

Expertise für die Bundesärztekammer

Krankheitsspezifische Ausprägung von Placeboeffekten

Autoren:

PD Dr. med. Karin Meissner, Prof. Dr. med. Klaus Linde

München, 31.3.2013

Kontakt:

PD Dr. med. Karin Meissner

Gollierstr. 53

80338 München

Email: Karin.Meissner@med.lmu.de

Prof. Dr. med. Klaus Linde

Wendelsteinstr. 12

82054 Sauerlach

Email: Klaus.Linde@lrz.tum.de

Inhalt

1.	Arbeitsauftrag	5
2.	Grundsätzliche Überlegungen zur Beantwortung der Fragen	6
2.1.	Kontext und Krankheitsspezifität	6
2.2.	Ermittlung der Größe von Placeboeffekten	7
3.	Methodisches Vorgehen	13
4.	Zu Frage 1: Gibt es Hinweise, dass der Placeboeffekt in Behandlungsstudien verschiedener Krankheiten bzw. Diagnosegruppen differiert?	15
4.1.	Höchste Evidenzstufe: Meta-Analysen zu dreiarmligen Studien	19
4.2.	Mittlere Evidenzstufe: Meta-Analysen zu indirekten Vergleichen	22
4.3.	Untere Evidenzstufe: Meta-Analysen zur Placeboresponse und zum Anteil der Placebo- an der Gesamtresponse	27
4.4.	Zusammenfassende Bewertung zur Evidenzlage bezüglich der Frage 1	28
5.	Zu den Fragen 2 und 3: Gibt es Unterschiede im Placeboeffekt hinsichtlich der gewählten Messmethoden, insbesondere im Vergleich biologischer - versus nicht biologischer Verfahren? Wirkt Placebo verschieden auf PRO (Patient-Reported Outcomes, Patientenzufriedenheit) und objektive Messgrößen?	28
5.1.	Höchste Evidenzstufe: Meta-Analysen zu dreiarmligen Studien	28
5.2.	Mittlere Evidenzstufe: Meta-Analysen zu indirekten Vergleichen	29
5.3.	Untere Evidenzstufe: Meta-Analysen zur Placeboresponse und zum Anteil der Placebo- and der Gesamtresponse	32
5.4.	Zusammenfassende Bewertung zur Evidenzlage bezüglich der Fragen 2 und 3	36
6.	Zur Frage 4: Finden sich Hinweise auf die Bedeutung der initialen Krankheitsschwere bzw. auch der spezifischen Behandlungsvorgeschichte (z. B. bei der Behandlung sog. therapieresistenter Fälle)?	37
6.1.	Untere Evidenzstufe: Meta-Analysen zur Placeboresponse und zum Anteil der Placebo- and der Gesamtresponse	37
6.2.	Zusammenfassende Bewertung zur Evidenzlage bezüglich der Frage 4	41

7.	Zu Frage 5: Wie wird der der Placebo-Wirkung zuzuschreibende Effekt interpretiert? Gibt es hier Hinweise auf Unterschiede nach Diagnosen bzw. Krankheiten?	42
8.	Zu Frage 6: Welche Forschungsfragen sollten noch gelöst werden? Welche haben eine besondere Priorität?	43
9.	Zusammenfassung	46
10.	Literatur	49

Anhang

Wichtige Abkürzungen

BIOs	biologische Maße (in der Regel biochemische, physiologische oder maschinell gemessene Parameter)
PROs	Patient-reported Outcomes
OROs	Observer-reported Outcomes
RR	Relatives Risiko bzw. Risk Ratio
SMD	Standardisierte Mittelwertsdifferenz

1. **Arbeitsauftrag**

Die Bundesärztekammer hat die Autoren im Herbst 2012 beauftragt, eine Expertise zum Thema ‚Krankheitsspezifische Ausprägung von Placebo-Effekten‘ zu erstellen, bei der folgende Fragestellungen bearbeitet werden sollten:

