befunde erschwert.

Deutliche Effekte zeigen sich in einem Fremdeinschätzungsverfahren, hier gibt es für die GPT II signifikant stärkere Symptomverbesserungen. Für die dichotomen Outcome-Kriterien Relapse und Re-

Hospitalisierungsraten lassen sich keine Effekte nachweisen. Hierfür war vermutlich die statistische Power zu gering. Für die Wirksamkeit der klientenzentrierten Intervention spricht zudem, dass der klientenzentrierte Ansatz im Vergleich zur CBT nicht unterlegen ist, obwohl die Autoren (und Therapeuten!) Vertreter der CBT sind.



**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

Stand: 15.05.2007

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

HK

**Studie:**

Autoren:

Tarrier, N., Kinney, C., McCarthy, E, Humphreys, L., Wittkowski, A. & Morris, J.

Titel:

Two-year-follow-up of cognitive-behavioral therapy and supportive counseling in the treatment of persistent symptoms in chronic schizophrenia

Quelle:

Journal of Consulting and Clinical Psychology (2000), 68, 917-922

## 1. Basisinformationen

1.1 ( ) **Benennung der Vorgehensweise**  
(Verfahren/Methode/Interventionstechnik)

Supportive Counseling (zusätzlich zur Routinebehandlung)

1.2 (6) **Fragestellung / Zielsetzung der Studie / Forschungshypothese**

Erweist sich kognitiv-behaviorale Therapie (plus Routinebehandlung) überlegen gegenüber Supportiver Beratung (plus Routinebehandlung) und Routinebehandlung allein.

1.3 (2) **Design**

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) **Evidenzstufe**

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Chronisch psychotische Störungen (akute Positivsymptomatik): Diagnose einer Schizophrenie, Schizoaffektiven Erkrankung oder Wahnstörung durch psychiatrisches Interview

|               | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|---------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K1.<br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>2 |
| K2.<br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>1 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

Klinische Diagnose einer Schizophrenie, Schizo-affektiven Erkrankung oder Wahnhafte Störung; Zustimmung zur Teilnahme durch Patienten (informed consent), keine Evidenz für organische Hirnerkrankung; Substanzabhängigkeit nicht identifiziert als primäres Problem; Alter zwischen 18 und 65 Jahren; aktuelle persistente Halluzinationen oder Wahninhalte seit mindestens 6 Monaten; gleichbleibende Medikation; keine psychologische oder psychosoziale Intervention

|               | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|---------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K3.<br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>1 |
| K4.<br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>2 |
| K5.<br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>3 |

### 2.3 () Setting/Patientenzugang

- Patienten, die vom britischen staatlichen Gesundheitssystem (National Health Service) in einer Datei geführt wurden.

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

|  |  |  |   |
|--|--|--|---|
|  |  |  | 2 |
|--|--|--|---|

**2.4 (7) Zentrale Patientenmerkmale**  
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

Keine Angaben (ähnlich wie in Tarrier et al., 1998)

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

28 Patienten pro Gruppe

|                       | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating   |
|-----------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K7.</b><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br>2 |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

Gescreent: 470 Patienten; Einschlußkriterien erfüllt: 138; randomisiert zu Treatmentbedingungen zuordenbar: 87; in Datenanalysen eingeschlossen: 61

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>2  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>2 |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br>1 |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br>2  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>2  |



**2.7 (10) Vergleichbarkeit der Gruppen**

Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen

Durch Randomisierung gegeben

Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)

Gegeben für die beiden Gruppen mit zusätzlicher psychotherapeutischer Intervention

Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)

Durch Weiterführung der medizinisch-psychiatrischen Routinebehandlung gegeben

Vergleichbarkeit der Therapieadhärenz

Gegeben: überprüft durch korrekte Zuordnung zufällig erhobener Audiotapes

|                        | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating       |
|------------------------|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K13.</b><br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | 1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen<br>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert<br>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt | B9<br><br>1  |
| <b>K14.</b><br>(A.18.) | Drop-out-Analysen                                                                         | 1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen<br>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt<br>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage                                                                                                                       | A18<br><br>2 |

### 3. Intervention

|                                                                                |                                                                                                                                                                                                                                                                                                                                            |
|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>3.1 (14) Studiendesign /<br/>Behandlungsarme</b>                            | <ul style="list-style-type: none"> <li>- Supportive Beratung + medizinisch-psychiatrische Routinebehandlung</li> <li>- Kognitiv-behaviorale Therapie + Routinebehandlung</li> <li>- medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                                                                |
| <b>3.2 (11) Behandlung Experimental-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Supportive Beratung (emotionale Unterstützung durch eine vertrauliche, wertschätzende Beziehung: „relationship fostering rapport and unconditional regard for the patient“); 20 Sitzungen, je eine Stunde, 2 mal wöchentlich</li> <li>- + medizinisch-psychiatrische Routinebehandlung</li> </ul> |
| <b>3.3 (12) Behandlung Kontroll-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b>     | <ul style="list-style-type: none"> <li>- Kognitiv-behaviorale Therapie (Verbesserung der Coping-Strategien, Problemlösetraining, Erlernen von Strategien zur Rückfallprophylaxe); 20 Sitzungen, je eine Stunde, 2 mal wöchentlich</li> <li>- + medizinisch-psychiatrische Routinebehandlung</li> </ul>                                     |
| <b>3.4 (13) Evtl. weitere Kontroll-<br/>gruppen<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                                                                                                                                                                                                           |

|                       | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                          | Rating      |
|-----------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K15.</b><br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | <ol style="list-style-type: none"> <li>1) ausführliches Therapiemanual</li> <li>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie</li> <li>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens</li> </ol>                                                                                                                                            | B3<br><br>2 |
| <b>K16.</b><br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung</li> <li>2) Ex post facto Beschreibung der Kontrollbedingungen</li> <li>3) keine Beschreibung der Kontrollbedingung</li> </ol>                                                                                                                                                                                            | B4<br><br>1 |
| <b>K17.</b><br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz</li> <li>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab</li> <li>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab</li> </ol> | B5<br><br>1 |
| <b>K18.</b><br>(B.6.) | Manualtreue, Treatment Integrity                                                     | <ol style="list-style-type: none"> <li>1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt</li> <li>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt</li> <li>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substanzielle Abweichungen</li> </ol>                                                                                                                       | B6<br><br>1 |
| <b>K19.</b><br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | <ol style="list-style-type: none"> <li>1) Intervention wie in klinischer Alltagspraxis</li> <li>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert</li> <li>3) Intervention gegenüber klinischer Alltagspraxis stark verändert</li> </ol>                                                                                                                                                                                | C4<br><br>2 |

|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Rating       |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br><br>2  |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                               | C6<br><br>1  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                  | C7a<br><br>1 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                 | C7b<br><br>1 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                 | C7c<br><br>1 |

**3.5 (19) Beobachtungszeitraum/ Messzeitpunkte/Katamnese**

- Pre-, Post-Erhebung; (12 Wochen), 1-Jahres-katamnese; 2-Jahres-Katamnese

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating       |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br><br>2  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br><br>2 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br><br>2 |

**3.6 (15) Anzahl der Zentren**

1

**3.7 (16) Allokation zu Interventions-**

randomisierte Zuteilung

gruppen (Randomisierung, self-selection, ...)



**3.8 (17) ggf. Concealment der Randomisierung**

Nicht gegeben; Fremdeinschätzer waren aber blind für die Gruppenzugehörigkeit der Patienten (Verwendung zweier getrennter Bürokomplexe)

|                | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                              | Rating  |
|----------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K26.<br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br>3 |
| K27.<br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br>2 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

- Present State Examination
- Brief Psychiatric Rating Scale

### 4.2 (21) Sekundäre Zielkriterien

- Relapse (Krankenhausaufenthalt von mind. 5 Tagen Dauer)
- Medikamentendosis
- Beck Depressions-Inventar
- Beck Hoffnungslosigkeitsskala
- Social Functioning Scale

### 4.3 () Verblindung der Erhebung der Zielkriterien

- gegeben durch Geheimhaltung der Treatmentzuweisung vor den Fremdeinschätzern

|                        | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K28.</b><br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | <ol style="list-style-type: none"> <li>1) a priori definiert unter Berücksichtigung der Wertigkeit</li> <li>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien</li> <li>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,</li> </ol>                                                                                                                                                                         | A7<br>2  |
| <b>K29.</b><br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | <ol style="list-style-type: none"> <li>1) reliable und valide Outcome-Verfahren</li> <li>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren</li> <li>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend</li> </ol>                                                                                                                                                                             | A8<br>1  |
| <b>K30.</b><br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | <ol style="list-style-type: none"> <li>1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität</li> <li>2) Zielkriterien beziehen sich nur auf eine Dimension</li> <li>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)</li> </ol>                                                                                                                       | C9<br>1  |
| <b>K31.</b><br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | <ol style="list-style-type: none"> <li>1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen</li> <li>2) zwei Informationsquellen</li> <li>3) eine Informationsquelle</li> </ol>                                                                                                                                                                                                                                                                        | A9<br>1  |
| <b>K32.</b><br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | <ol style="list-style-type: none"> <li>1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler</li> <li>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern</li> <li>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten</li> </ol> | A10<br>1 |

**4.4 ( ) Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

Kovarianzanalyse über aggregierten Follow-up-Zeitraum (3-Mon. + 1-Jahr + 2-Jahres-Katamnese); Häufigkeitsanalyse; Survivalkurve

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>2  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>2 |

**4.5 (22) Ergebnisse**

Über den 2Jahres-Katamnesenzeitraum hinweg zeigten sich sowohl supportive Beratung als auch kognitiv-behaviorale Therapie der medizinisch-psychiatrischen Routinebehandlung überlegen (hinsichtlich Positiv- sowie Negativsymptomatik), wohingegen sich die beiden psychotherapeutischen Interventionen nicht signifikant voneinander unterschieden.

Keine Unterschiede zwischen den Behandlungsgruppen fanden sich dagegen in Bezug auf die Häufigkeit von Relapse und die Medikamentendosierung.

Hinsichtlich der Unterschiede von Prä nach 2-Jahres-Katamnese in Hinblick auf BDI, BHI und SFS zeigte die supportive Beratung die höchsten Effektstärken im Vergleich zu kognitiv-behavioraler Therapie und mehr noch der medizinischen Routinebehandlung. Diese Überlegenheit der supportiven Beratung zum Zeitpunkt der 2-Jahres-Katamnese zeigte sich augenscheinlich auch in Bezug auf die schizophrene Positiv- und Negativsymptomatik

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>3 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>3 |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

- keine Angaben/Analysen

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                    | Rating       |
|------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br>3 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Kognitiv-behaviorale Therapie ist ein effektiver Zusatz zur routinemäßigen psychiatrischen Behandlung bei schizophrener Positivsymptomatik

|                       | Kriterium              | Operationalisierung                                                                    | Rating      |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|-------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br>1 |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |         |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1<br>1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2<br>1 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3<br>3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>3 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>1 |



## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

unbefriedigend: zwar hoher Aufwand bzgl. Der Stichprobenrekrutierung aber geringe Gruppengrößen, undifferenzierte Datenanalysen, unklare Ergebnisinterpretationsmöglichkeiten

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Über den 2-Jahres-Katamnesenzeitraum hinweg (3-Monats + 1-Jahres + 2-Jahreskatamnese) zeigten sich sowohl supportive Beratung als auch kognitiv-behaviorale Therapie der medizinisch-psychiatrischen Routinebehandlung überlegen (hinsichtlich Positiv- sowie Negativsymptomatik), wohingegen sich die beiden psychotherapeutischen Interventionen nicht signifikant voneinander unterschieden.

Keine Unterschiede zwischen den Behandlungsgruppen fanden sich dagegen in Bezug auf die Häufigkeit von Rückfällen und die Medikamentendosierung.

Hinsichtlich der Unterschiede von Prä nach 2-Jahres-Katamnese in Hinblick auf BDI, BHI und SFS zeigte die supportive Beratung die höchsten Effektstärken im Vergleich zu kognitiv-behavioraler Therapie und mehr noch zur medizinischen Routinebehandlung. Diese Überlegenheit der supportiven Beratung zum Zeitpunkt der 2-Jahres-Katamnese zeigte sich augenscheinlich (nach den Verlaufsgrafiken) auch in Bezug auf die schizophrene Positiv- und Negativsymptomatik

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

**Einschränkung:** Verbesserung der Gruppe der supportiven Beratung zur 2-Jahres-Katamnese kann nicht eindeutig auf das Treatment zurückgeführt werden

- Verbesserungen finden vor allem zwischen der 1-Jahres-Katamnese und der 2-Jahres-Katamnese statt
- Kleine Gruppen, fehlende Drop-out-Analysen

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

**Eingeschränkt:** relativ hohe Selektivität der Patienten (nur 18,5% der gescreenten Patienten gingen in die Randomisierung und 13% in die Follow-up-Auswertungen ein: Einschränkung auf psychotherapiebereite, medikamentenstabile chronisch Schizophrene)

### Stärken der Studie:

Umfassende Stichprobenerhebung und –beschreibung;

multimodale Erfassung von Outcome-Kriterien (Fremdeinschätzungen, Selbsteinschätzungen, Informationen aus medizinischer Dokumentation: (Krankenhausaufenthaltsstage, Medikamentendosierung)

### Schwächen der Studie:

Nur 3 Therapeuten, die beide Psychotherapiearme abdecken: unklar ob Unterschiede der Behandlungseffektivität auf Therapeutenmerkmale oder Behandlungskonzepte zurückzuführen sind.

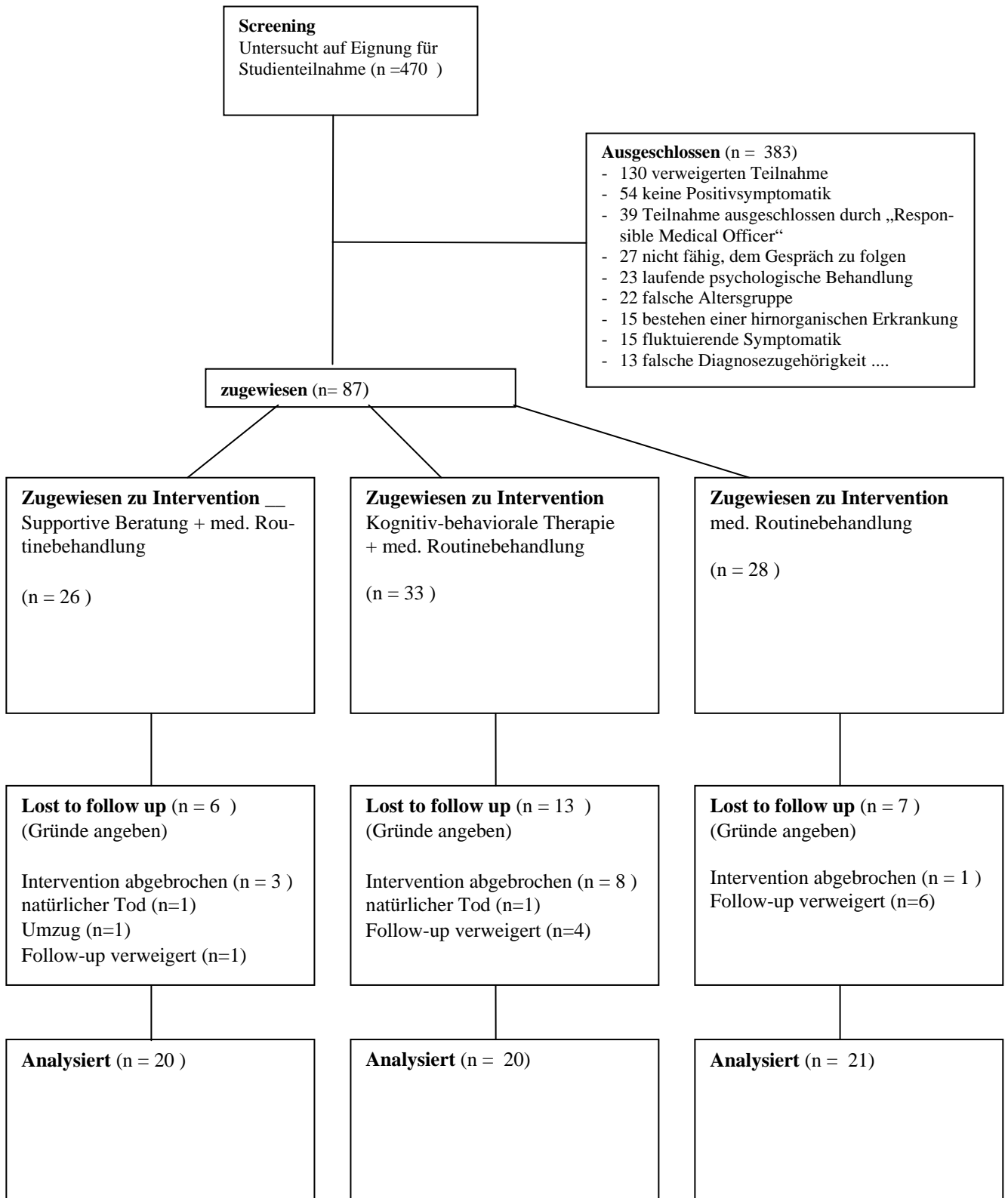
Relativ kleine Gruppengrößen zur 2-Jahreskatamnese (20, 20, 21)

Wenig differenzierte Datenanalysen (Summierung der Follow-up-Zeitpunkte, entweder nur deskriptive oder nur Inferenzstatistik).

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

**Eingeschränkt relevant:** supportive Beratung (GPT II) zeigt sich zur 2-Jahres-Katamnese sowohl der Routinebehandlung als auch in einigen sekundären Zielkriterien der kognitiv-behavioralen Therapie überlegen, doch diese Effekte scheinen sich erst nach der 1-Jahres-Katamnese einzustellen und können daher nicht eindeutig auf das Treatment zurückgeführt werden. Zudem sind die Gruppen recht klein (n=20) und es wurden keine Drop-out-Analysen (immerhin 30%) vorgenommen)

## 8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

HK

**Studie:**

Autoren:

Tarrier, N., Yusupoff, L., Kinney, C., McCarthy, E, Gledhill, A., Haddock, G. & Morris, J.

Titel:

Randomised controlled trial of intensive cognitive behaviour therapy for patients with chronic schizophrenia

Quelle:

British Medical Journal (1998), 317, 303-307

## 1. Basisinformationen

1.1 ( ) **Benennung der Vorgehensweise**  
(Verfahren/Methode/Interventionstechnik)

Supportive Counseling (zusätzlich zur Routinebehandlung)

1.2 (6) **Fragestellung / Zielsetzung der Studie / Forschungshypothese**

Erweist sich kognitiv-behaviorale Therapie (plus Routinebehandlung) überlegen gegenüber Supportiver Beratung (plus Routinebehandlung) und Routinebehandlung allein.

1.3 (2) **Design**

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

1.4 (3) **Evidenzstufe**

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.



## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Chronisch psychotische Störungen (akute Positivsymptomatik): Diagnose einer Schizophrenie, Schizoaffektiven Erkrankung oder Wahnstörung durch psychiatrisches Interview

|               | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|---------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K1.<br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>2 |
| K2.<br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>1 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

Klinische Diagnose einer Schizophrenie, Schizo-affektiven Erkrankung oder Wahnhafte Störung; Zustimmung zur Teilnahme durch Patienten (informed consent), keine Evidenz für organische Hirnerkrankung; Substanzabhängigkeit nicht identifiziert als primäres Problem; Alter zwischen 18 und 65 Jahren; aktuelle persistente Halluzinationen oder Wahninhalte seit mindestens 6 Monaten; gleichbleibende Medikation; keine psychologische oder psychosoziale Intervention

|               | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|---------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K3.<br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>1 |
| K4.<br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>2 |
| K5.<br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>3 |

### 2.3 () Setting/Patientenzugang

- Patienten, die vom britischen staatlichen Gesundheitssystem (National Health Service) in einer Datei geführt wurden.