- Frage 1: Gibt es Hinweise, dass der Placebo-Effekt in Behandlungsstudien verschiedener Krankheiten bzw. Diagnosegruppen differiert?
- Frage 2: Gibt es Unterschiede im Placeboeffekt hinsichtlich der gewählten Messmethoden, insbesondere im Vergleich biologischer versus nicht-biologischer Verfahren?
- Frage 3: Wirkt Placebo verschieden auf PRO (Patient-Reported Outcomes, Patientenzufriedenheit) und objektive Messgrößen?
- Frage 4: Finden sich Hinweise auf die Bedeutung der initialen Krankheitsschwere bzw. auch der spezifischen Behandlungsvorgeschichte (z. B. bei der Behandlung sog. therapieresistenter Fälle)?
- Frage 5: Wie wird der der Placebo-Wirkung zuzuschreibende Effekt interpretiert? Gibt es hier Hinweise auf Unterschiede nach Diagnosen bzw. Krankheiten?
- Frage 6: Welche Forschungsfragen sollten noch gelöst werden? Welche haben eine besondere Priorität?

Die Erteilung des Auftrags erfolgte nach Vorlage eines Antrags von Seiten der Antragsteller, in dem ausgeführt wurde, dass die Expertise in Anbetracht des zur Verfügung stehenden Erstellungszeitraumes (ca. 6 Monate) und Finanzierungsvolumens primär auf der Basis publizierter systematischer Übersichtsarbeiten zu klinischen Studien erstellt werden sollte.

2. Grundsätzliche Überlegungen zur Beantwortung der Fragen

2.1. Kontext und Krankheitsspezifität

Die experimentelle und klinische Forschung der letzten 15 Jahre hat gezeigt, dass Placeboeffekte reale psychobiologische Phänomene sind, die ihren Ursprung wesentlich im therapeutischen Kontext haben (Finniss et al., 2010). Es besteht Konsens, dass es nicht einen Placeboeffekt als solchen gibt, sondern eine Klasse von Placeboeffekten mit verschiedenen Auslösern und Wirkmechanismen (Benedetti, 2009). Die Gabe eines Placebos bzw. einer aktiven Therapie ist dabei nur ein Faktor, der zur Auslösung von Placeboeffekten führen kann. Ob und in welchem Umfang Placeboeffekte auftreten, hängt auch von Charakteristika des Patienten und des Behandlers, der Interaktion zwischen Patient und Behandler sowie dem medizinischen Umfeld ab (siehe Tabelle 1).

Tabelle 1: Kontextfaktoren in der Medizin (nach Di Blasi et al., 2001)

Charakteristika des Patienten	Charakteristika des Arztes	Arzt-Patient-Interaktion	Behandlungscharakteristika	Medizinisches Umfeld
z. B. Einstellungen und Erwartungen zu Krankheit und Therapie, Ängstlichkeit, Adhärenz	z. B. Status, Geschlecht, Einstellungen und Erwartungen zu Krankheit und Therapie	z. B. Verbale Suggestion, Beruhigung, Mitgefühl	z. B. Farbe, Größe, Form des Medikaments	z. B. Zu Hause, im Krankenhaus, Raumaufteilung

Diese so genannten „Kontextfaktoren“ sind in drei Bereichen der Medizin bedeutsam: Erstens in der klinischen Forschung, in der Placebointerventionen als Kontrolltherapien eingesetzt werden; hier erschweren große Placeboeffekte den Nachweis der spezifischen Wirksamkeit einer Therapie. Zweitens in der klinischen Praxis, wo Kontextfaktoren nachweislich die Wirkung von Standardtherapien modulieren (Bingel et al., 2011; Colloca et al., 2004), und Ärzte deshalb aufgerufen sind, Behandlungseffekte durch Optimierung der Kontextfaktoren zu maximieren (Bundesärztekammer, 2011). Drittens bilden Kontexteffekte auch eine wichtige Grundlage für den Einsatz von Placebointerventionen in der Praxis, der nicht nur in Deutschland (Meissner et al., 2012), sondern auch in vielen anderen Ländern in relevantem Umfang erfolgt (Fässler et al., 2010).

Die Bundesärztekammer hat jüngst ein Buch herausgegeben, in dem sie die bewusste Anwendung von reinen Placebos oder so genannten Pseudo-Placebos unter bestimmten

Voraussetzungen für vertretbar hält (Bundesärztekammer, 2011): (1) Es ist keine geprüfte wirksame (Pharmako-)Therapie vorhanden; (2) es handelt sich um relativ geringe Beschwerden und es liegt der ausdrückliche Wunsch des Patienten nach einer Behandlung vor; (3) es besteht Aussicht auf Erfolg einer Placebobehandlung bei dieser Erkrankung. Dieses Gutachten hat zum Ziel, das bisher verfügbare Wissen vor allem zu Punkt 3 dieser Empfehlungen zusammenzutragen: Was wissen wir heute zur Ausprägung von Placeboeffekten bei verschiedenen Erkrankungen, bzw. wo besteht noch Forschungsbedarf?

Aufgrund theoretischer Vorüberlegungen ist einerseits zu erwarten, dass bestimmte Erkrankungen besonders placebo-sensitiv sind, so z. B. Erkrankungen, die nachweislich durch psychologische Faktoren mit beeinflusst werden. Andererseits könnten innerhalb derselben Erkrankung aber auch bestimmte Parameter stärker, und andere weniger stark durch Kontextfaktoren beeinflusst werden. So zeigte eine neuere Studie, dass Asthmatiker, die subjektiv zufrieden stellend auf Placebogabe reagierten, objektiv keine Besserungen ihrer Lungenfunktion zeigten (Wechsler et al., 2011). Krankheit ist also sicher nur ein Faktor, der die Größe von Placeboeffekten mitbestimmt. Daher ist eine große Variabilität der Befunde zu erwarten.

Hinzu kommen eine Reihe von methodologischen Schwierigkeiten bei der Bearbeitung der Fragestellung, auf die im Folgenden kurz eingegangen wird, bevor die eigentlichen Fragen dieses Gutachtens bearbeitet werden.

2.2. Ermittlung der Größe von Placeboeffekten

2.2.1. Warum können Veränderungen in Placebogruppen klinischer Studien in der Regel nicht als Placeboeffekte interpretiert werden?

Um die vom wissenschaftlichen Beirat gestellten Fragen beantworten zu können, muss die Größe von Placeboeffekten quantifiziert werden. Die Erfassung der Größe von Placeboeffekten ist jedoch mit erheblichen Schwierigkeiten behaftet. Sehr häufig wurden und werden auch heute noch alle Änderungen, die unter Placebothherapie in randomisierten Studien auftreten, als Placeboeffekte bezeichnet und implizit ein kausaler Zusammenhang zwischen Placebogabe und beobachteter Änderung angenommen. Geht z.B. in einer Studie die Schmerzintensität um 40% zurück, wird dies als Placeboeffekt interpretiert. Dabei wird nicht berücksichtigt, dass Spontanbesserungen, statistische Artefakte (Regression zur Mitte) und Kointerventionen die beobachtete Veränderung verursacht haben können. In der neueren Placeboliteratur wird daher häufig zwischen *Placeboresponse* und *Placeboeffekt*

unterschieden (z. B. Enck et al., 2011b). Unter *Placeboresponse* werden dabei alle positiven Veränderungen im Zeitverlauf bei Patienten unter Placebothherapie verstanden, unter *Placeboeffekten* nur die nach Abzug von Faktoren wie Spontanveränderung, Regression zur Mitte etc. der Placebointervention kausal attribuierbare Veränderungen (siehe Abbildung 1).