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

|  |  |  |          |
|--|--|--|----------|
|  |  |  | <b>2</b> |
|--|--|--|----------|

|                                                                                                                                                |                                                                                                                                  |
|------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| <b>2.4 (7) Zentrale Patientenmerkmale</b><br>(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad) | Alter: M=38.6, SD=11.08.07<br>Männlich=69%<br>Singles=74%<br>Arbeitslos=87%<br>Median der Erkrankungsdauer=11 Jahre (Range=1-42) |
|------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|

|                                                          |                         |
|----------------------------------------------------------|-------------------------|
| <b>2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl</b> | 28 Patienten pro Gruppe |
|----------------------------------------------------------|-------------------------|

|                       | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating          |
|-----------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <b>K7.</b><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br><b>1</b> |

|                                                                                                                                                                                               |                                                                                                                                                      |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten</b><br>(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm) | Gescreent: 470 Patienten; Einschlusskriterien erfüllt: 138; randomisiert zu Treatmentbedingungen zuordenbar: 87; in Datenanalysen eingeschlossen: 79 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating          |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br><b>2</b>  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br><b>1</b> |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br><b>1</b> |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br><b>2</b>  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4              |

**2.7 (10) Vergleichbarkeit der Gruppen**

Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen

Durch Randomisierung gegeben

Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)

Gegeben für die beiden Gruppen mit zusätzlicher psychotherapeutischer Intervention

Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)

Durch Weiterführung der medizinisch-psychiatrischen Routinebehandlung gegeben

Vergleichbarkeit der Therapieadhärenz

Gegeben: überprüft durch korrekte Zuordnung zufällig erhobener Audiotapes

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating       |
|-----------------|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | 1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen<br>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert<br>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt | B9<br><br>1  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | 1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen<br>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt<br>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage                                                                                                                       | A18<br><br>2 |

### 3. Intervention

|                                                                                |                                                                                                                                                                                                                                                                                                                                            |
|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>3.1 (14) Studiendesign /<br/>Behandlungsarme</b>                            | <ul style="list-style-type: none"> <li>- Supportive Beratung + medizinisch-psychiatrische Routinebehandlung</li> <li>- Kognitiv-behaviorale Therapie + Routinebehandlung</li> <li>- medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                                                                |
| <b>3.2 (11) Behandlung Experimental-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Supportive Beratung (emotionale Unterstützung durch eine vertrauliche, wertschätzende Beziehung: „relationship fostering rapport and unconditional regard for the patient“); 20 Sitzungen, je eine Stunde, 2 mal wöchentlich</li> <li>- + medizinisch-psychiatrische Routinebehandlung</li> </ul> |
| <b>3.3 (12) Behandlung Kontroll-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b>     | <ul style="list-style-type: none"> <li>- Kognitiv-behaviorale Therapie (Verbesserung der Coping-Strategien, Problemlösetraining, Erlernen von Strategien zur Rückfallprophylaxe); 20 Sitzungen, je eine Stunde, 2 mal wöchentlich</li> <li>- + medizinisch-psychiatrische Routinebehandlung</li> </ul>                                     |
| <b>3.4 (13) Evtl. weitere Kontroll-<br/>gruppen<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                                                                                                                                                                                                           |

|                       | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                          | Rating      |
|-----------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K15.</b><br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | <ol style="list-style-type: none"> <li>1) ausführliches Therapiemanual</li> <li>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie</li> <li>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens</li> </ol>                                                                                                                                            | B3<br><br>2 |
| <b>K16.</b><br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung</li> <li>2) Ex post facto Beschreibung der Kontrollbedingungen</li> <li>3) keine Beschreibung der Kontrollbedingung</li> </ol>                                                                                                                                                                                            | B4<br><br>1 |
| <b>K17.</b><br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz</li> <li>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab</li> <li>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab</li> </ol> | B5<br><br>1 |
| <b>K18.</b><br>(B.6.) | Manualtreue, Treatment Integrity                                                     | <ol style="list-style-type: none"> <li>1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt</li> <li>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt</li> <li>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substanzielle Abweichungen</li> </ol>                                                                                                                       | B6<br><br>1 |
| <b>K19.</b><br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | <ol style="list-style-type: none"> <li>1) Intervention wie in klinischer Alltagspraxis</li> <li>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert</li> <li>3) Intervention gegenüber klinischer Alltagspraxis stark verändert</li> </ol>                                                                                                                                                                                | C4<br><br>2 |

|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating       |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br><br>2  |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br><br>1  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br><br>1 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br><br>1 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br><br>1 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

- Pre-, Post-Erhebung; (12 Wochen)

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating       |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br><br>2  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br><br>2 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br><br>3 |

**3.6 (15) Anzahl der Zentren**

1

**3.7 (16) Allokation zu Interventionsgruppen (Randomisierung,**

randomisierte Zuteilung

self-selection, ...)



**3.8 (17) ggf. Concealment der Randomisierung**

Nicht gegeben; Fremdeinschätzer waren aber blind für die Gruppenzugehörigkeit der Patienten (Verwendung zweier getrennter Bürokomplexe)

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                              | Rating  |
|-----------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br>3 |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br>2 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

- Present State Examination
- Brief Psychiatric Rating Scale

### 4.2 (21) Sekundäre Zielkriterien

-

### 4.3 () Verblindung der Erhebung der Zielkriterien

- gegeben durch Geheimhaltung der Treatmentzuweisung vor den Fremdeinschätzern

|                        | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K28.</b><br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>2  |
| <b>K29.</b><br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | A8<br>1  |
| <b>K30.</b><br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | C9<br>2  |
| <b>K31.</b><br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | A9<br>3  |
| <b>K32.</b><br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10<br>1 |

**4.4 ( ) Statistische Methodik  
(Subgruppenanalysen,  
Sensibilitätsanalysen)**

Varianzanalyse; Häufigkeitsanalyse; logistische Regression

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>1  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>2 |

**4.5 (22) Ergebnisse**

Die Treatmentgruppe unterschieden sich signifikant hinsichtlich der Reduktion der Positivsymptomatik, wobei sich signifikante Einzelunterschiede nur zwischen kognitiv-behavioraler und psychiatrischer Routinebehandlung ergaben.

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcomemaßen und zu allen relevanten Messzeitpunkten                                              | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |



**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

- keine Angaben/Analysen

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                    | Rating       |
|------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br>3 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Kognitiv-behaviorale Therapie ist ein effektiver Zusatz zur routinemäßigen psychiatrischen Behandlung bei schizophrener Positivsymptomatik

|                       | Kriterium              | Operationalisierung                                                                    | Rating      |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|-------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br>1 |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |         |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1<br>3 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2<br>3 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3<br>3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>3 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>1 |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Hoher Aufwand, ordentliche Methodik: umfangreiches Screening und Auswahl von Patienten; klar definierte Ein- und Ausschlusskriterien; Intention-to-treat-Analysen durchgeführt

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Die Behandlungsgruppen unterscheiden sich hinsichtlich der Reduktion der Positivsymptomatik statistisch signifikant. Im Einzelvergleich erweist sich aber nur die kognitiv-behaviorale Therapie der psychiatrischen Routinebehandlung als überlegen.

Immerhin verbessern sich auch die Patienten mit supportiver Unterstützung signifikant im Gegensatz zu den ausschließlich Routinebehandelten, allerdings fällt der entsprechende Gruppenvergleich nicht signifikant aus.

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

#### etwas eingeschränkt

gute Vergleichbarkeit der Gruppen und plausible Operationalisierung der Veränderungen aber unklar, ob unterschiedliche Effektivität der kognitiv-behavioralen versus supportiven Therapie wirklich auf unterschiedliche Behandlungskonzepte zurückzuführen ist, da die gesamten Therapien von nur 3 Therapeuten durchgeführt wurden und es unklar ist, wie die Behandlungen verteilt wurden und mit welcher Qualifikation und welchen Präferenzen die Behandlungen erfolgten.

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

#### etwas eingeschränkt

relativ hohe Selektivität der Patienten (nur 18,5% der gescreenten Patienten gingen in die Randomisierung ein: Einschränkung auf psychotherapiebereite chronisch Schizophrene)

### Stärken der Studie:

Umfassende Stichprobenerhebung und -beschreibung

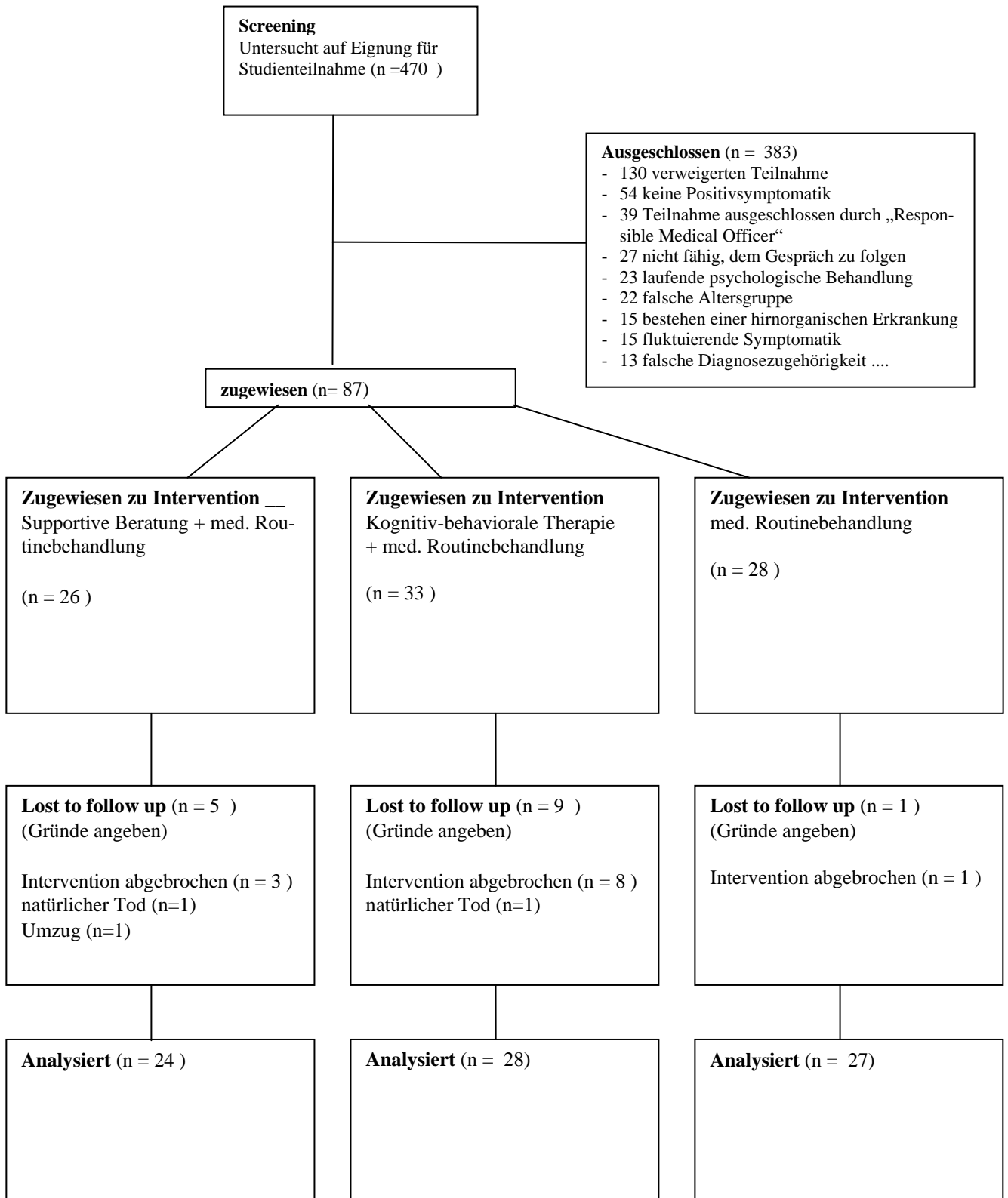
### Schwächen der Studie:

Nur 3 Therapeuten, die beide Psychotherapiearme abdecken; unklar ob Unterschiede der Behandlungseffektivität auf Therapeutenmerkmale oder Behandlungskonzepte zurückzuführen sind; fehlende Angaben zu Post-Hoc-Analysen.

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

**Studie liefert keinen Nachweis der Wirksamkeit**, da Ergebnisse zum Behandlungsende nicht für klientenzentrierten Ansatz (GPT II) sprechen, sondern für kognitiv-behaviorale Therapie, die bezüglich der Reduktion der Positivsymptomatik überlegen scheint; allerdings entsteht der Eindruck, dass diese Überlegenheit auf Präferenzen der Forscher zurück geführt werden könnte (im Einleitungsteil wird nur über kognitiv-behaviorale Therapie referiert).

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)



# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

HK

**Studie:**

Autoren:

Tarrier, N., Haddock, G., Lewis, S., Drake, R. & Gregg, L.

Titel:

Suicide behaviour over 18 months in recent onset schizophrenic patients: The effects of CBT.

Quelle:

Shizophrenia Research (2006), 83, 15-27.

## 1. Basisinformationen

### 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Supportive Counseling (zusätzlich zur Routinebehandlung)

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Erweist sich kognitiv-behaviorale Therapie (plus Routinebehandlung) im Langzeitverlauf hinsichtlich Suizidalität überlegen gegenüber Supportiver Beratung (plus Routinebehandlung) und Routinebehandlung allein bei Patienten mit erst- oder zweimalig auftretender akuter schizophrener Symptomatik?

### 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Nicht-Chronifizierte (erst- oder zweimalig auftretende ) akute schizophrene Episode

|                      | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|----------------------|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K1.</b><br>(A.2.) | Objektive und reliable Diagnostikstellung (mittels (teil-) standardisierter Interviews) | 1) Diagnostikstellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnostikstellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnostikstellung                                                                                                           | A2<br>2 |
| <b>K2.</b><br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                               | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>1 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

Klinische Diagnose einer Schizophrenie, Schizo-affektiven Erkrankung oder Wahnhafte Störung; erst- oder zweimaliges Auftreten/Behandlungersuchen; psychotische Positivsymptome seit mind. 4 Wochen; Wert von mind. 4 auf der Positive and Negative Syndrome Scale (PANS); Zielitem entweder Wahnvorstellungen oder Halluzinationen; Weder Substanzmissbrauch noch organische Faktoren als Hauptursache für psychotische Symptomatik feststellbar

|                      | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|----------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K3.</b><br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>2 |
| <b>K4.</b><br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>2 |
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>2 |

### 2.3 () Setting/Patientenzugang

- Patienten, die in einem der teilnehmenden psychiatrischen Einrichtungen vorstellig wurden.

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

|  |  |  |   |
|--|--|--|---|
|  |  |  | 1 |
|--|--|--|---|

**2.4 (7) Zentrale Patientenmerkmale**  
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

Keine Angaben

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

Keine Angaben

|                       | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating   |
|-----------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K7.</b><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br>1 |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

Analysiert: Baseline: 278; nach 6 Wochen: 210; nach 3 Monaten: 195; nach 18 Monaten: 218

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>1  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15      |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br>3 |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br>1  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>1  |

**2.7 (10) Vergleichbarkeit der Gruppen**

Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen

Durch Randomisierung gegeben

Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)

Gegeben für die beiden Gruppen mit zusätzlicher psychotherapeutischer Intervention

Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)

Durch Weiterführung der medizinisch-psychiatrischen Routinebehandlung gegeben

Vergleichbarkeit der Therapieadhärenz

Gegeben: angeblich manualisierte Behandlungsformen (keine Angabe spezifischer Literatur, nur Hinweis auf andere Studien: z.B. Tarrier et al., 1998); fortlaufende Supervision; empirisch überprüft durch korrekte Zuordnung zufällig erhobener Audiotapes sowie Einschätzungen auf der Cognitive Therapy Scale of Psychosis

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating       |
|-----------------|-------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | 1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen<br>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert<br>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt | B9<br><br>3  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | 1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen<br>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt<br>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage                                                                                                                       | A18<br><br>3 |

### 3. Intervention

|                                                                                |                                                                                                                                                                                                                                                                                                                     |
|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>3.1 (14) Studiendesign /<br/>Behandlungsarme</b>                            | <ul style="list-style-type: none"> <li>- Supportive Beratung + medizinisch-psychiatrische Routinebehandlung</li> <li>- Kognitiv-behaviorale Therapie + m.-p. Routinebehandlung</li> <li>- medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                                   |
| <b>3.2 (11) Behandlung Experimental-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Supportive Beratung (emotionale Unterstützung durch eine vertrauliche, wertschätzende Beziehung: „supportive relationship fostering rapport and unconditional regard for the patient“); 15-20 Stunden</li> <li>- + medizinisch-psychiatrische Routinebehandlung</li> </ul> |
| <b>3.3 (12) Behandlung Kontroll-<br/>gruppe<br/>(Art, Dauer, Frequenz)</b>     | <ul style="list-style-type: none"> <li>- Kognitiv-behaviorale Therapie (Verbesserung der Coping-Strategien, Problemlösetraining, Erlernen von Strategien zur Rückfallprophylaxe); 15-20 Stunden</li> <li>- + medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                |
| <b>3.4 (13) Evtl. weitere Kontroll-<br/>gruppen<br/>(Art, Dauer, Frequenz)</b> | <ul style="list-style-type: none"> <li>- Medizinisch-psychiatrische Routinebehandlung</li> </ul>                                                                                                                                                                                                                    |

|                       | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                            | Rating      |
|-----------------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K15.</b><br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | 1) ausführliches Therapiemanual<br>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie<br>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens                                                                                                                                            | B3<br><br>2 |
| <b>K16.</b><br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | 1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung<br>2) Ex post facto Beschreibung der Kontrollbedingungen<br>3) keine Beschreibung der Kontrollbedingung                                                                                                                                                                                            | B4<br><br>1 |
| <b>K17.</b><br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | 1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz<br>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab<br>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab | B5<br><br>1 |
| <b>K18.</b><br>(B.6.) | Manualtreue, Treatment Integrity                                                     | 1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt<br>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt<br>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substanzielle Abweichungen                                                                                                                       | B6<br><br>1 |
| <b>K19.</b><br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | 1) Intervention wie in klinischer Alltagspraxis<br>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert<br>3) Intervention gegenüber klinischer Alltagspraxis stark verändert                                                                                                                                                                                | C4<br><br>2 |

|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating       |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br><br>2  |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br><br>1  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br><br>1 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br><br>1 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br><br>1 |

**3.5 (19) Beobachtungszeitraum/ Messzeitpunkte/Katamnese**

- Pre-; 6-Wochen; Follow-Up (3 sowie 18 Monate nach Behandlungsbeginn)

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating       |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br><br>2  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br><br>1 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br><br>1 |

**3.6 (15) Anzahl der Zentren**

11 unabhängige Einrichtungen in 3 Regionen

**3.7 (16) Allokation zu Interventionsgruppen (Randomisierung,**

randomisierte Zuteilung

self-selection, ...)

**3.8 (17) ggf. Concealment der Randomisierung**

Gegeben; Fremdeinschätzer waren blind für die Gruppenzugehörigkeit der Patienten; Überprüfung durch Einschätzungen der Fremdeinschätzer bzgl. Der Treatmentzugehörigkeit der Patienten (keine signifikanten Übereinstimmungen mit tatsächlichen Zugehörigkeiten)

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                              | Rating  |
|-----------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br>3 |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br>1 |



## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

Krankenakten  
Festgestellte Suizide und Unfalltode

Fremdeinschätzungsverfahren  
- Scala "Non-Accidental Self Injury Scale" aus HoNOS (Health of the Nation Outcome Scales; Wing et al., 1998)

### 4.2 (21) Sekundäre Zielkriterien

### 4.3 () Verblindung der Erhebung der Zielkriterien

- gegeben für HoNOS durch Geheimhaltung der Treatmentzuweisung vor den Fremdeinschätzern

|                        | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating      |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K28.</b><br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | <ol style="list-style-type: none"> <li>1) a priori definiert unter Berücksichtigung der Wertigkeit</li> <li>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien</li> <li>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,</li> </ol>                                                                                                                                                                         | A7<br><br>2 |
| <b>K29.</b><br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | <ol style="list-style-type: none"> <li>1) reliable und valide Outcome-Verfahren</li> <li>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren</li> <li>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend</li> </ol>                                                                                                                                                                             | A8<br><br>2 |
| <b>K30.</b><br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | <ol style="list-style-type: none"> <li>1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität</li> <li>2) Zielkriterien beziehen sich nur auf eine Dimension</li> <li>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)</li> </ol>                                                                                                                       | C9<br><br>2 |
| <b>K31.</b><br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | <ol style="list-style-type: none"> <li>1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen</li> <li>2) zwei Informationsquellen</li> <li>3) eine Informationsquelle</li> </ol>                                                                                                                                                                                                                                                                        | A9<br><br>2 |
| <b>K32.</b><br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | <ol style="list-style-type: none"> <li>1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler</li> <li>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern</li> <li>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten</li> </ol> | A10         |

**4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

Kovarianzanalysen? Unklar was tatsächlich analysiert wurde: etwa eine abhängige dichotome Variable mit dem Varianzanalytischen Modell? Es scheint fast, als fehlte ein Stück Text!

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>1 |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13     |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14     |

**4.5 (22) Ergebnisse**

Hinsichtlich suizidalen Verhaltens (HoNOS-Skala) unterschieden sich die Gruppen nicht statistisch signifikant. Hinsichtlich tatsächlicher Suizide und Unfalltode waren die Fallzahlen zu gering (3 Suizide, 2 Unfalltode; davon 3 in der Supportive-Counseling und 2 in der CBT-Gruppe; keiner in der Treatment-Usual-Gruppe)

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-Maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>2 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

- keine Angaben/Analysen

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                    | Rating              |
|------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br><b>3</b> |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Kognitiv-behaviorale Therapie ist ein effektiver Zusatz zur routinemäßigen psychiatrischen Behandlung bei schizophrener Positivsymptomatik

|                       | Kriterium              | Operationalisierung                                                                    | Rating             |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br><b>1</b> |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |             |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1<br><br>3 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2          |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3          |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                                   | Rating       |
|-----------------|----------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br><br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br><br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br><br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br><br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br><br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br><br>1 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br><br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br><br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br><br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br><br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br><br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br><br>1 |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>1. wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>2. wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>3. wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                                  |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>1. die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>2. der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                       |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>1. das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>2. das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>3. bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>4. bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

#### schwach

- fehlende Angaben zur Größe der Behandlungsgruppen (dadurch auch CONSORT-Diagramm nur rudimentär ausfüllbar;
- relevante Datenanalyse (Tab. 3) unklar, evtl. fehlt ein Stück Text (vgl. in der Publikation Punkt 3.3), evtl. sogar falsches statistisches Verfahren angewandt?! Jedenfalls keine klare Bewertung möglich.