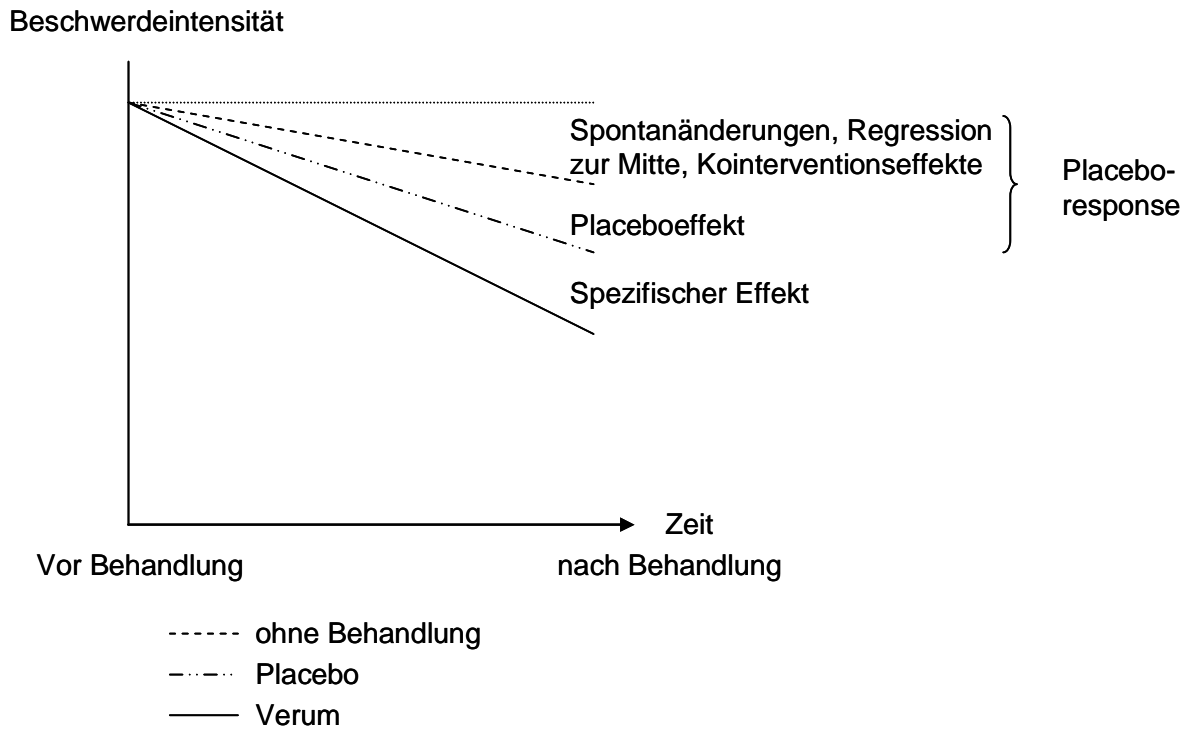


Abbildung 1

Schematische Darstellung zur Unterscheidung von spezifischen Effekten, Placeboeffekten und anderen Gründen für Veränderungen im Zeitverlauf

Spontanverbesserungen dürften v.a. bei akuten Erkrankungen ein wesentlicher Grund für Veränderungen über die Zeit in Placebogruppen sein. Stellt man sich z.B. vor, dass Patienten mit einem unkomplizierten Infekt der oberen Luftwege oder mit postoperativen Schmerzen, die ein Placebo erhalten, über einen Zeitraum von zwei Wochen beobachtet werden, so ist offensichtlich, dass ein relevanter oder gar kompletter Rückgang von Infektsymptomen bzw. Schmerzen auch ohne Placebogabe zu erwarten wäre. Bei chronischen Erkrankungen ist zwar grundsätzlich nicht regelhaft mit einer Spontanheilung zu rechnen, spontane Schwankungen des Krankheitsverlaufes sind jedoch auch hier zu erwarten. Auf der anderen Seite kann auch aus einer fehlenden Änderung über die Zeit nicht geschlossen werden, dass kein Placeboeffekt vorliegt. Beispielsweise könnte der Spontanverlauf progredient sein, durch einen Placeboeffekt jedoch eine Aufrechterhaltung des Zustandes vor Therapie erreicht werden.

Regression zur Mitte (engl. regression to the mean) dürfte ebenfalls in vielen Studien zu beobachteten Verbesserungen im Zeitverlauf beitragen. Z.B. werden in Studien zur Migräneprophylaxe (die üblicherweise nur bei Patienten mit durchschnittlich mindestens 2 Migräneattacken pro Monat als indiziert gilt) sehr häufig Patienten ausgeschlossen, die in einer vierwöchigen Baselinephase weniger als 2 Migräneattacken haben. Auch Patienten, die tatsächlich durchschnittlich mindestens 2 Attacken pro Monat haben, können zufällig in einzelnen Monaten nur eine Attacke erleiden oder gar attackenfrei bleiben. Werden diese Patienten systematisch ausgeschlossen, überschätzt der Mittelwert der Attackenhäufigkeit in den eingeschlossenen Patienten den eigentlichen Mittelwert in einer Population von Migränepatienten mit durchschnittlich mindestens 2 monatlichen Attacken. Wird in den Folgemonaten wiederum die Attackenhäufigkeit bestimmt, kommt es auch ohne jegliche Intervention zufällig wieder zu attackenfreien Monaten bzw. Monaten mit nur einer Attacke. Der entsprechende Mittelwert liegt daher unter dem in der Baselineperiode.

In vielen Studien ist es außerdem unmöglich, Patienten, die an einer randomisierten Studie teilnehmen und dabei eine Prüflintervention oder Placebo erhalten, andere effektive Therapien vorzuenthalten. Z.B. können Patienten mit chronischen Schmerzen bei Bedarf Analgetika einnehmen. Selbstverständlich können solche effektiven Kointerventionen Zielparameter in Studien beeinflussen.

2.2.2. Wie kann die Größe von Placeboeffekten quantifiziert werden?

Die beste Möglichkeit, um durch eine Placebointervention induzierte oder vermittelte Effekte hiervon abzugrenzen, ist das Mitführen einer unbehandelten Kontrollgruppe (no treatment Gruppe) bzw. einer Gruppe, die die gleichen Kointerventionen wie die Placebogruppe, jedoch keine Placebointervention, erhält (siehe Abb. 2) (Hrobjartsson et al., 2011). Als Placeboeffekt wird dementsprechend die Differenz der Veränderungen in beiden Gruppen bezeichnet (*nicht* die Veränderung in der Placebogruppe). Studien, die sowohl eine Placebo- wie auch eine unbehandelte Kontrollgruppe mitführen, stellen jedoch nur einen Bruchteil aller Studien dar. Auch derartige Studien können allerdings nicht verhindern, dass zumindest bei der Erfassung subjektiver Zielgrößen Verzerrungen auftreten können, z.B. durch soziale Erwünschtheit (z.B. ein Patient berichtet eine Verbesserung, obwohl er eine solche gar nicht empfunden hat).