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Hinsichtlich tatsächlicher Suizide zu geringe Fallzahlen für Datenanalysen: nur 3 verzeichnete Suizide und 2 Unfälle, bei denen Suizid nicht ausgeschlossen werden kann (von diesen 5 möglichen Suiziden sind 2 zum Supportive Counseling und 3 zur CBT zugeordnet; in Treatment-As-Usual-Gruppe kein Suizid/Unfalltod)

Hinsichtlich der Suizidalität abbildenden Skala aus der HoNOS (auf die sich die unklare Datenanalyse bezieht): keine signifikanten Unterschiede zwischen den Behandlungsgruppen.

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

**nicht gegeben**, da die relevante Datenanalyse unklar ist; weiterhin wurden regionale Unterschiede nicht analysiert, die allerdings in einer früheren Publikation auf Basis des gleichen Datensatzes relevant waren (vgl. Tarrier, 2004)

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

### eingeschränkt

Unklar welche Ein-/Ausschlusskriterien bestanden haben: zwar angeblich dieselben wie bei Tarrier, 2004, doch kleinerer Stichprobenumfang (hier: N-Baseline=278; bei Tarrier, 2004: N-Baseline= 309; was ist mit den fehlenden 31 Patienten?)

Weiterhin gelten dieselben Einschränkungen wie bei Tarrier, 2004: nur erst- bzw. zweimalig auftretende schizophrene Episoden, keine chronifizierte Störungen. Supportive Beratung wurde von ausgebildeten Verhaltenstherapeuten durchgeführt. Durchschnittlich lediglich 8 Stunden Therapiedauer (7.1 Support. Beratung; 8.6 CBT)

### Stärken der Studie:

Große Stichprobe, randomisierte Zuteilung, Concealment

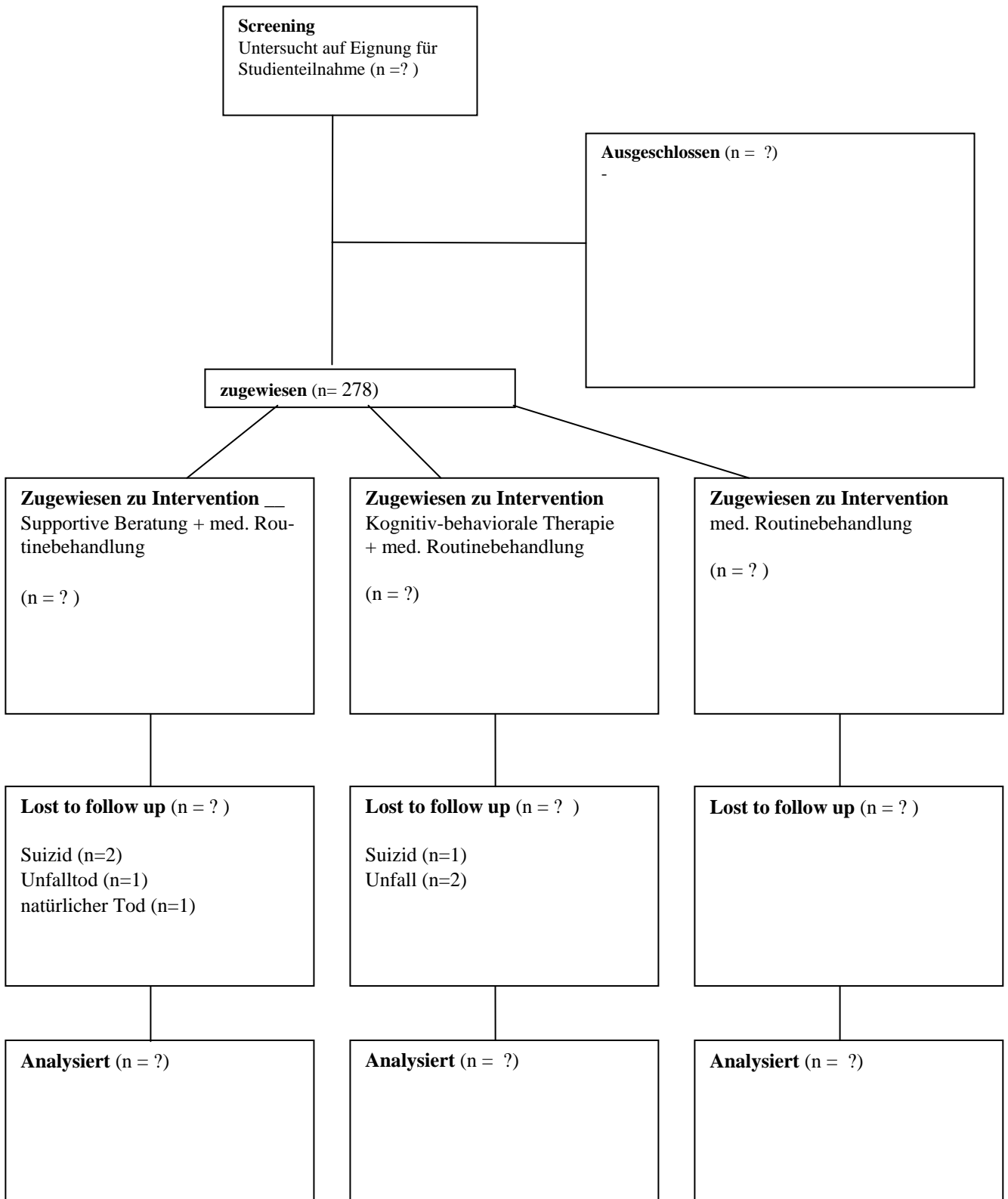
### Schwächen der Studie:

Nur 5 Therapeuten, die beide Psychotherapiearme abdecken; lediglich regional arbeitende Fremdeinschätzer, regional unterschiedliche Behandlungsdauer.

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

**Keine über die Publikation von Tarrier et al. (2004) hinausgehende Aussagekraft:** unklare Datenanalysen, unklare Ein-/Ausschlusskriterien, hinsichtlich Outcome-Kriterium verübte Suizide/Suizidalität keine Überlegenheit für Supportives Counseling (+ Treatment-As-Usual) gegenüber Treatment-As-Usual

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.



**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

Kirchmann

**Studie:**

Autoren:

Watson, J. C. et al.

Titel:

Comparing the effectiveness of process-experiential with cognitive-behavioral psychotherapy in the treatment of depression

Quelle:

Journal of Consulting and Clinical Psychology (2003), 71, 773-781.

## 1. Basisinformationen

### 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

„Process-Experiential Psychotherapy“:  
klientenzentrierte Beziehungsgestaltung mit zusätzlichen gestalttherapeutischen Elementen (z.B. empty chair) zur gezielten Exploration von emotionalen Erfahrungen/Prozessen Kurzzeittherapie mit 16 Sitzungen 1\*wöchentlich;  
(Manual basierend auf Greenberg et al., 1993)

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Vergleich der Effektivität der „Process-Experiential“ Psychotherapie mit Kognitiv-Behavioraler Therapie bei Depressionen; 16 Sitzungen, 1\*wöchentlich, manualisiert nach Beck et al., 1979.

### 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Major Depression

|                      | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating      |
|----------------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K1.</b><br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br><br>1 |
| <b>K2.</b><br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br><br>1 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

Einschluss: SCID-Diagnose: Major Depression;

Ausschluss:

- laufende psychotherapeutische oder pharmakologische Behandlung;  
Sprachprobleme; Suizidalität; komorbide Substanzabhängigkeit, Essstörung, psychotische oder bipolare Störung, Antisoziale, Borderline oder Schizotypische Persönlichkeitsstörung

|                      | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|----------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K3.</b><br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>1 |
| <b>K4.</b><br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>2 |
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>2 |

### 2.3 () Setting/Patientenzugang

Zeitungsanzeigen, Radiowerbung

|                      | Kriterium                           | Operationalisierung                                                                                                                                                                                                                                                                                                                                            | Rating  |
|----------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K6.</b><br>(C.2.) | Art der Rekrutierung der Stichprobe | 1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege<br>2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege<br>3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung) | C2<br>3 |

### 2.4 (7) Zentrale Patientenmerkmale (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

Alter: M=41.52; Geschlecht: 67% weiblich; Ethnizität: 91% European, 6% Asian; 3% Hispanic;  
Symptomschwere nach BDI: mild-moderat: 18%; moderat-schwer: 58%; extrem schwer: 24%  
Anzahl früherer Episoden: gegenwärtig 1. Episode: 6%; bisher 2-4 Episoden: 26%; bisher mehr als 5 Episoden: 62%  
Länge der gegenwärtigen Episode: weniger als 6 Monate: 29%; 6 Monate bis 9 Jahre: 51%; länger als 9 Jahre: 12%

### 2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl

Angaben erst im Diskussionsteil: bei power=.80; alpha=.05 und mittlere bis große Effekte: N= 26 bis 62

|                       | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating   |
|-----------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K7.</b><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br>2 |

**2.6 (9) Anzahl gescreener, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

368 (Telefoninterview); 182 (persönliches Interview mit SKID), 93 randomisiert zugewiesen auf zwei Treatmentbedingungen (PET: 43; CBT: 50); Therapie abgeschlossen: 66 (PET: 33; CBT:33)  
- allerdings Stichprobenfluktuationen je nach Outcome-Variable (28-33)

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>2  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>1 |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br>2 |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br>2  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4       |

**2.7 (10) Vergleichbarkeit der Gruppen**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen</p> <p>- Gegeben: Randomisierung, Pre-Treatment-Analysen n.s. bzw. statistisch kontrolliert</p> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)</p> <p>- Gegeben: alle Patienten erhielten 16 Sitzungen 1*wöchentlich</p> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)</p> <p>- Gegeben durch Ausschluss</p> <p>Vergleichbarkeit der Therapieadhärenz</p> <p>Gegeben: entsprechende Prüfungen durch Fremdeinschätzungen</p> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Rating              |
|-----------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | <p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanten Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p> | <p>B9</p> <p>1</p>  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | <p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>                                                                                                                       | <p>A18</p> <p>2</p> |



### 3. Intervention

**3.1 (14) Studiendesign /  
Behandlungsarme**

a) „Process-Experiential Psychotherapy“  
vs.  
b) Kognitiv-Behaviorale Therapie

**3.2 (11) Behandlung Experimental-  
gruppe  
(Art, Dauer, Frequenz)**

„Process-Experiential Psychotherapy“:  
klientenzentrierte Beziehungsgestaltung mit zusätzlichen gestalttherapeutischen Elementen zur gezielten Exploration von emotionalen Erfahrungen/Prozessen Kurzzeittherapie mit 16 Sitzungen 1\*wöchentlich;  
(Manual basierend auf Greenberg et al., 1993)

**3.3 (12) Behandlung Kontroll-  
gruppe  
(Art, Dauer, Frequenz)**

Kognitiv-Behaviorale Therapie:  
„kognitive Therapie mit einigen behavioralen Komponenten wie aufschreiben der alltäglichen Aktivitäten und Verhaltensexperimente“  
(nach Beck et al., 1979)

**3.4 (13) Evtl. weitere Kontroll-  
gruppen  
(Art, Dauer, Frequenz)**

|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K15.</b><br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe)                                                    | 1) ausführliches Therapiemanual<br>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie<br>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens                                                                                                                                                                                                                                                                                                                | B3<br>2  |
| <b>K16.</b><br>(B.4.) | Operationale Definition der Kontrollbedingungen                                                                                         | 1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung<br>2) Ex post facto Beschreibung der Kontrollbedingungen<br>3) keine Beschreibung der Kontrollbedingung                                                                                                                                                                                                                                                                                                                                                                | B4<br>1  |
| <b>K17.</b><br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                                                                         | 1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz<br>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab<br>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab                                                                                                                                                                     | B5<br>1  |
| <b>K18.</b><br>(B.6.) | Manualtreue, Treatment Integrity                                                                                                        | 1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt<br>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt<br>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen                                                                                                                                                                                                                                                                                           | B6<br>1  |
| <b>K19.</b><br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer                                                            | 1) Intervention wie in klinischer Alltagspraxis<br>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert<br>3) Intervention gegenüber klinischer Alltagspraxis stark verändert                                                                                                                                                                                                                                                                                                                                                    | C4<br>1  |
|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantiellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br>1  |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br>3  |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)                                                                  | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br>2 |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br>2 |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br>3 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

Nur Prä-Post-Erhebung

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating       |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br><br>2  |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br><br>2 |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br><br>3 |

**3.6 (15) Anzahl der Zentren**

1

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

Randomisierung

**3.8 (17) ggf. Concealment der Randomisierung**

Randomisierung vorgenommen von „research assistants“, unabhängig vom Erstautor.

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating      |
|-----------------------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                         | C8<br><br>3 |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br><br>1 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

BDI

### 4.2 (21) Sekundäre Zielkriterien

IIP  
 Rosenberg Self-Esteem Scale  
 SCL-90  
 Dysfunctional Attitudes Scale  
 Problem-Focused Style of Coping

### 4.3 () Verblindung der Erhebung der Zielkriterien

Keine Fremdeinschätzungsverfahren

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating  |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>2 |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | A8<br>1 |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | C9<br>1 |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | A9<br>3 |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10     |

4.4 () **Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

Varianzanalysen mit Messwiederholung (Einschluss einer Kovariate)

|                 | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|-----------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K33.<br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>1  |
| K34.<br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| K35.<br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>2 |

4.5 (22) **Ergebnisse**

Die beiden Behandlungsgruppen unterscheiden sich nicht hinsichtlich des Therapieerfolges im BDI. Beide Gruppen erreichen große Effekte (PET: ES= 1.73; CBT: ES=1.69); Reliable Veränderungen entsprechen RCI wurden von 57.57% der PET-Gruppe und 51.51% der CBT-Gruppe erreicht.

Hinsichtlich der sekundären Zielkriterien wurden ebenfalls statistisch signifikante, aber geringere Effektstärken in beiden Behandlungsgruppen erreicht (ES=0.30 bis 1.04)

Lediglich hinsichtlich des IIP unterschieden sich die Behandlungsgruppen signifikant im Therapieerfolg (PET besser als CBT)

Im „Intention-to-Treat-Sample war die Ergebnislage sehr ähnlich

|                 | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K36.<br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcomemaßen und zu allen relevanten Messzeitpunkten                                              | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| K37.<br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

- keine Angaben

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating              |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br><b>3</b> |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

Symptombezogen waren beide Therapieformen äquivalent, bezüglich interpersonaler Probleme zeigte sich die PET der CBT überlegen.

|                       | Kriterium              | Operationalisierung                                                                    | Rating             |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br><b>1</b> |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |         |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1<br>3 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2<br>1 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3<br>1 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>1 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>1 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>1 |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |



## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Gut: Aufwändige Therapiestudie (z.B. Tätigkeit von 15 verschiedenen Therapeuten, Erhebung von SKID-Diagnosen, manualisierte Therapien mit vorhergehendem Therapeutentraining und begleitender Adherence-Prüfung)

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

Beide Behandlungsgruppen erreichten große Effekte in der Reduktion depressiver Symptome nach dem BDI (ES für PET=1.73, für CBT=1.69). Auch wenn der Post-Treatment-Mittelwert der PET-Gruppe mit 9.03 ziemlich genau am BDI-Cut-Off für „Recovery“ (BDI=9) liegt, scheint dies doch eine klinisch relevante Symptomreduktion zu sein.

Leider liegen keine Katamnesedaten vor. Für eine längerfristige Effektivität liegen Hinweise vor in Form von ebenfalls verbesserten Werten in Dysfunktionalen Einstellungen, den erhobenen Coping-Skalen sowie den interpersonalen Problemen.

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Gut

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Eingeschränkt: Gewinnung der Probanden durch Werbung in Zeitungen/im Radio (evtl. hoch motivierte Personen); relativ hohe Rate an Therapieabbrechern in beiden Behandlungsgruppen (25-30%)

### Stärken der Studie:

Gute allgemeine methodische Qualität (siehe oben);

### Schwächen der Studie:

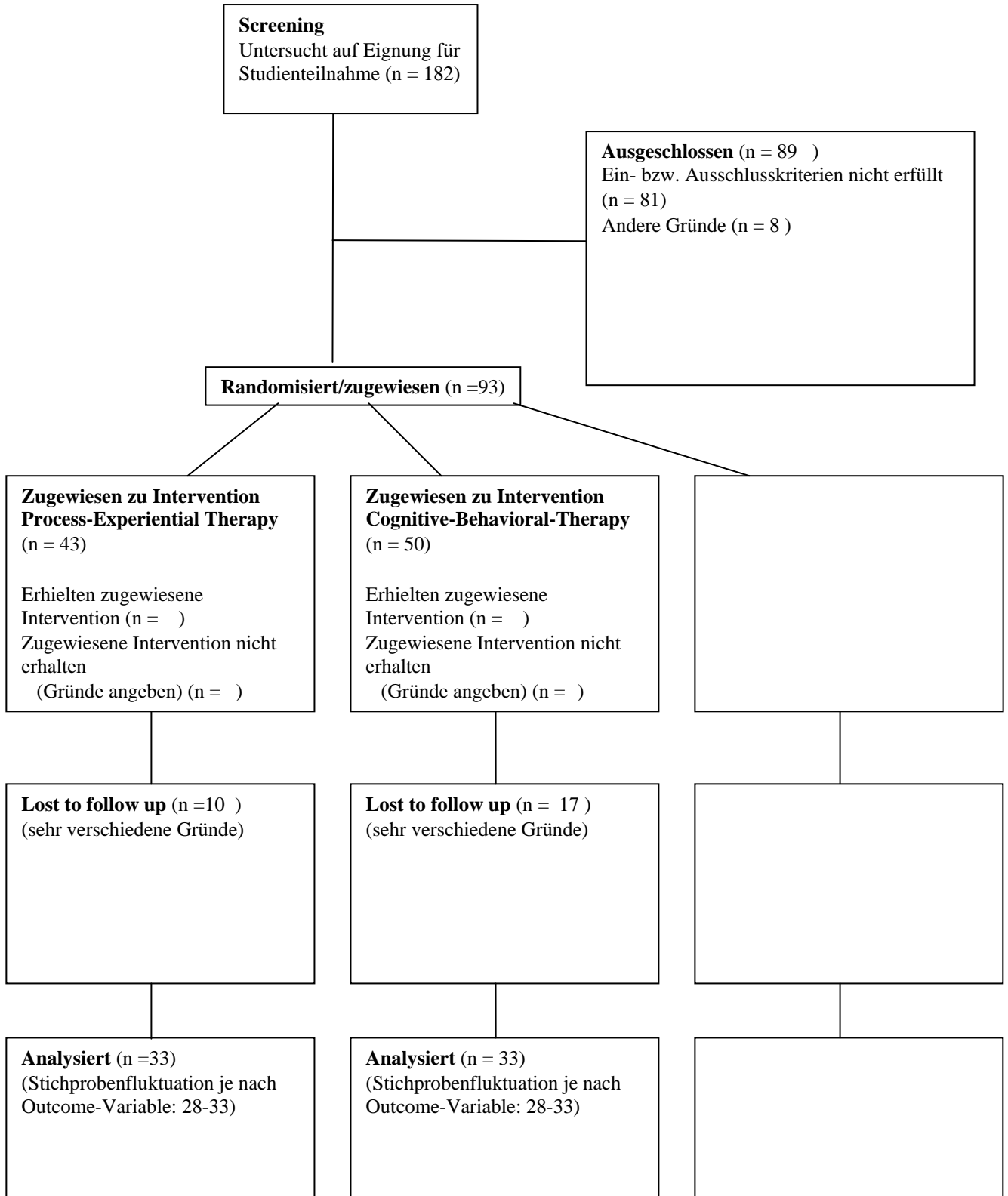
Nur Selbsteinschätzungsverfahren verwendet; keine Katamnese

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

**Relevant:** Hinsichtlich des primären Zielkriteriums BDI-Wert zeigte die „Process-Experiential Psychotherapy“: (klientenzentrierte Beziehungsgestaltung mit zusätzlichen gestalttherapeutischen Elementen) gleichwertige Effektivität im Pre-Post-Vergleich und allen erhobenen sekundären Zielkriterien.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**

**Gründe für Drop-out schlecht beschrieben: keine Differenzierung zwischen den Behandlungsarmen**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

MT

**Studie:**

Autoren:

Teusch et al.

Titel:

Konfliktzentrierte Monotherapie oder Methodenintegration? Veränderungsprozesse von Gesprächspsychotherapie mit und ohne verhaltenstherapeutische Reizkonfrontation bei Agoraphobie mit Panikstörung

Quelle:

Nervenarzt, 2001, 72: 31-39

## 1. Basisinformationen

### 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Gesprächspsychotherapie (GPT) nach Manual für die Behandlung von Panik und Agoraphobie von Teusch & Finke (*Teusch L, Finke J; 1995; Psychotherapeut*)

- Einzel- & Gruppentherapie
- Ergänzend Bewegungstherapie, Gestaltungstherapie, Ergotherapie, Entspannungstechniken

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Vergleich der Wirksamkeit von GPT versus GPT + behaviorale Exposition (EXP) bei Agoraphobie mit Panikstörung

Fragestellungen:

- Unterschiedliche Einflüsse auf die Leitsymptomatik und auf die Persönlichkeitsentwicklung?
- charakteristische Unterschiede dieser Veränderungen im zeitlichen Verlauf?
- kommt es bei GPT+EXP zu einer Addition der Effekte beider Behandlungsverfahren?
- gibt es überlegene Wirkeffekte bei ausschließlicher GPT?
- fördert ausschließliche GPT die Beziehungs- und Konfliktfähigkeit, wie aufgrund der Krankheits- und Therapietheorie erwartet werden könnte?

### 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Agoraphobie und Panikstörung (DSM-III-R, No. 300.21)  
- klinische Diagnose (Therapeut), Überprüfung durch SCID (klinischer Psychologe)

|               | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating |
|---------------|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| K1.<br>(A.2.) | Objektive und reliable Diagnostikstellung (mittels (teil-) standardisierter Interviews) | 1) Diagnostikstellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnostikstellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnostikstellung                                                                                                           | 1 A2   |
| K2.<br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                               | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | 1 C1   |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

- Voraussetzung: Agoraphobie und Panikstörung (DSM-III-R, No. 300.21)  
- Ausschluss: psychotische Störung, hirnorganische Störung, manifeste Substanzabhängigkeit

|               | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating |
|---------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| K3.<br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 1 B1   |
| K4.<br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | 1 B2   |
| K5.<br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1 C3   |



**2.3 ( ) Setting/Patientenzugang**

Überweisung durch niedergelassene Psychiater und Therapeuten

|                          | Kriterium                           | Operationalisierung                                                                                                                                                                                                                                                                                                                                            | Rating |
|--------------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>K6.</b><br><br>(C.2.) | Art der Rekrutierung der Stichprobe | 1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege<br>2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege<br>3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung) | 1 C    |

**2.4 (7) Zentrale Patientenmerkmale**  
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- überwiegend Frauen (25 m, 43 w)
- mittleres Alter 32,4 Jahre (SD 8,2)
- meistens mit Partner lebend (N = ?)
- Schulabschluss: Hauptschule (N = 30)/ mittl. Reife (N = 23)/ Abitur (N = 15)
- Erkrankungsdauer vor Einschluss: 5,5 Jahre (SD 5,0)

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

Keine Angaben

|                           | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating |
|---------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>K7.</b><br><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | 2 A16  |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- Keine Angaben über die Anzahl gescreenter Patienten
- 82 Patienten eingeschlossen
- 14 Dropouts in Woche 1-4 werden analytisch nicht in die Studie aufgenommen
- die verbleibenden 68 Patienten wie folgt eingeschlossen:
  - 40 Patienten eingeschlossen und je 20 randomisiert einer der Vergleichsgruppen zugewiesen
  - 28 weitere Patienten wurden aus therapeutischen oder administrativen Gründen nicht randomisiert, sondern nach klinischen Gesichtspunkten einer der beiden Bedingungen zugewiesen („comprehensive cohort design“)
- keine Dropouts zu den 2 Katamnesezeitpunkten
- 68 Patienten ausgewertet

|                 | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|-----------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K8.<br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>2  |
| K9.<br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>3 |
| K10.<br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A7<br>2  |
| K11.<br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | 2<br>A3  |
| K12.<br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>1  |

## 2.7 (10) Vergleichbarkeit der Gruppen

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen:</p> <ul style="list-style-type: none"> <li>- Gruppen statistisch vergleichbar bzgl. der erhobenen sozialen und demographischen Faktoren</li> <li>- relevant erscheinen dennoch Unterschiede bzgl. der Geschlechterverteilung (m: GPT - 29 %, GPT+EXP - 43%) und der Bildung (GPT – 38% Hauptschule &amp; 29% Abitur, GPT+EXP - 48% Hauptschule &amp; 18% Abitur)</li> <li>- d.h. augenscheinlich sind in der GPT+EXP Gruppe relativ mehr männliche Patienten und Patienten mit einem niedrigen Bildungsgrad als in der GPT Gruppe</li> <li>- den Ausgangswerten nach erscheint die GPT Gruppe etwas stärker gestört (etwas stärkere Agoraphobie gemessen über FSS und Unterschiede in drei Skalen des Gießen-Tests: Soziale Resonanz, Kontrolle, Summe aller Diskrepanzen zw. Selbstbild und Idealbild)</li> </ul> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer):</p> <ul style="list-style-type: none"> <li>- CCT+EXP-Gruppe zeitlich intensivere Behandlung (siehe 3.2 &amp; 3.3)</li> </ul> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen):</p> <ul style="list-style-type: none"> <li>- Nicht bekannt</li> </ul> <p>Vergleichbarkeit der Therapieadhärenz:</p> <ul style="list-style-type: none"> <li>- ist gegeben (Dropouts: 6 GTP, 8 GTP+EXP)</li> </ul> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Rating   |
|-----------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | <p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p> | B9<br>2  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | <p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>                                                                                                                       | A18<br>3 |

### 3. Intervention

**3.1 (14) Studiendesign /  
Behandlungsarme**

Therapiestudie mit randomisierter Vergleichsgruppe, 2 Behandlungsarme: 1. Gesprächspsychotherapie (GPT) mit zusätzlichem Expositionstraining und Hausaufgaben (EXP) 2. GPT

**3.2 (11) Behandlung Experimental-  
gruppe  
(Art, Dauer, Frequenz)**

GPT nach Manual für die Behandlung von Panik und Agoraphobie von Teusch & Finke (*Teusch L, Finke J; 1995; Psychotherapeut*)  
- 1x/Woche Einzeltherapie, 3x/Woche Gruppentherapie

Expositionstherapie (EXP)  
- tägliche Angstexposition und 1x/pro Woche Gruppenexpositionstraining unter therapeutischer Anleitung nach dem Manual von Mathews et al. (1988)

- stationärer (zum Teil tagesklinischer) Aufenthalt für 88+/- 20 Tage

**3.3 (12) Behandlung Kontroll-  
gruppe  
(Art, Dauer, Frequenz)**

GPT (s.o.)  
- 1x/Woche Einzeltherapie, 4x/Woche Gruppentherapie

- stationärer (zum Teil tagesklinischer) Aufenthalt für 88+/- 20 Tage

**3.4 (13) Evtl. weitere Kontroll-  
gruppen  
(Art, Dauer, Frequenz)**

keine

|                | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                            | Rating  |
|----------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K15.<br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | 1) ausführliches Therapiemanual<br>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie<br>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens                                                                                                                                            | B3<br>1 |
| K16.<br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | 1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung<br>2) Ex post facto Beschreibung der Kontrollbedingungen<br>3) keine Beschreibung der Kontrollbedingung                                                                                                                                                                                            | B4      |
| K17.<br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | 1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz<br>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab<br>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab | B5<br>2 |
| K18.<br>(B.6.) | Manualtreue, Treatment Integrity                                                     | 1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt<br>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt<br>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen                                                                                                                       | B6<br>9 |
| K19.<br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | 1) Intervention wie in klinischer Alltagspraxis<br>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert<br>3) Intervention gegenüber klinischer Alltagspraxis stark verändert                                                                                                                                                                                | C4<br>2 |

|                | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K20.<br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantiellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br>9  |
| K21.<br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br>9  |
| K22.<br>(C.7.) | Qualifikation der Behandler<br>(insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                     | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br>2 |
|                | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br>9 |
|                | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br>9 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

Ca. 12 Wochen Behandlungszeitraum, Untersuchung bei Aufnahme in die Studie und Entlassung aus der Behandlung, zwei Katamnese Untersuchungen (6 und 12 Monate nach Entlassung), d.h. ca. 15 Monate Beobachtungszeitraum

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating  |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br>3 |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10     |
| <b>K25.</b><br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11     |

**3.6 (15) Anzahl der Zentren**

1

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

Teilweise Randomisierung, teilweise Zuweisung nach klinischen Gesichtspunkten

**3.8 (17) ggf. Concealment der Randomisierung**

Keine Angabe

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                              | Rating  |
|-----------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8      |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br>2 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

-Reduktion der Schwere der Panikstörung (SKID) und der Agoraphobie (SKID & FSS)

### 4.2 (21) Sekundäre Zielkriterien

- Persönlichkeitsentwicklung (FPI-R, Gießen-Test), insb. Beziehungs- und Konfliktfähigkeit

FPI-R, (gemessene) Skalen: Lebenszufriedenheit, Soziale Orientierung, Leistungsorientierung, Gehemmtheit, Erregbarkeit, Aggressivität, Beanspruchung, Körperbeschwerden, Gesundheitsorgen, Offenheit, Extraversion, Emotionalität

Gießen-Test, (gemessene) Skalen: Soziale Resonanz, Dominanz, Kontrolle, Grundstimmung, Durchlässigkeit, Soziale Potenz, Mittel (zentrale Werte), Extreme (extreme Werte), Selbst-Ideal-Diskrepanz

### 4.3 () Verblindung der Erhebung der Zielkriterien

Nicht bekannt

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>2  |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | 1A8      |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | 1C9      |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | 1A9      |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10<br>9 |

**4.4 ( ) Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

Jeweils für alle Messinstrumente:  
 - Multivariate Varianzanalyse mit wiederholten Messungen (MANOVA) über 4 Zeitpunkte und 2 Gruppen  
 - bei signifikanten Veränderungen zwischen Zeitpunkten abhängig von der Gruppe (Zeitpunkt x Gruppe) wurde für das jeweilige Instrument mittels einer ANOVA die Effektstärke zwischen den Gruppen verglichen

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | 1 A6     |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>3 |

**4.5 (22) Ergebnisse**

**GPT und GPT+EXP**  
 - zwischen Prä- und Postmessung hoch signifikante Verbesserung von Panik und Agoraphobie (SKID) im zeitlichen Verlauf über die vier Messzeitpunkte hinweg sowie Besserungen in fast allen Persönlichkeitsvariablen (FPI-R und Gießen-Test) hinweg (darunter Lebenszufriedenheit, Gehemmtheit, Beanspruchung, Gesundheits Sorgen, Körperbeschwerden, soziale Orientierung)

**GPT vs. GPT+EXP**  
 - hinsichtlich der Leitsymptome Panik und Agoraphobie keine Unterschiede zwischen den Gruppen im zeitlichen Verlauf  
 - in vier Variablen berichten die Autoren Unterschiede zwischen den Verfahren:  
 1. GPT besser GPT+EXP:  
 - Abhängigkeit/ Anklammerung (soziale Orientierung) sinkt kontinuierlich, während Parameter bei GPT+EXP zum letzten Messzeitpunkt hin wieder ansteigt.  
 - Gefühl der Beanspruchung und Körperbeschwerden sinken bei GPT kontinuierlich, während beide bei GPT+EXP erst im Verlauf abnehmen  
 2. GPT+EXP besser GPT:  
 - Gefühl, von anderen sozial akzeptiert zu werden, steigt schneller



|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |

**4.6. (23) Ergebnisse zu unerwünschten Therapie-  
wirkungen**

Nicht bekannt oder nicht berichtet

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                    | Rating   |
|------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignisse, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br>3 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

- Wirksamkeit von GPT und GPT+EXP bei Agoraphobie und Panikstörung  
 - GPT- Mono-Therapie hat sich für die Entwicklung wichtiger Persönlichkeitsbereiche als günstiger als eine Kombination mit Verhaltenstherapie erwiesen

|                       | Kriterium              | Operationalisierung                                                                    | Rating |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1     |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating                       |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C <sub>2</sub> <sup>10</sup> |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C <sub>1</sub>               |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C <sub>1</sub> <sup>12</sup> |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C <sub>2</sub> <sup>13</sup> |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Die allgemeine methodische Qualität ist eingeschränkt.

Die Diagnosestellung erfolgt objektiv und reliabel. Es gibt 2 Katamnese-Erhebungen. Es werden reliable und valide Outcome-Verfahren genutzt.

Das Design erlaubt keinen Nachweis einer eventuellen Überlegenheit der GPT, da die Vergleichsgruppe GPT + Exposition bekommt. Es gibt keine Angaben über die Anzahl gescreener Patienten. Die statistische Power ist mäßig. Trotz relativ hoher Dropout-Raten, werden keine IIT- und Dropout-Analysen durchgeführt

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

#GPT und GPT+EXP führen zu nachhaltigen Therapieeffekten

#es werden positive differentielle Effekte einer Mono-GPT auf verschiedenen Persönlichkeitsvariablen gezeigt (siehe Kritik der methodischen Qualität der Studie)

#die Angstsymptomatik verbessert sich deutlich oder verschwindet

#die Befunde sind konsistent

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Die Zuordnung zu den Therapiegruppen erfolgt nur überwiegend randomisiert („comprehensive cohort design“). Obwohl ein Unterschied nicht statistisch ausgewiesen wird, erscheinen die Gruppen deskriptiv nicht ganz vergleichbar bzgl. Geschlechterverteilung und Bildungsniveau.

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Die externe Validität bzgl. Setting und Frequenz ist eingeschränkt (stationär, 4-5x/ Woche) und nicht auf ein ambulantes Setting im dt. Versorgungssystem übertragbar. Auch die Behandler sind nicht repräsentativ (eng supervidierte Ausbildungskandidaten).

### Stärken der Studie:

2 Katamneseerhebungen

### Schwächen der Studie:

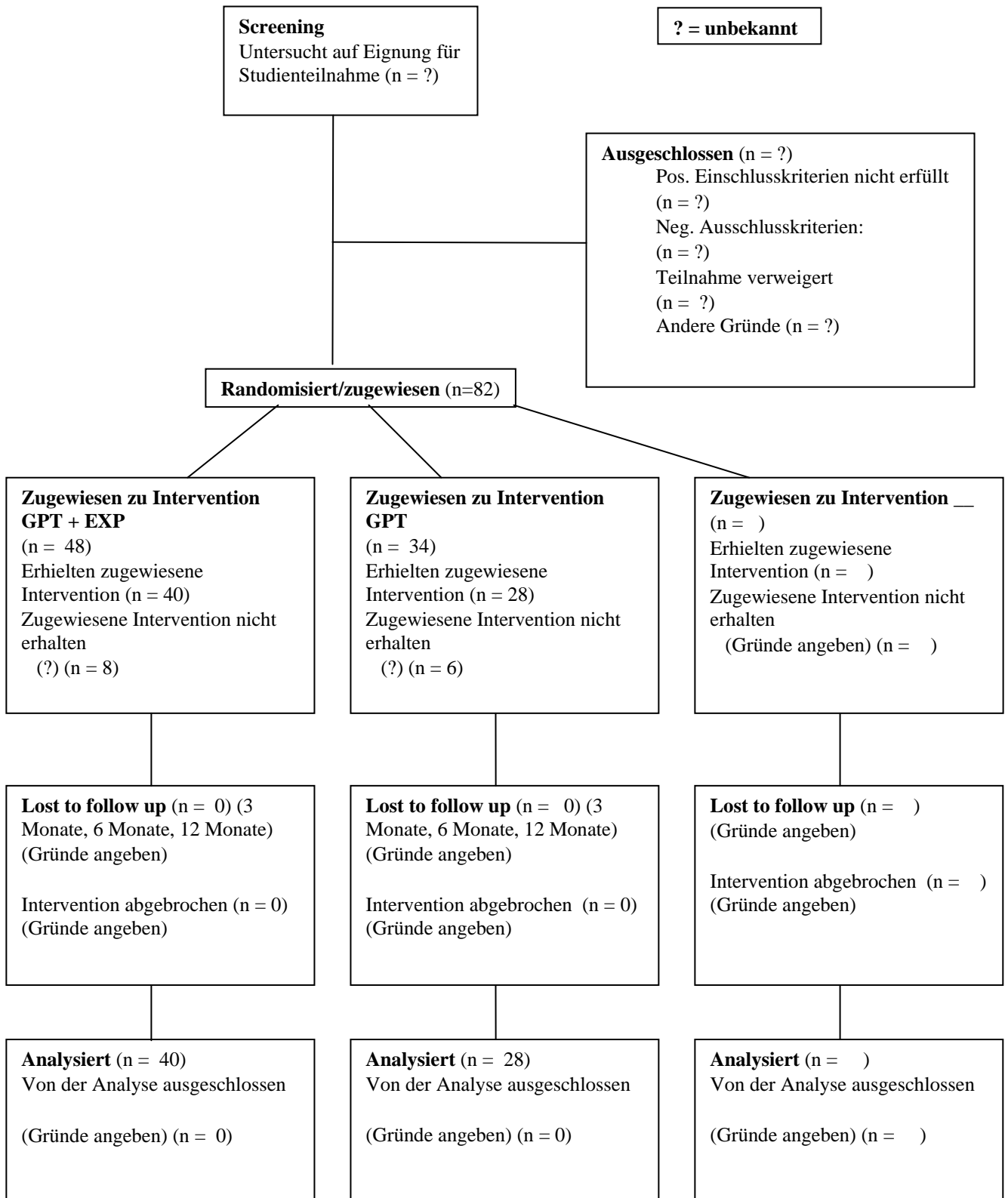
Die Zuordnung zu den Gruppen erfolgt nur teilweise randomisiert. Die Wahl der Versuchsgruppen (GPT vs. GPT + EXP) erlaubt keine Beurteilung der spezifischen Wirksamkeit der GPT. In beiden Gruppen wurde jeweils stationär ein Bündel verschiedener Maßnahmen realisiert, so dass nicht beurteilt werden kann, worauf der Effekt genau zurückzuführen ist. Die externe Validität ist durch mangelnde Übertragbarkeit von Therapie Setting und Frequenz eingeschränkt.

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie ist aufgrund methodischer Einschränkungen nicht für einen Nachweis der Wirksamkeit der GT bei Angststörungen geeignet. Das Design lässt Aussagen über die Wirksamkeit der GT nur auf Evidenzstufe IV zu, da es für die hier interessierende Fragestellung keine geeignete Vergleichsgruppe gibt. Sie liefert jedoch Hinweise auf eine Wirksamkeit der GT in einem stationären Setting.

Die spezifische Wirksamkeit kann nicht beurteilt werden, da es 1. keine echte Kontrollgruppe gibt und 2. stationär in jeder Gruppe ein Bündel von Maßnahmen realisiert wurde. Die allgemeine methodische Qualität und die externe Validität bzgl. Therapiesetting, -frequenz sind eingeschränkt.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.



## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

MT

**Studie:**

Autoren:

Teusch, L & Böhme, H

Titel:

Is the Exposure Principle really Crucial in Agoraphobia? The Influence of Client-Centered "NonPrescriptive" Treatment on Exposure

Quelle:

Psychother. Res., 1999, 9(1): 115-123

## 1. Basisinformationen

### 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Klientenzentrierte Therapie (CCT) nach Manual für die Behandlung von Panik und Agoraphobie von Teusch & Finke (*Teusch L, Finke J; 1995; Psychotherapeut; engl. Version 1996*) & „Occupational Therapy“

- Einzel- & Gruppentherapie
- Psychogymnastik & kreative Elemente
- Keine: direktiven Interventionen, Exposition, Hausaufgaben

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

Differentielle Wirksamkeit von CCT vs. CCT + behaviorale Exposition vs. behaviorale Exposition

- Ist Exposition für eine wirksame Behandlung einer Agoraphobie notwendig?