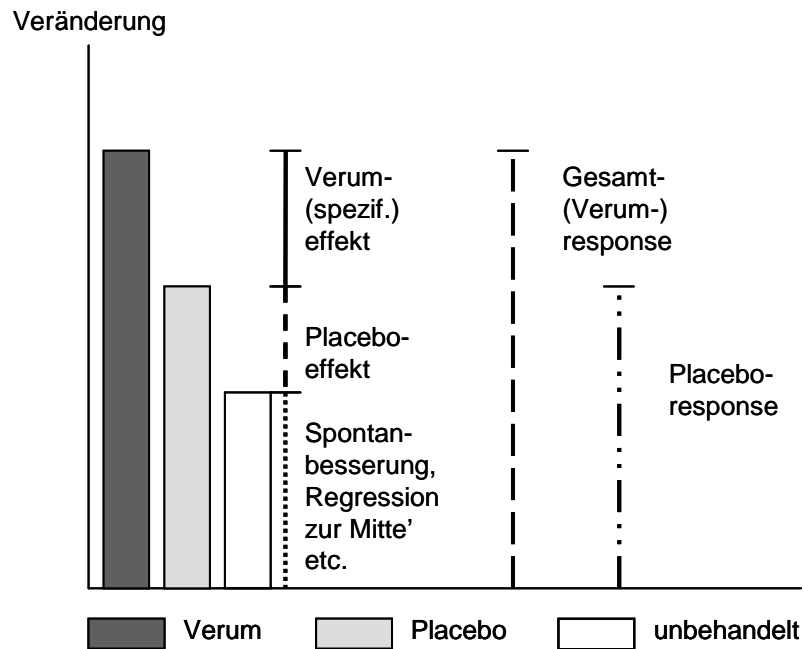


Abbildung 2

Typische dreiarmlige Studie zur direkten Untersuchung von (Verum- und) Placeboeffekten.

Da randomisierte Studien, in denen neben einer Placebogruppe auch eine unbehandelte Kontrollgruppe mitgeführt wird, selten sind, wird in systematischen Übersichten gelegentlich zu indirekten Vergleichen gegriffen. Dies bedeutet, dass die Größe der vorher-nachher Veränderungen in den Placebogruppen einer Gruppe von placebokontrollierten Studien und in den unbehandelten Kontrollgruppen einer Gruppe von *anderen* Studien mit einer solchen Kontrollgruppe berechnet werden und dann die Differenz dieser beider Größen ermittelt wird (Abbildung 3). Das Hauptproblem bei dieser Vorgehensweise ist, dass sich placebo-kontrollierte und no-treatment-kontrollierte Studien häufig auch in anderen Merkmalen unterscheiden, die sich auf die vorher-nachher-Unterschiede innerhalb der Gruppen auswirken können. No-treatment-Kontrollen werden häufig bei nichtmedikamentösen Studien eingesetzt, Placebokontrollen besonders häufig bei Arzneimittelstudien. Beispielsweise unterscheiden sich psychologische und Arzneimittelstudien in der Depressionsforschung in Selektionskriterien, Beobachtungsdauer, Messinstrumenten und Kointerventionen – alles Faktoren, die vorher-nachher Unterschiede beeinflussen können. Indirekte Vergleiche sind die zweitbeste Herangehensweise, wenn die Größe von Placeboeffekten krankheitsübergreifend verglichen werden sollen, sie müssen aber mit großer Zurückhaltung interpretiert werden. Derartige Analysen sind außerdem sehr selten. Grundsätzlich ist es möglich, indirekte und direkte Vergleiche in so genannten Netzwerk-Meta-Analysen zu kombinieren (Caldwell et al., 2005). Ein solches Vorgehen würde die vorhandene Evidenz am besten nutzen, ist aber in hohem Maße aufwändig und anspruchsvoll. Damit die

Ergebnisse von Netzwerk-Meta-Analysen valide sind, müssen außerdem eine Reihe von Voraussetzungen erfüllt sein. Die Autoren dieser Expertise haben dieses Vorgehen in begrenztem Umfang einmal für eine Placebofragestellung verwendet (Publikationsmanuskript eingereicht), weitere Anwendungen dieser Methodik zu Placebothemen sind ihnen jedoch nicht bekannt.