- bei Panik und Agoraphobie

- bezogen auf Angstsymptome (spezif. die Bereitschaft, sich phobischen Reizen auszusetzen, die Beeinträchtigung durch die Störung)

### 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Agoraphobie mit Panikstörung (DSM-III-R, No. 300.21)  
- klinische Diagnose (Therapeut), Überprüfung durch SCID (klinischer Psychologe)

|                      | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating |
|----------------------|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>K1.</b><br>(A.2.) | Objektive und reliable Diagnostikstellung (mittels (teil-) standardisierter Interviews) | 1) Diagnostikstellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnostikstellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnostikstellung                                                                                                           | 1 A2   |
| <b>K2.</b><br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                               | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | 1 C1   |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

- bekannt für CCT und CCT+EXP, nicht für EXP  
- Voraussetzung: „Panikstörung mit Agoraphobie“ als dominante Diagnose  
- Ausschluss: psychotische Störungen, Drogen- oder Alkoholabhängigkeit

|                      | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating |
|----------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>K3.</b><br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 9 B1   |
| <b>K4.</b><br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | 9 B2   |
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 9 C3   |

**2.3 ( ) Setting/Patientenzugang**

- bekannt für CCT und CCT+EXP, nicht für EXP  
 - Überweisung durch niedergelassene Psychiater und Therapeuten

|                   | Kriterium                           | Operationalisierung                                                                                                                                                                                                                                                                                                                                            | Rating |
|-------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| K6.<br><br>(C.2.) | Art der Rekrutierung der Stichprobe | 1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege<br>2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege<br>3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung) | 9 C    |

**2.4 (7) Zentrale Patientenmerkmale**  
 (Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- 32% Männer, 68% Frauen, mittleres Alter 33 Jahre (SD ca. 8 Jahre)  
 - Dauer der Angststörung vor Einschluss: ca. 6-7 Jahre

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

Keine Angaben

|                    | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating |
|--------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| K7.<br><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | 2A16   |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
 (Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

Für EXP:  
 - Anzahl gescreenter und eingeschlossener Patienten und Dropouts unbekannt  
 - 75 Patienten ausgewertet

Für CCT und CCT+EXP:  
 - Keine Angaben über die Anzahl gescreenter Patienten  
 - 40 Patienten eingeschlossen und je 20 einer der Gruppen zugewiesen  
 - 13 Dropouts in Woche 1-4  
 - Stichprobe entsprechend dem Studiendesign wieder auf 40 Patienten aufgefüllt  
 - keine Dropouts zu den 3 Katamnesezeitpunkten

Insgesamt:  
 - 115 Patienten ausgewertet

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>2  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>9 |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A7<br>9  |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | 9<br>A3  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>9  |

## 2.7 (10) Vergleichbarkeit der Gruppen

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen:</p> <ul style="list-style-type: none"> <li>- Gruppen vergleichbar bzgl. der erhobenen sozialen und demographischen Faktoren und der Ausgangswerte in den erhobenen Kriterien</li> </ul> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer):</p> <ul style="list-style-type: none"> <li>- CCT+EXP zeitlich intensivere Behandlung als CCT (siehe 3.2 und 3.3)</li> <li>- nicht bekannt für EXP</li> </ul> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen):</p> <ul style="list-style-type: none"> <li>- event. geringer differentieller Einfluss zusätzlicher psychotroper Medikation während der Behandlungsphase (CCT - N = 3, CCT+EXP - N = 6)</li> <li>- nicht bekannt für EXP</li> </ul> <p>Vergleichbarkeit der Therapieadhärenz:</p> <p>Gegeben (Dropouts: 6 CCT, 7 CCT+EXP), nicht bekannt für EXP</p> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                        | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Rating   |
|------------------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K13.</b><br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | <p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p> | B9<br>1  |
| <b>K14.</b><br>(A.18.) | Drop-out-Analysen                                                                         | <p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>                                                                                                                       | A18<br>9 |

### 3. Intervention

**3.1 (14) Studiendesign /  
Behandlungsarme**

Therapiestudie mit zwei randomisierten Vergleichsgruppen und einer konsekutiven Vergleichsgruppe, 3 Behandlungsarme: 1. Klientenzentrierte Psychotherapie (CCT) mit zusätzlichem Expositionstraining und Hausaufgaben (EXP) 2. CCT 3. EXP

**3.2 (11) Behandlung Experimental-  
gruppe  
(Art, Dauer, Frequenz)**

CCT nach Manual für die Behandlung von Panik und Agoraphobie von Teusch & Finke (*Teusch L, Finke J; 1995; Psychotherapeut*) & „Occupational Therapy“  
- 1x/Woche Einzeltherapie, 3x/Woche Gruppentherapie

Expositionstherapie (EXP)  
- tägliches Einzelexpositionstraining umfasste Hausaufgaben nach dem Manual von Mathews et al. (1981)  
- Gruppenexpositionstraining unter therapeutischer Anleitung: 1x/Woche

- stationärer Aufenthalt für 10-14 Wochen (die letzten 4 Wochen optional tagesklinisch)

**3.3 (12) Behandlung Kontroll-  
gruppe  
(Art, Dauer, Frequenz)**

CCT (s.o.)  
- Psychogymnastik & kreative Elemente  
- Keine: direktiven Interventionen, Exposition, Hausaufgaben  
- 1x/Woche Einzeltherapie, 4x/Woche Gruppentherapie

- stationärer Aufenthalt für 10-14 Wochen (die letzten 4 Wochen optional tagesklinisch)

**3.4 (13) Evtl. weitere Kontroll-  
gruppen  
(Art, Dauer, Frequenz)**

EXP  
- nach Manual von Mathews et al. (1981) (s.o.)  
- genauer Ablauf/ Frequenz/ Dauer hier nicht angegeben

- ambulante Behandlung



|                | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                            | Rating  |
|----------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K15.<br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | 1) ausführliches Therapiemanual<br>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie<br>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens                                                                                                                                            | B3<br>1 |
| K16.<br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | 1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung<br>2) Ex post facto Beschreibung der Kontrollbedingungen<br>3) keine Beschreibung der Kontrollbedingung                                                                                                                                                                                            | 2 B4    |
| K17.<br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | 1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz<br>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab<br>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab | 3 B5    |
| K18.<br>(B.6.) | Manualtreue, Treatment Integrity                                                     | 1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt<br>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt<br>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen                                                                                                                       | B6<br>9 |
| K19.<br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | 1) Intervention wie in klinischer Alltagspraxis<br>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert<br>3) Intervention gegenüber klinischer Alltagspraxis stark verändert                                                                                                                                                                                | C9<br>4 |

|                | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K20.<br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantiellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br>9  |
| K21.<br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C9<br>6  |
| K22.<br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br>9 |
|                | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br>9 |
|                | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br>9 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

Für CCT und CCT+EXP:  
10-14 Wochen Behandlungszeitraum (Spielraum entsteht durch Ersetzen der Dropouts der Wochen 1-4), Messzeitpunkte bei Aufnahme in die Studie und Entlassung aus der Behandlung, drei Katamnese Untersuchungen (3, 6 und 12 Monate nach Entlassung), d.h. ca. 15-16 Monate Beobachtungszeitraum

Für EXP:  
Behandlungszeitraum nicht bekannt, Messzeitpunkte bei Aufnahme in die Studie und Entlassung aus der Behandlung, drei Katamnese Untersuchungen (nach 6 und normalerweise 12 Monaten (konnte scheinbar aber auch 1-4 Jahre nach Entlassung sein))

|                 | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating   |
|-----------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K23.<br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br>9  |
| K24.<br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10      |
| K25.<br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br>1 |

**3.6 (15) Anzahl der Zentren**

>1 (nicht genau gegeben)

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

Randomisierung (CCT & CCT+EXP), natürliche Stichprobe (EXP)

**3.8 (17) ggf. Concealment der Randomisierung**

Keine Angabe

|                | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                   | Rating  |
|----------------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K26.<br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                         | C8<br>2 |
| K27.<br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße (n>30/Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße n<30/Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br>2 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

- Anstieg der Bereitschaft, phobischen Situationen ausgesetzt zu werden (FSS-3, Resistenz)  
- Reduktion der Beeinträchtigung durch die Angst (FSS-4, Interferenz)

### 4.2 (21) Sekundäre Zielkriterien

keine

### 4.3 () Verblindung der Erhebung der Zielkriterien

Nein, für EXP nicht bekannt

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>2  |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | 1A8      |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | 1C9      |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | 3A9      |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10<br>9 |

**4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

- Gepaarte T-Tests zwischen Zeitpunkten für jede der drei Versuchsgruppen  
 - Univariate Varianzanalyse mit wiederholten Messungen (ANOVA) für jeden der 4 Zeitpunkte über 3 Gruppen für jedes der beiden Maße

|                 | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|-----------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K33.<br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | 2 A6     |
| K34.<br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| K35.<br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | 3 A14    |

**4.5 (22) Ergebnisse**

- in allen Vergleichsgruppen zwischen Prä- und Postmessung signifikante Verbesserung in beiden Zielkriterien (d.h. Resistenz gegen Exposition, Beeinträchtigung durch die Angststörung)

Vergleich der Gruppen zur Postmessung:  
 - nach EXP oder CCT+EXP größere Verbesserung im Kriterium Resistenz  
 - keine Unterschiede bzgl. der Beeinträchtigung

Follow-up (1-4 Jahre):  
 - alle Gruppen gleich bzgl. des Kriteriums Resistenz  
 - signifikante Verbesserung im Kriterium Beeinträchtigung nur in der CCT Bedingung

|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |

**4.6. (23) Ergebnisse zu unerwünschten Therapieeffekten**

Nicht erfasst oder nicht berichtet

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br>3 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

- Nachweis der Wirksamkeit von CCT, EXP und CCT+EXP bei Agoraphobie und Panikstörung

- kurz nach der Behandlung sind CCT+EXP und EXP der CCT im Kriterium Resistenz überlegen

- da zur 1-Jahres-Katamnese eine „internale Problembearbeitung“ stattgefunden habe, werden diese Coping Mechanismen nicht mehr gebraucht und die Wirksamkeit der Therapievarianten ist gleich, bezogen auf das Kriterium Beeinträchtigung ist CCT besser

- die spezifische Wirksamkeit von Exposition bei Panikstörung und Agoraphobie wird in Frage gestellt

|                       | Kriterium              | Operationalisierung                                                                    | Rating  |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|---------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br>1 |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C2<br>10 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C3<br>1  |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>2 |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                     |  |  |  |
|-------------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität<br>(Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität<br>(Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9)<br>(Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13)<br>(Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>1. wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>2. wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>3. wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                                  |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>1. die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>2. der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                       |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>1. das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>2. das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>3. bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>4. bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |  |



## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Die allgemeine methodische Qualität ist eingeschränkt (kann für ergänzte EXP Stichprobe nicht beurteilt werden).

Die Diagnosestellung erfolgt objektiv und reliabel. Es gibt 3 Katamnese-Erhebungen. Es werden reliable und valide Outcome-Verfahren genutzt.

Das Design erlaubt keinen Nachweis einer eventuellen Überlegenheit der CCT, da die Vergleichsgruppe CCT + Exposition bekommt. Es gibt keine Angaben über die Anzahl gescreener Patienten. Die statistische Power ist mäßig. Trotz relativ hoher Dropout-Raten, werden keine IIT- und Dropout-Analysen durchgeführt.

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

#CCT, CCT+EXP und EXP wirksam (d.h. Resistenz gegen Exposition, Beeinträchtigung durch die Angststörung sinken signifikant)

#Postmessung: nach EXP oder CCT+EXP größere Verbesserung im Kriterium Resistenz, Gruppen gleich im Kriterium Beeinträchtigung

#Follow-up: alle Gruppen gleich bzgl. des Kriteriums Resistenz, signifikante Verbesserung im Kriterium Beeinträchtigung nur in der CCT Bedingung

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Die interne Validität ist grundsätzlich gut. Die Gruppen (auch die externe EXP Vergleichsgruppe) sind zum Zeitpunkt prä und bzgl. sozio-demographischer Daten äquivalent. Die Äquivalenz der Therapiebedingungen ist allerdings nicht gegeben: Während CCT und CCT+EXP stationär oder teilstationär behandelt werden, erfahren Patienten der EXP Gruppe eine ambulante Behandlung.

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Für CCT und CCT+EXP: Die externe Validität bzgl. Setting und Frequenz ist eingeschränkt (stationär, 4-5x/ Woche) und nicht auf ein ambulantes Setting im dt. Versorgungssystem übertragbar.

### Stärken der Studie:

3 Katamnesezeitpunkte

### Schwächen der Studie:

Teil dieser Studie ist eine prospektive randomisierte Studie (siehe Teusch et al., 1997, PsychotherPsycho-som), ergänzt um eine weitere Vergleichsgruppe (nur Exposition) (Daten publiziert in Hand et al., 1986). Die Methodik in Teusch et al., 1999 ist nur verkürzt beschrieben (und auf Teusch et al., 1997 verweisend). Die Methodik bzgl. der zusätzlichen Versuchsgruppe (z.B. Einschlusskriterien) ist ohne zusätzliches Studium von Hand et al., 1986 schwer einzuschätzen.

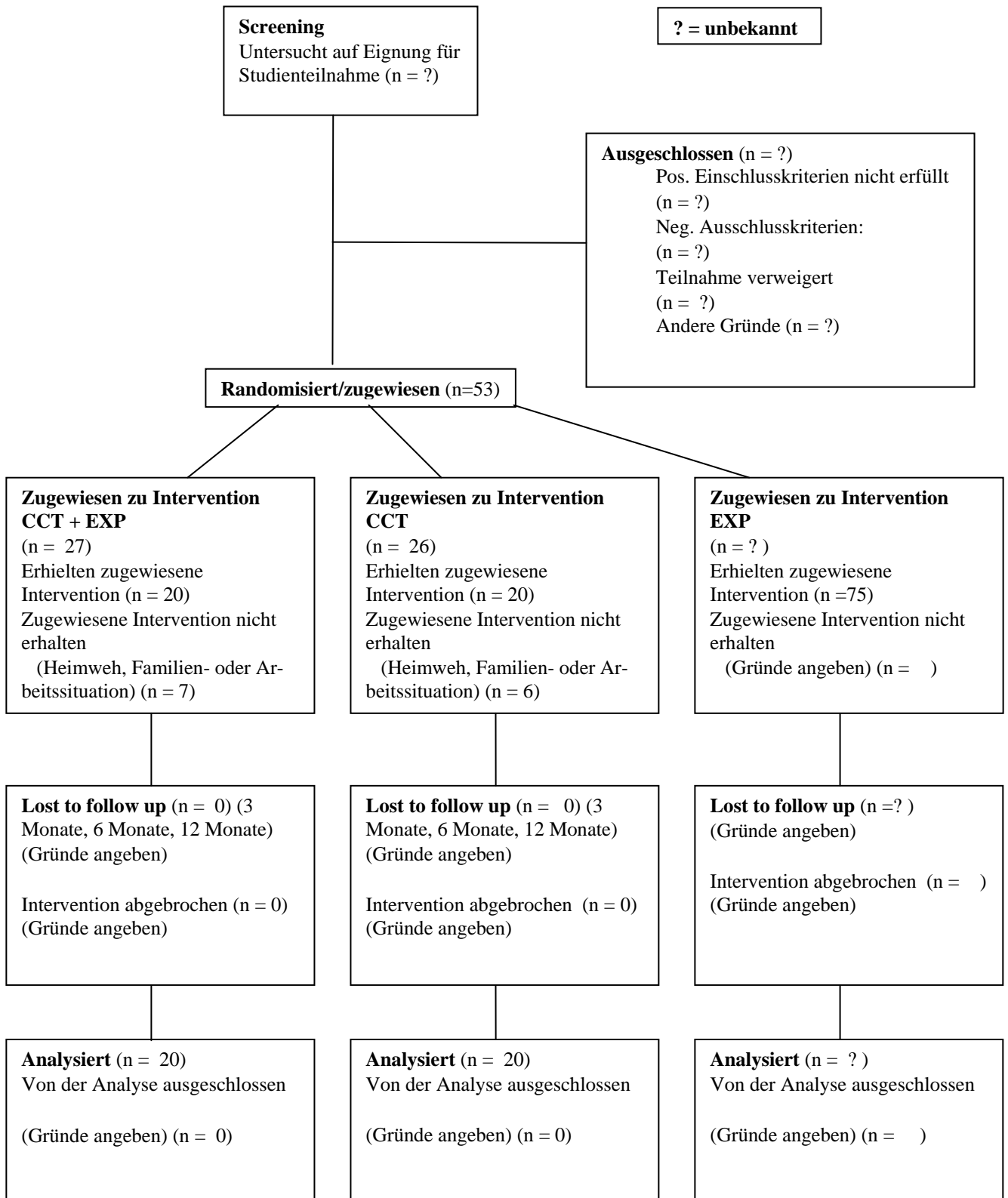
Die Zuweisung der Patienten erfolgte nur zum Teil randomisiert. Die Settingbedingungen sind nicht äquivalent (bei CCT und CCT+EXP stationäre oder teilstationäre Behandlung vs. bei EXP ambulante Behandlung).

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie ist aufgrund methodischer Einschränkungen nicht für einen Nachweis der Wirksamkeit der GT bei Angststörungen geeignet, liefert jedoch Hinweise auf eine Wirksamkeit der GT in einem stationären Setting.

Wirksamkeit der CCT weicht nicht von der einer reinen Expositionsbehandlung ab. Einschränkungen ergeben sich insbesondere aus den nicht vergleichbaren Rahmenbedingungen der Behandlung. Während EXP ambulant stattfand, fand CCT stationär statt.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

MT

**Studie:**

Autoren:

Teusch et al.

Titel:

The Benefit of an Insight-Oriented and Experiential Approach on Panic and Agoraphobia Symptoms: Results of a Controlled Comparison of Client-Centered Therapy Alone and in Combination with Behavioral Exposure

Quelle:

Psychother. Psychosom., 1997, 66: 293-301

## 1. Basisinformationen

### 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Klientenzentrierte Therapie (CCT) nach Manual für die Behandlung von Panik und Agoraphobie von Teusch & Finke (Teusch L, Finke J; 1995; *Psychotherapeut*) & „Occupational Therapy“

- Einzel- & Gruppentherapie
- Psychogymnastik & kreative Elemente
- Keine: direktiven Interventionen, Exposition, Hausaufgaben

### 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

- 1) Wirksamkeit von CCT
- 2) Differentielle Wirksamkeit von CCT + behaviorale Exposition

bezogen auf zentrale Angstsymptome (+ depressive Symptome) bei Agoraphobie mit Panikstörung

### 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

### 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

Agoraphobie und Panikstörung (DSM-III-R, No. 300.21)  
- klinische Diagnose (Therapeut), Überprüfung durch SCID (klinischer Psychologe)

|                      | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating |
|----------------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>K1.</b><br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | 1 A2   |
| <b>K2.</b><br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | 1 C1   |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

- Voraussetzung: „Agoraphobie und Panikstörung“ Hauptdiagnose  
- Ausschluss: psychotische Störungen, Drogen- oder Alkoholabhängigkeit

|                      | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating |
|----------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>K3.</b><br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 1 B1   |
| <b>K4.</b><br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | 1 B2   |
| <b>K5.</b><br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1 C3   |



**2.3 ( ) Setting/Patientenzugang**

Überweisung durch niedergelassene Psychiater und Therapeuten

|                          | Kriterium                           | Operationalisierung                                                                                                                                                                                                                                                                                                                                            | Rating |
|--------------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>K6.</b><br><br>(C.2.) | Art der Rekrutierung der Stichprobe | 1) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); keine Selektionseffekt aufgrund der Zugangswege<br>2) Patientenzugang durch gängige klinische Routinen (Überweisung, Primärzugang, etc); Selektionseffekte aufgrund der Zugangswege<br>3) Patientenzugang über Aufforderungen der Forschergruppe (z. B. Anzeigenwerbung) | 1 C    |

**2.4 (7) Zentrale Patientenmerkmale**  
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- 16 Männer, 24 Frauen, mittleres Alter 33 Jahre (SD ca. 9 Jahre)
- meistens mit Partner lebend (N = 28) oder Single/ allein lebend (N = 10)
- Schulabschluss mittl. Reife oder Abitur (N = 19), Schulabschluss niedriger oder nicht vorhanden (N = 21)
- meistens Lebensunterhalt selbst verdienend (N = 25)
- Sozioökonom. Status am häufigsten „Mittelklasse“ (N = 26) oder „Oberklasse“ (N = 11)
- Mittlere tägliche Arbeitszeit während letzter 4 Wochen: 2,8 h
  
- Erkrankungsdauer vor Einschluss: ca. 6-7 Jahre
- Dauer der erfolglosen Behandlung der Angststörung vor Einschluss: ca. 4 Jahre
- 87,5% der Patienten – pharmakologische Behandlung (Antidepressiva, Benzodiazepine oder Neuroleptika), meistens Mehrfachmedikation

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

Keine Angaben

|                           | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating |
|---------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>K7.</b><br><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | 3 A16  |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

- Keine Angaben über die Anzahl gescreenter Patienten
- 40 Patienten eingeschlossen und je 20 randomisiert einer der Gruppen zugewiesen
- 13 Dropouts in Woche 1-4
- Stichprobe entsprechend dem Studiendesign wieder auf 40 Patienten aufgefüllt
- keine Dropouts zu den 3 Katamnesezeitpunkten
- 40 Patienten ausgewertet

|                 | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating   |
|-----------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K8.<br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br>2  |
| K9.<br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br>3 |
| K10.<br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A7<br>1  |
| K11.<br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | 9<br>A3  |
| K12.<br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br>1  |

## 2.7 (10) Vergleichbarkeit der Gruppen

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen:<br/>- Gruppen vergleichbar bzgl. der erhobenen sozialen und demographischen Faktoren und der Ausgangswerte in den erhobenen Kriterien</p> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer):<br/>- CCT+EXP-Gruppe erfährt zeitlich intensivere Behandlung (siehe 3.2 und 3.3)</p> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen):<br/>- event. geringer differentieller Einfluss zusätzlicher psychotroper Medikation während der Behandlungsphase (CCT - N = 3, CCT+EXP - N = 6)</p> <p>Vergleichbarkeit der Therapieadhärenz:<br/>- ist gegeben (Dropouts: 6 CCT, 7 CCT+EXP), Gruppe wieder „aufgefüllt“</p> |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                        | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Rating   |
|------------------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K13.</b><br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | <p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p> | B9<br>1  |
| <b>K14.</b><br>(A.18.) | Drop-out-Analysen                                                                         | <p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>                                                                                                                       | A18<br>3 |

### 3. Intervention

**3.1 (14) Studiendesign /  
Behandlungsarme**

Therapiestudie mit randomisierter Vergleichsgruppe, 2 Behandlungsarme: 1. Klientenzentrierte Psychotherapie (CCT) mit zusätzlichem Expositionstraining und Hausaufgaben (EXP) 2. CCT

**3.2 (11) Behandlung Experimental-  
gruppe  
(Art, Dauer, Frequenz)**

CCT nach Manual für die Behandlung von Panik und Agoraphobie von Teusch & Finke (*Teusch L, Finke J; 1995; Psychotherapeut*) & „Occupational Therapy“  
- 1x/Woche Einzeltherapie, 3x/Woche Gruppentherapie

Expositionstherapie (EXP)  
- tägliches Einzelexpositionstraining umfasste Hausaufgaben nach dem Manual von Mathews et al. (1981)  
- Gruppenexpositionstraining unter therapeutischer Anleitung: 1x/Woche

- stationärer Aufenthalt für 10-14 Wochen (die letzten 4 Wochen optional tagesklinisch)

**3.3 (12) Behandlung Kontroll-  
gruppe  
(Art, Dauer, Frequenz)**

CCT (s.o.)  
- Psychogymnastik & kreative Elemente  
- Keine: direktiven Interventionen, Exposition, Hausaufgaben  
- 1x/Woche Einzeltherapie, 4x/Woche Gruppentherapie

- stationärer Aufenthalt für 10-14 Wochen (die letzten 4 Wochen optional tagesklinisch)

**3.4 (13) Evtl. weitere Kontroll-  
gruppen  
(Art, Dauer, Frequenz)**

keine

|                | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                            | Rating  |
|----------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K15.<br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | 1) ausführliches Therapiemanual<br>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie<br>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens                                                                                                                                            | B3<br>1 |
| K16.<br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | 1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung<br>2) Ex post facto Beschreibung der Kontrollbedingungen<br>3) keine Beschreibung der Kontrollbedingung                                                                                                                                                                                            | B4      |
| K17.<br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | 1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz<br>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab<br>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab | B5<br>2 |
| K18.<br>(B.6.) | Manualtreue, Treatment Integrity                                                     | 1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt<br>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt<br>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen                                                                                                                       | B6<br>3 |
| K19.<br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | 1) Intervention wie in klinischer Alltagspraxis<br>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert<br>3) Intervention gegenüber klinischer Alltagspraxis stark verändert                                                                                                                                                                                | C4<br>2 |

|                | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Rating   |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K20.<br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substantziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br>2  |
| K21.<br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                               | C6<br>1  |
| K22.<br>(C.7.) | Qualifikation der Behandler<br>(insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                     | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. Behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                  | C7a<br>9 |
|                | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                 | C7b<br>9 |
|                | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                 | C7c<br>9 |

**3.5 (19) Beobachtungszeitraum/  
Messzeitpunkte/Katamnese**

10-14 Wochen Behandlungszeitraum (Spielraum entsteht durch Ersetzen der Dropouts der Wochen 1-4), Untersuchung bei Aufnahme in die Studie und Entlassung aus der Behandlung, drei Katamnese Untersuchungen (3, 6 und 12 Monate nach Entlassung), d.h. ca. 15-16 Monate Beobachtungszeitraum

|                 | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating   |
|-----------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K23.<br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br>1  |
| K24.<br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10      |
| K25.<br>(B.11.) | Follow-up-Messung                                                      | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe (in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese                                                                                                                                                                                                           | B11<br>1 |

**3.6 (15) Anzahl der Zentren**

1

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

Randomisierung

**3.8 (17) ggf. Concealment der Randomisierung**

Keine Angabe

|                | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                              | Rating  |
|----------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K26.<br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br>3 |
| K27.<br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br>2 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

-Reduktion der Schwere der Panikstörung (SKID), der Agoraphobie (SKID & FSS-1), der Angst (HAMA) und der Depression (HAMD)

### 4.2 (21) Sekundäre Zielkriterien

- Reduktion anderer Phobien (FSS-2-1)  
 - Anstieg der Bereitschaft, phobischen Situationen ausgesetzt zu werden (FSS-3, Resistenz)  
 - Reduktion der Beeinträchtigung durch die Angst (FSS-4, Interferenz)

### 4.3 () Verblindung der Erhebung der Zielkriterien

nein

|                        | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K28.</b><br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>2  |
| <b>K29.</b><br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | 1A8      |
| <b>K30.</b><br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | 1C9      |
| <b>K31.</b><br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | 1A9      |
| <b>K32.</b><br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10<br>2 |

**4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

Jeweils für alle Messinstrumente:

- Gepaarte T-Tests zwischen Zeitpunkten gepoolt für beide Versuchsgruppen
- Multivariate Varianzanalyse mit wiederholten Messungen (MANOVA) über 5 Zeitpunkte und 2 Gruppen
- Varianzanalyse (ANOVA) über 2 Gruppen zu 4 Zeitpunkten mit der Prä-Messung als Kovariate

|                 | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|-----------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K33.<br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | 1 A6     |
| K34.<br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>2 |
| K35.<br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>3 |

**4.5 (22) Ergebnisse**

**CCT**

- zwischen Prä- und Postmessung signifikante Verbesserung (oder Verschwinden der Symptomatik) in allen Zielkriterien (d.h. bzgl. der Schwere der Panikstörung, der Agoraphobie, der Angst, der Depression, der Resistenz gegen Exposition und der Beeinträchtigung durch die Angststörung)
- zwischen Postmessung und 1-Jahres-Katamnese signif. Verbesserung (oder Verschwinden der Symptomatik) der Agoraphobie (im SKID, aber nicht im FSS), der Angst (HAMA) und der Depression (HAMD) (die Autoren berichten mehr Effekte, die aber nach Korrektur für multiple Tests wegfallen)

**CCT vs. CCT+EXP**

- nach 1 Jahr gibt es keine Unterschiede mehr zwischen den Gruppen
- zur Postmessung und den früheren Katamnese-Zeitpunkten (3-M, 6-M) ist CCT+EXP der CCT in wenigen Kriterien überlegen:
  1. höhere Bereitschaft, phobischen Situationen ausgesetzt zu werden (FSS-3) (Postmessung & 6-M-Katamnese)
  2. bessere agoraphobische Wirkung (3-M-Katamnese)

1-J-Katamnese - Verschwinden der Symptomatik  
 nach CCT: Panik (40% d. Patienten), Agoraphobie (25%)  
 nach CCT+EXP: Panik (45% ...), Agoraphobie (40%)



|                        | Kriterium                                                                                                                                          | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten                                             | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet                                                                                                                                                                                                                                                                                                                                             | A11<br>1 |
| <b>K37.</b><br>(B.12.) | Erzielte Veränderungen auf den primären und sekundären Zielkriterien im Vergleich zur Kontrollgruppe (Signifikanz, Größe und Relevanz der Effekte) | 1) vollständige Darstellung der erzielten Veränderungen auf den Zielkriterien inklusive der Signifikanz, Größe der Effektmaße und Ausmaß der klinisch relevanten Zielerreichung im Vergleich zur Kontrollgruppe<br>2) Darstellung des Behandlungsergebnisses nur durch Veränderungs- oder Zielerreichungsmaße oder beides ist im Vergleich zur Kontrollgruppe bei einigen Kriterien unvollständig<br>3) weitgehend unvollständige oder inadäquate Darstellung der Outcome-Kriterien im Vergleich zur KG | B12<br>1 |

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

keine

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating   |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br>1 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

- Nachweis der Wirksamkeit von CCT und CCT+EXP bei Agoraphobie und Panikstörung  
 - in den ersten 6 Monaten der Follow-up Periode ist CCT+EXP der CCT in einzelnen Aspekten überlegen  
 - diese Unterschiede sind den Autoren nach auf durch die EXP gebotenen Mechanismen des Umgangs mit der Angst zurückzuführen  
 - da zur 1-Jahres-Katamnese eine „internale Problembearbeitung“ stattgefunden habe, werden diese Coping Mechanismen nicht mehr gebraucht und die Wirksamkeit beider Therapievarianten ist gleich  
 - die spezifische Wirksamkeit von Exposition und kognitiven Treatments bei Panikstörung und Agoraphobie wird auf Grundlage der Ergebnisse der 1-Jahres-Katamnese vorsichtig in Frage gestellt

|                       | Kriterium              | Operationalisierung                                                                    | Rating |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|--------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1     |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | 2  |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating                       |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C <sub>2</sub> <sup>10</sup> |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C <sub>3</sub>               |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C <sub>1</sub> <sup>12</sup> |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C <sub>2</sub> <sup>13</sup> |

## 7. Gesamtbewertung der Studie

### Bewertungsergebnisse zur Studienqualität anhand des Kriterienkatalogs

|                                                                                  |  |  |  |
|----------------------------------------------------------------------------------|--|--|--|
| Durchschnittswert für A) Allgemeine methodische Qualität (Summe / 18)            |  |  |  |
| Durchschnittswert für B) Interne Validität (Summe / 12)                          |  |  |  |
| Durchschnittswert für C) Externe Validität (C1-C9) (Summe / 11)                  |  |  |  |
| Durchschnittswert für C) Externe Validität: Praxistransfer (C10-C13) (Summe / 4) |  |  |  |

### Bewertung der Studie unter Berücksichtigung der Behandlungsergebnisse

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p>Kriterien für '<b>A. Allgemeine methodische Qualität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>wenn die Kriterien A1 (Manipulation) und A2 (Diagnosestellung) und A8 (Messung Zielkriterien) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn das Kriterium C1 (klinische Stichprobe) mit 1 oder 2 bewertet wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>wenn der Durchschnitt der Bewertung aller A-Kriterien (A1-A18) kleiner als 2,25 ist.</li> </ol>                                                                                               |  |
| <p>Kriterien für '<b>B. Interne Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>die Kriterien B8 (Gruppenzuweisung) und B12 (Darstellung der Outcome-Maße) auf Stufe 1 oder 2 eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>der Durchschnitt der Bewertung aller B-Kriterien (B1-B12) kleiner als 2,25 ist.</li> </ol>                                                                                                                                                                                                                                                                 |  |
| <p>Kriterien für '<b>C. Externe Validität</b>' sind erfüllt, wenn:</p> <ol style="list-style-type: none"> <li>das Kriterium B12 (Darstellung der Outcome-Maße) mit 1 oder 2 bewertet wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>das Kriterium C9 (Outcome-Messung) auf Stufe 2 oder besser eingeschätzt wurden</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C1-C9 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde</li> </ol> <p>und</p> <ol style="list-style-type: none"> <li>bei den Kriterien C10-C13 ein Wert von durchschnittlich höchstens 2,25 erreicht wurde.</li> </ol> |  |
| <p>Die <b>Ergebnisse der Studie</b> belegen die Wirksamkeit der Methode, wenn: D1 oder D2 oder D3 mit 1 bewertet wurden.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |  |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

Die allgemeine methodische Qualität ist eingeschränkt.

Die Diagnosestellung erfolgt objektiv und reliabel. Es gibt 3 Katamnese-Erhebungen. Es werden reliable und valide Outcome-Verfahren genutzt.

Das Design erlaubt keinen Nachweis einer eventuellen Überlegenheit der CCT, da die Vergleichsgruppe CCT + Exposition bekommt. Es gibt keine Angaben über die Anzahl gescreener Patienten. Die statistische Power ist gering. Trotz relativ hoher Dropout-Raten, werden keine IIT- und Dropout-Analysen durchgeführt.

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

#CCT und CCT+EXP führen zu nachhaltigen Therapieeffekten

#keine Nebenwirkungen

#die Angstsymptomatik verbessert sich deutlich oder verschwindet, zwischen den Messzeitpunkten wird eine stetige Besserung bzgl. der Symptomatik verzeichnet.

#die Befunde sind konsistent

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

Die interne Validität ist gut. Es findet eine angemessene Randomisierung statt, die Vergleichsgruppen sind äquivalent.

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

Die externe Validität bzgl. Setting und Frequenz ist eingeschränkt (stationär, 4-5x/ Woche) und nicht auf ein ambulantes Setting im dt. Versorgungssystem übertragbar.

### Stärken der Studie:

3 Katamnesezeitpunkte

### Schwächen der Studie:

Die Wahl der Versuchsgruppen (CCT vs. CCT + EXP) erlaubt keine Beurteilung der spezifischen Wirksamkeit der CCT. In beiden Gruppen wurde jeweils stationär ein Bündel verschiedener Maßnahmen realisiert, so dass nicht beurteilt werden kann, worauf der Effekt genau zurückzuführen ist. Die statistische Power ist gering.

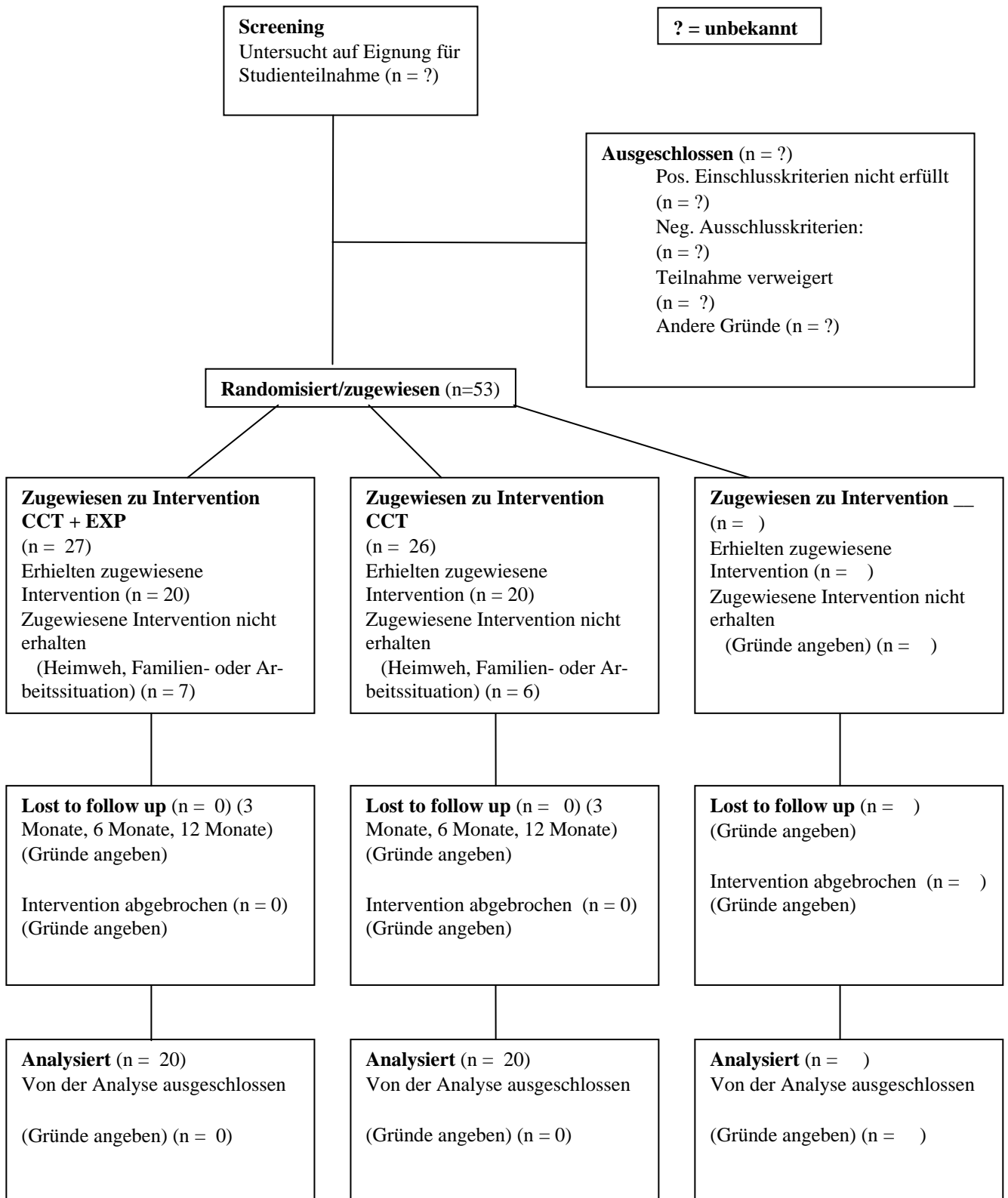
Die externe Validität ist durch mangelnde Übertragbarkeit von Therapie Setting und Frequenz eingeschränkt.

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Die Studie ist aufgrund methodischer Einschränkungen nicht für einen Nachweis der Wirksamkeit der GT bei Angststörungen geeignet. Das Design lässt Aussagen über die Wirksamkeit der GT nur auf Evidenzstufe IV zu, da es für die hier interessierende Fragestellung keine geeignete Vergleichsgruppe gibt. Sie liefert jedoch Hinweise auf eine Wirksamkeit der GT in einem stationären Setting.

Die spezifische Wirksamkeit kann nicht beurteilt werden, da es 1. keine echte Kontrollgruppe gibt und 2. stationär in jeder Gruppe ein Bündel von Maßnahmen realisiert wurde. Die allgemeine methodische Qualität und die externe Validität bzgl. Therapiesetting, -frequenz sind eingeschränkt.

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**



## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.



## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

# Protokoll- und Beurteilungsbogen für die Auswertung von Psychotherapiestudien

## Bewertung von Studien zur Wirksamkeit der Gesprächspsychotherapie

*Bitte nur die grau unterlegten Felder beschreiben! Bitte alle grauen Felder ausfüllen.*

*Falls ein Aspekt der Studienqualität anhand des Kriterienkatalogs nicht zu beurteilen ist,  
bitte in der Regel „9“ eintragen.*

**Projekt:** Beurteilung der Methode /des Verfahrens:

Gesprächspsychotherapie

**Beurteiler:**

JS

**Studie:**

Autoren:

Wuchner

Titel:

Behandlungsergebnisse von Borderline-Patienten nach klientenzentrierter Psychotherapie

Quelle:

Internationale Hochschulschriften

# 1. Basisinformationen

## 1.1 () Benennung der Vorgehensweise (Verfahren/Methode/Interventionstechnik)

Klientenzentrierte Gruppenpsychotherapie (GT)

## 1.2 (6) Fragestellung / Zielsetzung der Studie / Forschungshypothese

- Veränderung von Patienten mit Borderline-Persönlichkeitsstörung (PS) nach GT im Verhältnis zu neurotisch-depressiven und schizophrenen Patienten?
- Hypothesen:
- Verlauf der Entwicklung unterscheidet sich im borderlinespezifischen Erleben und Verhalten, in allgemeiner Psychopathologie, in Veränderungs- und Persönlichkeitsdimensionen, in der psychosozialen Integration der Patienten

## 1.3 (2) Design

- Therapiestudie mit randomisierter Vergleichsgruppe
- Therapiestudie mit nicht-randomisierter Vergleichsgruppe
- Therapiestudie mit Vergleichen über Zeit und Ort (z. B. historische Kontrollen)
- Fall-Kontroll-Studien
- Kohorten-Studien
- Therapiestudie ohne Vergleichsgruppen (auch „Vorher-Nachher-Studien“)
- Fallserie
- Fallbericht / Kasuistik (case report)
- Nicht eindeutig zuzuordnen:

## 1.4 (3) Evidenzstufe

- Ib: Randomisierte klinische Studien
- IIb: Prospektiv vergleichende Kohortenstudien
- III: Retrospektiv vergleichende Studien
- IV: Fallserien und nicht-vergleichende Studien
- V: Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte u. a.; nicht mit Studien belegte Meinungen anerkannter Experten, Bericht von Expertenkomitees und Konsensuskonferenzen.