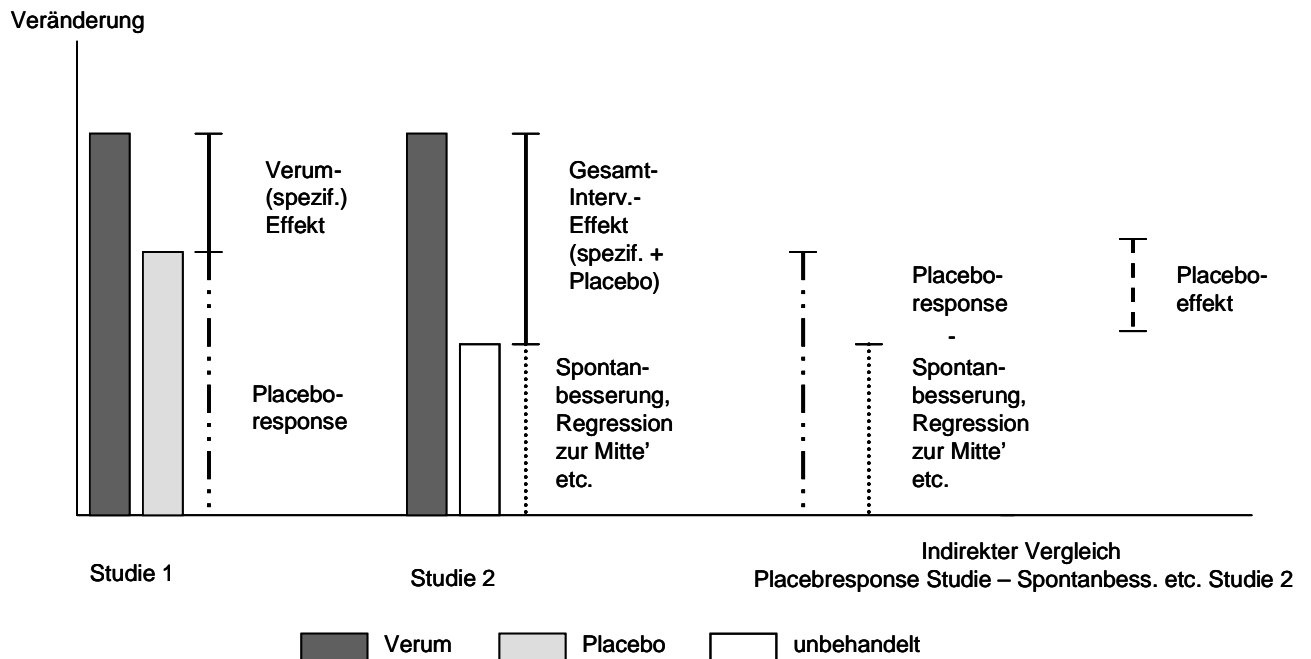


Abbildung 3

Indirekter Vergleich zur Untersuchung der Größe von Placeboeffekten. Studie 1 vergleicht Verum und Placebo, Studie 2 Verum und eine unbehandelte Kontrollgruppe; der Placeboeffekt wird ermittelt, indem die sonstigen Veränderungen (Spontanbesserung, Regression zu Mitte etc.) aus Studie 2 von der Placebosresponse in Studie 1 subtrahiert werden

Aus Sicht der Autoren ist auch der Anteil der Placeboresponse an der Gesamtresponse (Placeboresponse + spezifischer Effekt) von Interesse, obwohl sich hieraus keine direkte Aussagen zur Größe des Placeboeffektes ableiten lassen. Dieser Anteil gibt jedoch gerade für den Vergleich einzelner Erkrankungen einen intuitiven Eindruck der Relevanz unspezifischer Faktoren für die beobachteten Gesamtveränderungen im Zeitverlauf.

Will man die Placeboresponse, also die einfachen Veränderungsmessungen in Placebogruppen (siehe auch Abbildung 4), annähernd für die Untersuchung der Größe von Placeboeffekten bei unterschiedlichen Erkrankungen heranziehen, müssten bestimmte Voraussetzungen gegeben sein: will man z.B. die Wahrscheinlichkeit von Spontanverbesserungen gering halten, sollte die Symptomsschwere bei den Patienten über

einen längeren Zeitraum stabil sein. Um das Auftreten von Regression zur Mitte zu vermeiden, dürfte die Höhe eines Messwertes bei Baseline kein Einschlusskriterium sein. Kointerventionen dürften entweder nicht erfolgen oder müssten in ihrer Auswirkung genau einschätzbar sein. In der Realität dürften diese Voraussetzungen praktisch nie gegeben sein.