## 2. Population

### 2.1 (5) Indikation(sbereich) (inkl. diagnostisches Vorgehen)

B orderline-PS  
 o Diagnostik mit „Diagnostischem Interview für Borderline-Patienten“ und Rohrschachtest von unabhängigen Beurteilern

|               | Kriterium                                                                             | Operationalisierung                                                                                                                                                                                                                                                                                                                                             | Rating  |
|---------------|---------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K1.<br>(A.2.) | Objektive und reliable Diagnosestellung (mittels (teil-) standardisierter Interviews) | 1) Diagnosestellung mittels strukturiertem klinischen bzw. voll standardisiertem Interview (z. B. SKID, DIPS)<br>2) Diagnosestellung mittels Diagnosechecklisten oder nachvollziehbarem klinischen Urteil<br>3) keine adäquate Diagnosestellung                                                                                                                 | A2<br>1 |
| K2.<br>(C.1.) | Stichprobe von Patienten mit Störungen mit Krankheitswert                             | 1) ausschließlich Patienten mit Störung mit Krankheitswert (z. B. ICD-, DSM-Diagnosen)<br>2) Stichprobe von Patienten mit wahrscheinlicher klinischer Störung (z. B. Menschen nach Trauma) oder bis zu (maximal) 20% der Patienten mit lediglich erhöhter Symptomausprägung, z. T. subklinisch<br>3) Patienten ohne identifizierbare Störung mit Krankheitswert | C1<br>1 |

### 2.2 (7) Relevante Ein- und Ausschlusskriterien

o IN: eindeutige Diagnose Borderline/Neurose/Psychose

|               | Kriterium                                                                                | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating  |
|---------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| K3.<br>(B.1.) | Spezifizierung der Einschluss- und Ausschlusskriterien                                   | 1) eindeutige Spezifizierung der Ein- und Ausschlusskriterien<br>2) Ein- oder Ausschlusskriterien teilweise uneindeutig beschrieben<br>3) Ein- und Ausschlusskriterien sind nicht eindeutig definiert                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | B1<br>1 |
| K4.<br>(B.2.) | Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden | 1) Die Ein- und Ausschlusskriterien sind sämtlich klar operationalisiert und werden mittels valider Methoden erfasst (z. B. komorbide Störungen als Ausschluss werden mittels strukturiertem klinischen Interview erfasst; Ausschlusskriterien beziehen sich auf eindeutig objektivierbare Merkmale wie Alter, Geschlecht etc.)<br>2) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist teilweise eingeschränkt (z. B. komorbide Störungen als Ausschluss werden mittels globalem klinischen Urteil eingeschätzt) und wirkt sich jedoch nur in geringem Umfang auf die Zusammensetzung der Stichprobe aus.<br>3) Die Validität der Erhebungen von Teilen der Ein- oder Ausschlusskriterien ist deutlich eingeschränkt und wirkt sich differenziell auf die Zusammensetzung der Behandlungsgruppen aus. | B2<br>1 |
| K5.<br>(C.3.) | Selektivität der Stichprobe aufgrund der Ausschlusskriterien                             | 1) keine Selektionseffekte aufgrund der Ausschlusskriterien: Einschluss aller Patienten<br>2) mittlere Selektionseffekte aufgrund der Ausschlusskriterien (z. B. Ausschluss einiger epidemiologisch relevanter komorbider Störungen)<br>3) deutliche Selektionseffekte aufgrund der Ausschlusskriterien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | C3<br>1 |

### 2.3 () Setting/Patientenzugang

Rekrutierung durch Forschergruppe

|  | Kriterium | Operationalisierung | Rating |
|--|-----------|---------------------|--------|
|--|-----------|---------------------|--------|

|  |  |  |          |
|--|--|--|----------|
|  |  |  | <b>3</b> |
|--|--|--|----------|

**2.4 (7) Zentrale Patientenmerkmale**  
(Alter, Geschlecht, Sozioökonomischer Status, Ethnizität, Komorbidität, Chronizität, Schweregrad)

- 29 (Range 18-53)
- 16 männlich, 27 weiblich
- Borderline: 85,7% ledig, 8 von 14 Abitur; Schizophrenie: 61,5% ledig, 2 von 13 Abitur; Depression: 68,7% ledig, 3 von 13 Abitur
- keine Angaben
- keine Angaben
- keine Angaben
- keine Angaben

**2.5 (8) Fallzahlplanung, inkl. geplanter Fallzahl**

keine Angaben

|                       | Kriterium                                                                      | Operationalisierung                                                                                                                                                          | Rating          |
|-----------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <b>K7.</b><br>(A.16.) | Statistische Power der Vergleiche bei Vergleich mit bewährter Treatment-Gruppe | 1) adäquate statistische Power der Vergleiche<br>2) Eingeschränkte Power der statistischen Vergleiche (.50-.80)<br>3) unzureichende statistische Power der Vergleiche (<.50) | A16<br><b>2</b> |

**2.6 (9) Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten**  
(Drop-outs, IIT-Analyse, PP-Analyse, Darstellung des Patientenflusses nach CONSORT-Diagramm)

43 Patienten eingeschlossen; 14 Borderline, 16 Depression (davon 7 aus Lübeck), 13 Schizophrenie (davon 6 aus Lübeck)

|                        | Kriterium                                                                       | Operationalisierung                                                                                                                                                                                                                                                | Rating          |
|------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <b>K8.</b><br>(A.5.)   | Stichprobengröße pro Gruppe                                                     | 1) n pro Gruppe > 30<br>2) n pro Gruppe 10-30<br>3) n pro Gruppe < 10<br>(Abweichungen von diesem Schema sind bei sehr seltenen Erkrankungen möglich. Maßgeblich ist die kleinste Vergleichsgruppe)                                                                | A5<br><b>2</b>  |
| <b>K9.</b><br>(A.15.)  | Intention to treat – Analysen durchgeführt                                      | 1) ITT-Analysen durchgeführt<br>2) Keine ITT-Analysen bei geringem Risiko für attrition bias<br>3) Keine ITT-Analysen bei deutlichem Risiko für einen attrition bias                                                                                               | A15<br><b>2</b> |
| <b>K10.</b><br>(A.17.) | Vollständige Beschreibung der Drop-Outs                                         | 1) vollständige Beschreibung aller Drop-outs, inkl. Gründe und Zeitpunkt des Drop-out<br>2) unvollständige Angaben zu Gründen oder Zeitpunkten des Drop-out<br>3) keine Beschreibung der Drop-outs oder definitive Angabe der Anzahl der drop-outs in einer Gruppe | A17<br><b>3</b> |
| <b>K11.</b><br>(A.3.)  | Höhe der Drop-out-Quote zu Behandlungsende (sofern nicht Erfolgskriterium)      | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A3<br><b>9</b>  |
| <b>K12.</b><br>(A.4.)  | Höhe der Studien-Drop-outs zur Katamnese (falls Katamneseerhebung durchgeführt) | 1) deutlich besser als in Studien mit vergleichbaren Patientengruppen<br>2) Drop-out-Quote vergleichbar mit Studien entsprechender Patientengruppen<br>3) deutlich schlechter als in Studien mit vergleichbaren Patientengruppen                                   | A4<br><b>9</b>  |

**2.7 (10) Vergleichbarkeit der Gruppen**

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Vergleichbarkeit hinsichtlich prognostisch relevanter Variablen</p> <ul style="list-style-type: none"> <li>o Gruppen unterscheiden sich in borderline-spezifischer Symptomausprägung</li> </ul> <p>Vergleichbarkeit hinsichtlich Intensität der Behandlung (Zeit, Dauer)</p> <ul style="list-style-type: none"> <li>o Gruppentherapie war auf die Behandlung der Borderline-PS ausgerichtet – die anderen Patienten erhielten also keine störungsspezifische Intervention</li> </ul> <p>Vergleichbarkeit weiterer Einflussfaktoren während der Behandlung (z. B. Pharmakotherapie, weitere Behandlungen)</p> <ul style="list-style-type: none"> <li>o weitere Behandlungen (psychotherapeutisch, pharmakologisch u.a.) waren nicht ausgeschlossen</li> </ul> <p>Vergleichbarkeit der Therapieadhärenz</p> <ul style="list-style-type: none"> <li>o ja, da überall der gleiche Therapeut</li> </ul> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                 | Kriterium                                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Rating              |
|-----------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| K13.<br>(B.9.)  | Vergleichbarkeit der Gruppen zur Baseline in Hinblick auf prognostisch relevante Merkmale | <p>1) Weder statistisch noch klinisch relevante Unterschiede zwischen den Gruppen hinsichtlich prognostisch relevanter oder potentiell konfundierender Variablen</p> <p>2) Vergleichbarkeit hinsichtlich der meisten prognostisch relevanten Merkmale ist weitgehend gegeben; signifikante Unterschiede hinsichtlich relevanter prognostischer Merkmale zwischen den Gruppen werden statistisch angemessen kontrolliert</p> <p>3) Keine angemessene Überprüfung der Vergleichbarkeit oder Vergleichsgruppen unterscheiden sich erheblich hinsichtlich mehrerer prognostisch relevanter Merkmale und eine angemessene statistische Kontrolle des Einflusses dieser Merkmale fehlt</p> | <p>B9</p> <p>3</p>  |
| K14.<br>(A.18.) | Drop-out-Analysen                                                                         | <p>1) Drop-out-Analysen unter Berücksichtigung der wichtigsten prognostischen Faktoren durchgeführt; keine signifikanten Unterschiede zwischen den Vergleichsgruppen</p> <p>2) Drop-out-Analysen unter Berücksichtigung wichtigster prognostischer Faktoren durchgeführt; trotz sign. Unterschiede ist die Validität der Ergebnisse nicht wesentlich eingeschränkt</p> <p>3) Drop-out-Analysen nicht oder unter Vernachlässigung relevanter prognostischer Merkmale durchgeführt; Drop-out-Analysen stellen die Validität der Ergebnisse deutlich in Frage</p>                                                                                                                       | <p>A18</p> <p>3</p> |

### 3. Intervention

**3.1 (14) Studiendesign /  
Behandlungsarme**

1. Behandlungsgruppe Borderline-PS
2. Behandlungsgruppe Neurose
3. Behandlungsgruppe Psychose

**3.2 (11) Behandlung Experimental-  
gruppe  
(Art, Dauer, Frequenz)**

1. Borderline
  - o Teilnahme an ambulanter Gruppenpsychotherapie (zusammen mit Neurose- und Psychose-Patienten in geschlossener Gruppe) (100h 2x wöchentlich in rund einem Jahr)
  - o männlicher Therapeut und meistens weibliche Ko-Therapeutin

**3.3 (12) Behandlung Kontroll-  
gruppe  
(Art, Dauer, Frequenz)**

2. Neurose
  - o teilweise Teilnahme an ambulanter Gruppenpsychotherapie (zusammen mit Borderline- und Psychose-Patienten in geschlossener Gruppe) (100h 2x wöchentlich in rund einem Jahr), teilweise stationäre Behandlung
  - o männlicher Therapeut und meistens weibliche Ko-Therapeutin

**3.4 (13) Evtl. weitere Kontroll-  
gruppen  
(Art, Dauer, Frequenz)**

3. Psychose
  - o teilweise Teilnahme an ambulanter Gruppenpsychotherapie (zusammen mit Borderline- und Neurose-Patienten in geschlossener Gruppe) (100h 2x wöchentlich in rund einem Jahr), teilweise stationäre Behandlung
  - o männlicher Therapeut und meistens weibliche Ko-Therapeutin

|                       | Kriterium                                                                            | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                          | Rating      |
|-----------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K15.</b><br>(B.3.) | Operationale Definition der Interventionen (Experimental- und ggfls. Kontrollgruppe) | <ol style="list-style-type: none"> <li>1) ausführliches Therapiemanual</li> <li>2) Therapiebeschreibung mit Spezifikation der wesentlichen Bestandteile der Therapie</li> <li>3) Die Intervention ist nicht klar beschrieben, beschränkt sich auf die Benennung des Therapieverfahrens</li> </ol>                                                                                                                                            | B3<br><br>3 |
| <b>K16.</b><br>(B.4.) | Operationale Definition der Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) Prospektive Festlegung und umfassende Beschreibung der Kontrollbedingung</li> <li>2) Ex post facto Beschreibung der Kontrollbedingungen</li> <li>3) keine Beschreibung der Kontrollbedingung</li> </ol>                                                                                                                                                                                            | B4<br><br>1 |
| <b>K17.</b><br>(B.5.) | Strukturelle Äquivalenz bei Kontrollbedingungen                                      | <ol style="list-style-type: none"> <li>1) hinsichtlich des Umfangs an therapeutischer Zuwendung und der Settingbedingungen in der KG besteht Äquivalenz</li> <li>2) der Umfang der therapeutischen Zuwendung in der KG ist reduziert, die Settingbedingungen weichen von der IG ab</li> <li>3) der Umfang der therapeutischen Zuwendung in der KG ist deutlich reduziert, die Settingbedingungen weichen wesentlich von der IG ab</li> </ol> | B5<br><br>2 |
| <b>K18.</b><br>(B.6.) | Manualtreue, Treatment Integrity                                                     | <ol style="list-style-type: none"> <li>1) Manualtreue/Treatmentintegrität durch externe Beobachter (z. B. videogestützt) belegt</li> <li>2) Manualtreue/Treatmentintegrität durch Fragebögen belegt</li> <li>3) keine Maßnahmen zum Monitoring der Manualtreue oder Hinweise auf substantielle Abweichungen</li> </ol>                                                                                                                       | B6<br><br>3 |
| <b>K19.</b><br>(C.4.) | Klinische Repräsentativität der Intervention hinsichtlich Vorgehen und Dauer         | <ol style="list-style-type: none"> <li>1) Intervention wie in klinischer Alltagspraxis</li> <li>2) Intervention gegenüber klinischer Alltagspraxis teilweise verändert</li> <li>3) Intervention gegenüber klinischer Alltagspraxis stark verändert</li> </ol>                                                                                                                                                                                | C4<br><br>2 |

|                       | Kriterium                                                                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Rating    |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>K20.</b><br>(B.7.) | Zulässigkeit, Dokumentation und Analyse des Einflusses begleitender nicht-randomisierter Interventionen (insbesondere Pharmakotherapie) | 1) Ausschluss begleitender nicht-randomisierter Interventionen<br>2) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch detailliert dokumentiert und die Analysen weisen auf keinen substanziellen, differenziellen Einfluss der begleitenden Interventionen hin<br>3) begleitende nicht-randomisierte Interventionen sind zulässig, werden jedoch nicht dokumentiert oder die Analysen weisen auf eine differenzielle Inanspruchnahme von begleitenden Interventionen und deren Einfluss auf das Behandlungsergebnis hin | B7<br>3   |
| <b>K21.</b><br>(C.6.) | Zulässigkeit begleitender Interventionen (z. B. Pharmakotherapie)                                                                       | 1) keine Einschränkungen<br>2) begleitende in der Routine-Praxis übliche Interventionen teilweise ausgeschlossen<br>3) alle begleitenden in der Routine-Praxis üblichen Interventionen ausgeschlossen                                                                                                                                                                                                                                                                                                                                              | C6<br>1   |
| <b>K22.</b><br>(C.7.) | Qualifikation der Behandler (insb. der zu bewertenden Therapiemethode)<br>a) Klinische Tätigkeit der Therapeuten                        | 1) Therapeuten sind praktizierende Kliniker<br>2) klinische Forscher, die überwiegend Forschung betreiben und seltener auch Pat. behandeln; Ausbildungskandidaten)<br>3) keine Kliniker oder Kliniker, die keine Psychotherapeuten sind (z. B. Allgemeinmediziner)                                                                                                                                                                                                                                                                                 | C7a<br>1  |
|                       | b) Breite der Klinischen Tätigkeit der Therapeuten (Problemheterogenität)                                                               | 1) Therapeut behandelt Pat. mit verschiedenen Problemen innerhalb und außerhalb der Studie<br>2) Therapeut behandelt überwiegend Patienten mit ähnlichen Problemen innerhalb und außerhalb der Studie<br>3) Therapeut behandelt nur Pat. mit ähnlichen Problemen innerhalb und außerhalb der Studie (z. B. Therapeut, der nur Schmerzpatienten in einer Schmerzklinik behandelt; Psychotherapeuten, die außerhalb der Studie keine Pat. behandeln).                                                                                                | C7b<br>1  |
|                       | c) Spezifisches Training der Psychotherapeuten in einer Behandlungsmethode für die Studie                                               | 1) kein spezifisches Training für die Studie (z. B. Therapeuten wenden die von ihnen üblicherweise angewendete Therapie an)<br>2) kurzes Training für die Studie / intensives Training einiger Therapeuten<br>3) intensives Training vor der Studie                                                                                                                                                                                                                                                                                                | C7c<br>9* |

### 3.5 (19) Beobachtungszeitraum/ Messzeitpunkte/Katamnese

- Beobachtungszeitraum: frühe 80er Jahre
- MZPe:
  - Teilstichprobe Hamburg für Borderline-Patienten wurde Post und katamnestisch untersucht – Katamnese-Werte wurden berücksichtigt (durschn. 3,6 Jahre nach Behandlung)
  - Teilstichprobe Hamburg für Kontrollgruppe (Neurose, Psychose) wurde Post und katamnestisch untersucht – Post-Werte wurden berücksichtigt (durschn. 4 Jahre nach Behandlung)
  - Teilstichprobe Lübeck für Kontrollgruppe (Neurose, Psychose) wurde nur katamnestisch untersucht – Katamnese-Werte berücksichtigt (durschn. 4 Jahre nach Behandlung)

|                        | Kriterium                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Rating    |
|------------------------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>K23.</b><br>(C.5.)  | Art des Therapie-Monitorings (Einfluss auf Therapeutenverhalten)       | 1) keine Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (z. B. durch Therapie-Supervision; Ausnahmen: Audio- oder Video-Aufzeichnungen ohne Feedback an Therapeuten, nur zur späteren Auswertung sind erlaubt)<br>2) mittlere Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (gelegentliche Rückmeldungen an Therapeuten)<br>3) starke Veränderung des Therapeutenverhaltens durch Therapie-Monitoring (durch kontinuierliche Rückmeldungen) | C5<br>1   |
| <b>K24.</b><br>(B.10.) | Definition der Messzeitpunkte (Prospektive Messung; Follow-up-Messung) | 1) mehrere vorab festgelegte Messzeitpunkte über den Therapieverlauf incl. prä-post-Messungen und follow-up<br>2) ausschließlich prä-post-Messung<br>3) ausschließlich Post-Messung                                                                                                                                                                                                                                                                                            | B10<br>2* |



|                        |                   |                                                                                                                                                                                                                                                                         |              |
|------------------------|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K25.</b><br>(B.11.) | Follow-up-Messung | 1) zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung der Stichprobe<br>(in Ausnahmefällen mind. 6 Monate; in der Regel mind. 12 Monate)<br>2) Katamnese mit fraglich angemessenem Zeitraum bzw. niedriger Ausschöpfung der Stichprobe<br>3) keine Katamnese | B11<br><br>1 |
|------------------------|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|

**3.6 (15) Anzahl der Zentren**

**3.7 (16) Allokation zu Interventionsgruppen** (Randomisierung, self-selection, ...)

**3.8 (17) ggf. Concealment der Randomisierung**

|                       | Kriterium                                                                               | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                              | Rating      |
|-----------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>K26.</b><br>(C.8.) | Repräsentativität der patientenseitigen Freiheit hinsichtlich der Wahl der Intervention | 1) Patienten entscheiden sich selbst für eine der angebotenen Therapieformen<br>2) ein Teil der Patienten (z. B. Randomisierungswillige) wird der Therapie zufällig zugewiesen<br>3) alle Patienten werden der Therapie (zufällig) zugewiesen                                                                                                                                                                                                    | C8<br><br>1 |
| <b>K27.</b><br>(B.8.) | Gruppenzuweisung                                                                        | 1) angemessene Randomisierung (inkl. Cluster-Randomisierung[1]) bei ausreichender Stichprobengröße ( $n > 30$ /Gruppe[2]), die das Gelingen der Randomisierung hinsichtlich bekannter und unbekannter (nicht erfasster) prognostisch relevanter Merkmale sicherstellt<br>2) Parallelisierung oder teilweise randomisiert oder quasi-randomisiert oder Stichprobengröße $n < 30$ /Gruppe<br>3) keine randomisierte oder parallelisierte Zuweisung | B8<br><br>3 |

## 4. Ergebnisse

### 4.1 (20) Primäre Zielkriterien

- Fremdbeurteilungen:
  - ICD-9
  - Diagnostisches Interview für Borderline-Patienten (DIB)
  - Global Assessment Scale (GAS)
- Selbstbeurteilungen:
  - Borderline-Syndrom-Index
  - Paranoid-Depressivitäts-Skala (PDS)
  - Veränderungsfragebogen des Erlebens und Verhaltens (VEV)
  - Freiburger Persönlichkeitsinventar (FPI)
  - Gießen-Test (Selbst + Ideal)
- Biographische und Sozialdaten

### 4.2 (21) Sekundäre Zielkriterien

/

### 4.3 () Verblindung der Erhebung der Zielkriterien

keine Angaben

|                 | Kriterium                                                                                                                                    | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                       | Rating   |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K28.<br>(A.7.)  | a priori Definition der primären und sekundären Zielkriterien                                                                                | 1) a priori definiert unter Berücksichtigung der Wertigkeit<br>2) a priori Nennung der Zielkriterien ohne Differenzierung in primäre und sekundäre Zielkriterien<br>3) a posteriori Definition der Zielkriterien; keine Angaben zu Zielkriterien,                                                                                                                                                                         | A7<br>2  |
| K29.<br>(A.8.)  | Reliable und valide Messung zumindest der primären Zielkriterien                                                                             | 1) reliable und valide Outcome-Verfahren<br>2) nur eingeschränkte Reliabilität und/oder Validität der Messverfahren<br>3) Reliabilität und Validität der Messverfahren nicht überprüft oder Gütekriterien der Messverfahren sind unzureichend                                                                                                                                                                             | A8<br>1  |
| K30.<br>(C.9.)  | Zielkriterien beziehen sich auf patientenrelevante Parameter (insbesondere Schwere der Symptomatik, Leiden, Beeinträchtigung/Lebensqualität) | 1) Zielkriterien beziehen sich auf mehrere Dimensionen patientenrelevanter Parameter unter Einbezug von Beeinträchtigung/Lebensqualität<br>2) Zielkriterien beziehen sich nur auf eine Dimension<br>3) Zielkriterien beziehen sich ausschließlich auf Surrogatparameter (z. B. Kontrollüberzeugung)                                                                                                                       | C9<br>1  |
| K31.<br>(A.9.)  | Multiple Informationsquellen                                                                                                                 | 1) multidimensionale Erfassung der Zielkriterien – Drei oder mehr Informationsquellen<br>2) zwei Informationsquellen<br>3) eine Informationsquelle                                                                                                                                                                                                                                                                        | A9<br>2  |
| K32.<br>(A.10.) | Sofern Fremdeinschätzungsverfahren: externe Beurteiler (blind für die Gruppenzugehörigkeit)                                                  | 1) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), für die Gruppenbedingungen blinde externe Beurteiler<br>2) validiertes Fremdeinschätzungsverfahren angewendet von trainierten (erfahrenen), nicht-blinden externen Beurteilern<br>3) validiertes Fremdeinschätzungsverfahren angewendet – Rater sind weder trainiert (erfahren) noch blind für die Gruppenzugehörigkeit der Patienten | A10<br>1 |

**4.4 () Statistische Methodik (Subgruppenanalysen, Sensibilitätsanalysen)**

- deskriptive, korrelationsstatistische und interferenzstatistische Verfahren kommen zum Einsatz
- hypothesenprüfend genutzt wurden Chi-Quadrat-Test, t-Test für unabhängige Stichproben und ein- bzw. mehrfaktorielle Varianzanalysen

|                        | Kriterium                                                                              | Operationalisierung                                                                                                                                                                                                                                                                                                                                                                                                                | Rating   |
|------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K33.</b><br>(A.6.)  | Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert                | 1) a priori Definition der Vergleiche der Behandlungsgruppen und Messzeitpunkte erfüllt<br>2) teilweise post-hoc Definition der Vergleiche<br>3) Wirksamkeit ausschließlich bei post-hoc definierten Vergleichen                                                                                                                                                                                                                   | A6<br>2  |
| <b>K34.</b><br>(A.13.) | Anwendungsvoraussetzungen für statistische Modelle geprüft und erfüllt                 | 1) Anwendungsvoraussetzungen geprüft und erfüllt<br>2) Anwendungsvoraussetzungen geprüft und lediglich leichte Verletzungen der Voraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei gleichzeitiger Robustheit der angewendeten statistischen Verfahren<br>3) deutliche Verletzungen der Anwendungsvoraussetzungen oder keine Prüfung der Anwendungsvoraussetzungen bei substanziellem Risiko für deren Verletzung | A13<br>1 |
| <b>K35.</b><br>(A.14.) | Angemessenheit der statistischen Analysen (inklusive der Korrektur für multiple Tests) | 1) adäquate und umfassende statistische Analysen<br>2) weitgehend adäquate statistische Analysen<br>3) unangemessene statistische Analysen (fehlende Korrektur für multiple Tests, inadäquate statistische Methoden)                                                                                                                                                                                                               | A14<br>1 |

**4.5 (22) Ergebnisse**

- Abnahme der Impulskontrolle in allen drei Gruppen, Tendenz zu höherer Abnahme über die Zeit bei Borderline
- Abnahme psychotischer Symptome in allen drei Gruppen; sign. stärkere Abnahme in Borderline-Gruppe
- alle drei Gruppen weisen Verbesserungen in zwischenmenschlichen Beziehungen, PDS und Depressivitäts-Skala auf – keine Unterschiede zwischen den Gruppen
- DIB insgesamt: sign. stärkere Verbesserungen in Borderline-Gruppe in klinischen und diagnostischen Ausmaß – nur noch 2 von 13 Patienten haben DIB-Wert  $\geq 7$ , der diagnoserelevant ist
- im BSI keine Unterschiede zw. den Gruppen
- im GAS Borderline- und Depressions-Patienten sign. Verbesserung im Gegensatz zur Verschlechterung bei Schizophrenen
- laut VEV haben sich alle drei Gruppen stark verbessert, die Depressions-Patienten am meisten
- auch im FPI und Gießen-Test sind Verbesserungen in allen drei Gruppen da, bei Borderlinern höchste Annäherung von idealem und realem Selbstbild

|                        | Kriterium                                                                                              | Operationalisierung                                                                                                                                         | Rating   |
|------------------------|--------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| <b>K36.</b><br>(A.11.) | Vollständige Darstellung der Ergebnisse zu allen Outcome-maßen und zu allen relevanten Messzeitpunkten | 1) für alle Outcome-Variablen berichtet<br>2) ausschließlich für die primären Zielkriterien berichtet<br>3) nicht für alle primären Zielkriterien berichtet | A11<br>1 |

|  |  |  |   |
|--|--|--|---|
|  |  |  | 1 |
|--|--|--|---|

**4.6. (23) Ergebnisse zu unerwünschten Therapiewirkungen**

werden berichtet

|                        | Kriterium                                                                 | Operationalisierung                                                                                                                                                                                                                                                                                                                                     | Rating       |
|------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| <b>K38.</b><br>(A.12.) | Erfassung unerwünschter Ereignisse, Nebenwirkungen und Verschlechterungen | 1) Systematische Erfassung und Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>2) unsystematische Erfassung und/oder unvollständiger Bericht von unerwünschten Ereignissen, Nebenwirkungen, Verschlechterungen<br>3) Unerwünschte Ereignisse, Nebenwirkungen, Verschlechterungen wurden nicht erfasst oder nicht berichtet | A12<br><br>2 |

**4.7 (24) Fazit/Schlussfolgerungen der Autoren**

zeigt Wirksamkeit der Behandlung auf mittelfristige Sicht

|                       | Kriterium              | Operationalisierung                                                                    | Rating      |
|-----------------------|------------------------|----------------------------------------------------------------------------------------|-------------|
| <b>K39.</b><br>(A.1.) | Manipulation der Daten | 1) keine Hinweise auf Ergebnismanipulation<br><br>3) Hinweise auf Ergebnismanipulation | A1<br><br>1 |

| 5. Beurteilung der Ergebnisse der Studie |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| D1                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt signifikant größer als bei einer unbehandelten Kontrollgruppe, einer Placebo-Kontrollgruppe oder einer Kontrollgruppe mit "treatment as usual".<br>(Falls in der Studie primäre Zielkriterien nicht angegeben wurden, werden diese bei der Beurteilung festgelegt)<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                   | D1 |
| D2                                       | In der Interventionsgruppe ist die Verbesserung in den primären Zielkriterien zwischen Prä- und Post-Zeitpunkt mindestens gleich groß wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 (Statistische Power der Vergleiche) mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen)                                                                      | D2 |
| D3                                       | In der Interventionsgruppe wurde das Behandlungsziel einer Heilung oder Besserung der vorliegenden Störung signifikant besser bzw. bei signifikant mehr Patienten erreicht als bei den genannten Kontrollgruppen bzw. mindestens gleich häufig wie bei einer Vergleichsgruppe mit einer bereits als wirksam nachgewiesenen Behandlung, sofern das Kriterium A 16 mit eins oder zwei bewertet wurde.<br>(ja = 1; nein = 3; nicht zutreffend = freilassen) | D3 |

## 6. Praxistransfer

|                 | Kriterium                                                                  | Operationalisierung                                                                                                                                                                                                                                                                           | Rating   |
|-----------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| K40.<br>(C.10.) | Spezifikation und Herstellbarkeit notwendiger Settingbedingungen           | 1) Notwendige Settingbedingungen herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>2) Notwendige Settingbedingungen nur begrenzt herstellbar (z. B. Infrastruktur, Kooperation, Team)<br>3) Notwendige Settingbedingungen nicht herstellbar                                             | C10<br>2 |
| K41.<br>(C.11.) | Spezifikation und Herstellbarkeit der notwendigen Behandlungsqualifikation | 1) Notwendige Behandlungsqualifikation eindeutig beschrieben und herstellbar<br>2) Notwendige Behandlungsqualifikation eindeutig beschrieben, aber nur mit sehr großen Zeitaufwand herstellbar<br>3) Notwendige Behandlungsqualifikation nicht beschrieben oder praktisch nicht herstellbar   | C11<br>2 |
| K42.<br>(C.12.) | Spezifikation und Erfassbarkeit relevanter Patientenmerkmale               | 1) Relevante Patientenmerkmale (z. B. Alter, genetische Marker) praktisch erfassbar<br>2) Relevante Patientenmerkmale nur mit erheblichen Aufwand erfassbar<br>3) Relevante Patientenmerkmale praktisch nicht erfassbar                                                                       | C12<br>1 |
| K43.<br>(C.13.) | Spezifikation und Herstellbarkeit relevanter Treatmentaspekte              | 1) Relevante Treatmentmerkmale (Art der Interventionen, Reihenfolge, Dauer) praktisch herstellbar (z. B. durch Manual)<br>2) Relevante Treatmentmerkmale nur schwer herstellbar (z. B. tägliche Behandlung, Parallelbehandlung)<br>3) Relevante Treatmentmerkmale praktisch nicht herstellbar | C13<br>2 |

## Abschließende Bewertung

### Allgemeine methodische Qualität der Studie

- Wahl der Stichprobe ist ungewöhnlich – post-hoc Definition der Vergleiche
- Messinstrumente gut gewählt
- Kontroll-Gruppen sind ungewöhnlich – keine unbehandelten Patienten gleicher Störung, sondern unspezifisch behandelte Patienten anderer Störung

### Zusammenfassung der Ergebnisse

(Klinische Relevanz, Konsistenz der Befunde, Nachhaltigkeit der Therapiewirkungen, Nebenwirkungen/Risiken)

- zeigt Wirksamkeit der GT in allen drei Gruppen und vermehrt in der Borderline-Gruppe – klinisch relevante Verbesserungen (nur noch 2 von 13 Patienten überschreiten klinischen Wert im DIB)

### Interne Validität (Sicherheit kausaler Schlussfolgerungen)

- eingeschränkt, da in der Studie Patienten mit einer Borderlinestörung mit Patienten anderer klinischer Diagnosen verglichen wurden und insofern keine Daten einer adäquaten Vergleichsgruppe zu Verfügung stehen. Entsprechend weist nur die Borderline-Gruppe auch ausgeprägte Borderline-Symptomatik auf, während die Kontrollgruppen nur in sehr starkem Maße eine entsprechende Symptomatik berichten (keine Vergleichbarkeit der Gruppen vor Behandlungsbeginn)
- Einfluss weiterer zusätzlicher Interventionen wurde nicht kontrolliert (obwohl zugelassen)
- keine Dropout-Analysen
- Post-Werte werden mit Katamnese-Werten verglichen

### Externe Validität

Selektivität der Patienten, Repräsentativität der Behandler, Repräsentativität der Intervention, Übertragbarkeit auf das deutsche Versorgungssystem

- keine Beschreibung der Intervention
- hohe Selektivität der Patienten – Stichprobenziehung „im Nachhinein“
- Intervention ist ungewöhnlich und nicht ohne Weiteres auf ein anderes Setting übertragbar (Gruppenzusammensetzung, männlicher und weiblicher Therapeut)

### Stärken der Studie:

- Detaillierte Erfassung der Borderlinesymptomatik und für den Zeitraum der Studie angemessene Operationalisierung der Diagnosestellung

### Schwächen der Studie:

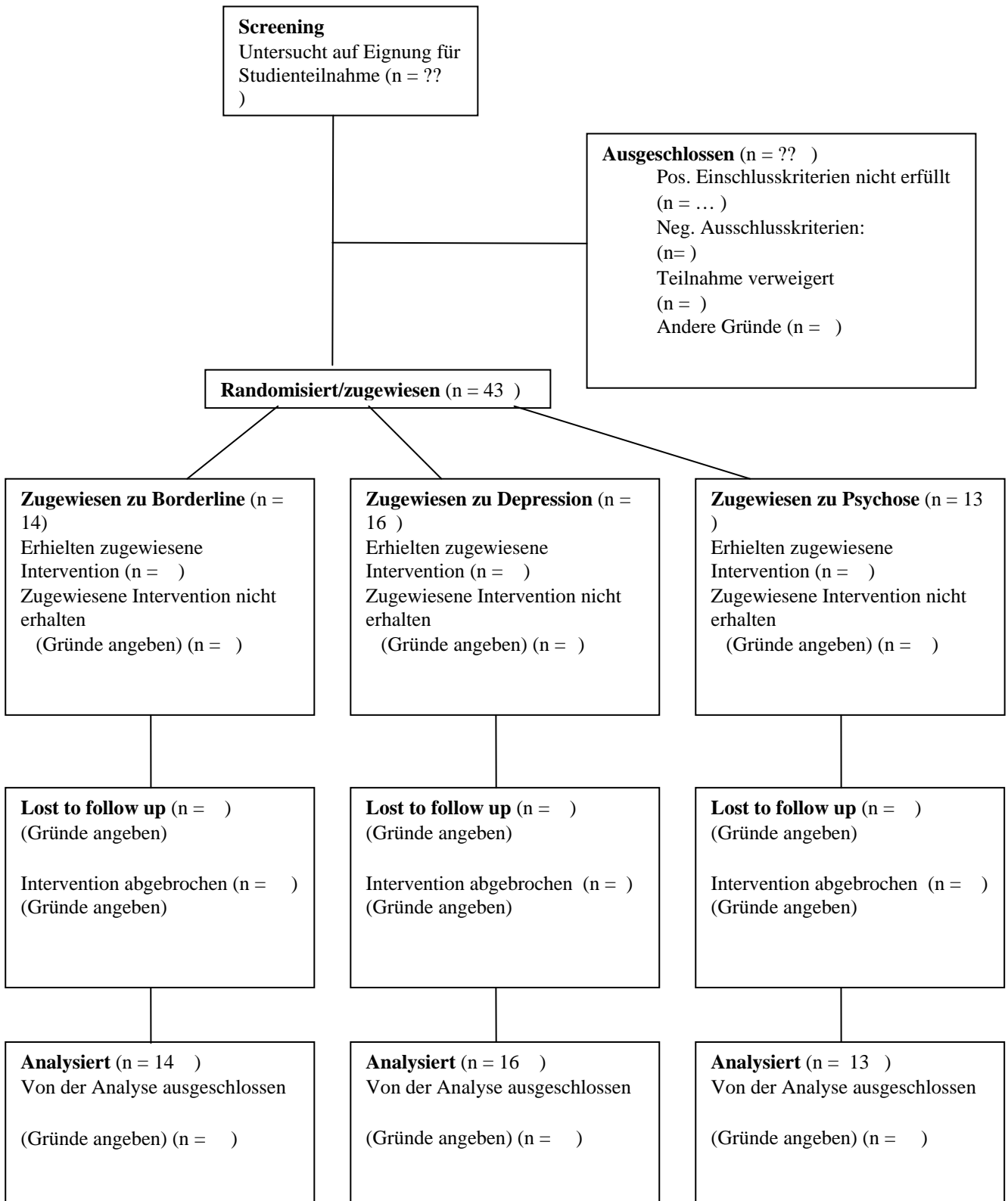
- Verwendung inadäquater Vergleichsgruppen

### Bedeutung der Studie für die Bewertung der Psychotherapiemethode

Eingeschränkt relevant

- aufgrund der inadäquaten Vergleichsgruppen sind sichere Aussagen über die Wirksamkeit der GT-Behandlung nicht möglich
- zeigt eine Wirksamkeit bei Borderline-Patienten, die auch klinisch relevant ist; jedoch ist aufgrund der methodischen Mängel (Stichprobenzusammensetzung und –größe, heterogene Messzeitpunkte post-hoc definiert) die Aussagekraft limitiert – somit liefert diese Studie einen Hinweis auf eine Wirksamkeit von GT bei Borderline-Persönlichkeitsstörungen

**8. Anhang zu Feld 6 aus Abschnitt 2.1: CONSORT-Diagramm**





## Erläuterungen zur Anwendung des Studienbewertungsbogens

Der Studienbewertungsbogen stellt eine Synthese aus dem Kriterienkatalog des Wissenschaftlichen Beirats Psychotherapie zur Bewertung der allgemeinen methodischen Qualität, der internen und externen Validität von Psychotherapiestudien und dem Bewertungsbogen des G-BA (angelehnt an CONSORT) dar.

Das Kernstück des Studienbewertungsbogens ist in vier größere Abschnitte gegliedert:

1. Basisinformation, 2. Population, 3. Intervention und 4. Ergebnisse. Für die Bewertung der externen Qualität kommt noch der Abschnitt 6. Praxistransfer hinzu.

Die einzelnen Abschnitte sind so aufgebaut, dass zunächst die relevanten Informationen zu Merkmalen der Studie extrahiert werden, bevor die Bewertung der Studienqualität anhand der zugehörigen Kriterien erfolgt. Dieses Vorgehen soll zum einen die Nachvollziehbarkeit von Bewertungen einzelner Kriterien erhöhen. Zum anderen soll eine Verbesserung der Qualität der Bewertung erreicht werden, dadurch dass die zur Beurteilung eines Kriteriums wichtigsten Informationen hinreichend umfassend extrahiert und bei der Bewertung systematisch berücksichtigt werden können. Um den Bewertungsprozess anhand der Kriterien ökonomisch zu halten, folgt die Bewertung eines Kriteriums, sobald die vorgegebene Extraktion eine ausreichende Bewertungsgrundlage für ein Kriterium liefert.

In einigen Fällen kann es jedoch für eine valide Bewertung eines Kriteriums sinnvoll sein, im Bogen weiter unten aufgeführte Informationen über Merkmale einer Studie mit zu berücksichtigen. Dies gilt insbesondere für die folgenden Kriterien:

**K5** (Selektivität der Stichprobe aufgrund der Ausschlusskriterien):

Hier bedarf es einer zusätzlichen Berücksichtigung der Informationen aus den Punkten 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener und ausgewerteter Patienten). Die tatsächliche Selektivität der Stichprobe aufgrund von Ausschlusskriterien lässt sich besser beurteilen, wenn nicht nur die spezifizierten Ausschlusskriterien betrachtet werden, sondern die Merkmale der resultierenden Stichprobe (z. B. Anteil komorbider Patienten, Schweregrad, etc.) und der Patientenfluss (wie viele Patienten werden tatsächlich ausgeschlossen und aus welchen Gründen) mitbetrachtet werden.

**K14** (Drop-out-Analysen):

Das Kriterium K14 prüft die Kontrolle und die Wahrscheinlichkeit möglicher Verzerrungen der Studienergebnisse durch drop-outs von Patienten im Verlauf der Studie. Bei der Bewertung dieses Kriteriums sind neben den Informationen aus 2.7 (Vergleichbarkeit der Gruppen) insbesondere auch die Informationen aus 2.4 (Zentrale Patientenmerkmale) und 2.6 (Anzahl gescreenter, eingeschlossener, zugewiesener und ausgewerteter Patienten – Patientenfluss nach CONSORT-Diagramm) mitzubetrachten.

**K33** (Vergleiche der Behandlungsgruppen und Messzeitpunkte a priori definiert):

Bei der Beurteilung dieses Kriteriums sind neben den Informationen aus 4.4 (Statistische Methodik – Subgruppenanalysen, Sensibilitätsanalysen) insbesondere die Informationen aus 3.5 (Beobachtungszeitraum/Messzeitpunkte/Katamnese) mitzubetrachten.

**K39** (Manipulation der Daten)

Für die Beurteilung dieses Kriteriums, ob Hinweise auf Manipulation der Ergebnisse der Studie vorliegen, sind sämtliche verfügbaren Informationen zu einer Studie heranzuziehen.

Zur besseren Nachvollziehbarkeit der jeweiligen Herkunft und ursprünglichen Position eines Extraktionspunkts bzw. eines Kriteriums finden sich in Klammern jeweils die Originalnummerierung des Extraktionspunktes in dem Bewertungsbogen des G-BA bzw. des Kriteriums in der ursprünglichen Version des Kriterienkatalogs des WBP. Um bei der Bewertung der einzelnen Kriterien eine schnellere Zuordnung der Kriterien zu den Bereichen „allgemeine methodische Qualität“ (A), „Interne Validität“ (B) und „Externe Validität“ (C) zu ermöglichen, ist zusätzlich in dem Ausfüllfeld des Kriteriums die Originalbezeichnung jeweils grau hinterlegt angegeben.

## **Anmerkungen zur durchgeführten Studienbewertungen**

(Hier kann insbesondere auf bei der Bewertung einzelner Kriterien entstandener Unklarheiten und Probleme hingewiesen werden. Bitte stets zunächst das betreffende Kriterium benennen. Außerdem kann hier in Ergänzung zum Abschnitt „Abschließende Bewertung auf Uneindeutigkeiten hinsichtlich der Zuordnung der untersuchten Intervention zur Gesprächspsychotherapie eingegangen werden)

zu K22c: es werden keine Angaben dazu gemacht, inwieweit sich der Therapeut für diese Therapie trainiert hat

zu K24: nur prä-katamnestisch