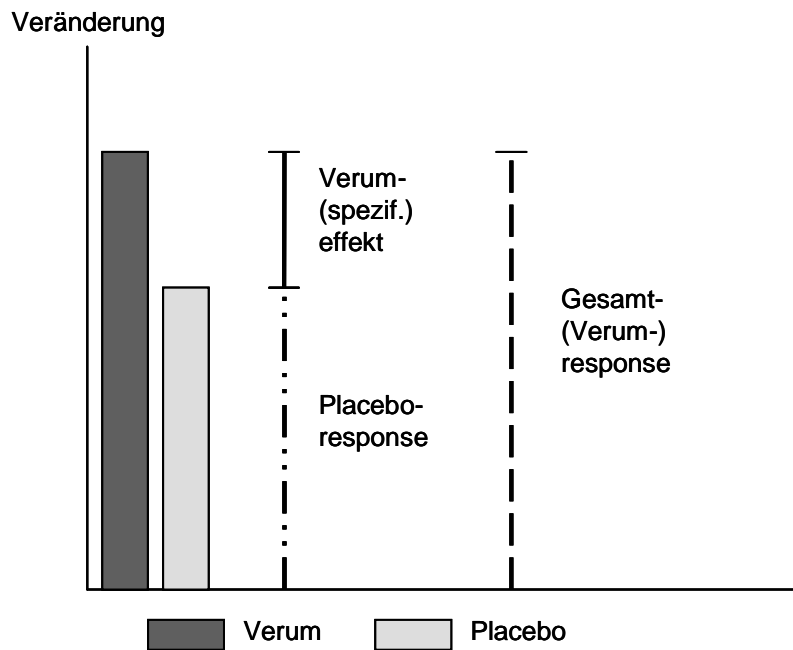


Abbildung 4

Übliche zweiarmige placebokontrollierte Studie. Eine Untersuchung des Placeboeffektes ist mit diesem Design nicht möglich

3. Methodisches Vorgehen

Für die Erstellung der Expertise wurde primär auf publizierte systematische Übersichtsarbeiten zum Thema Placeboeffekte zurückgegriffen. Die von den beiden Gutachtern über Jahre zusammengestellte Sammlung derartiger Arbeiten wurde durch eine Recherche in Medline aktualisiert und ergänzt. Hierzu wurden verschiedene Suchbegriffe für „Placebo“ und „Placeboeffekt“ mit Suchbegriffen für Meta-Analysen und systematische Reviews kombiniert.

So weit möglich wurden die Ergebnisse aus den zusammengetragenen Übersichtsarbeiten einer von drei Evidenzstufen zugeordnet:

- **Höchste Evidenzstufe:** Effektgrößen aus Meta-Analysen auf der Basis direkter Vergleiche, d.h. von Studien, die neben einer Placebogruppe auch eine unbehandelte Kontrollgruppe oder eine Gruppe, die die gleichen Kointerventionen wie die Placebogruppe, jedoch keine Placebointervention, erhielt, mitführten. In der Folge wird der Einfachheit halber immer von Vergleichen mit unbehandelten Kontrollgruppen gesprochen.
- **Mittlere Evidenzstufe:** Effektgrößen aus Meta-Analysen, bei denen eine Abschätzung des Placeboeffekts auf der Basis indirekter Vergleiche erfolgte (d.h. Ermittlung der Placebo-Response in einer Gruppe von Studien, Ermittlung der Veränderungen in unbehandelten Kontrollgruppen in separaten Studien).
- **Untere Evidenzstufe:** Effektgrößen aus Meta-Analysen zur Placeboresponse und zum Anteil der Placeboresponse an den Gesamtresponse unter Verum.

Die im Gutachten verwendeten Effektmaße sind im folgenden Kasten erläutert.

An dieser Stelle sei noch darauf hingewiesen, dass im Rahmen der vorliegenden Expertise angesichts des Arbeitsauftrages nicht systematisch auf negative Auswirkungen bzw. Nebenwirkungen von Placebointerventionen (Noceboeffekte) eingegangen wurde.

Maße zur Bestimmung der Größe des Placeboeffektes bzw. der Placeboresponse in Meta-Analysen

1. Für kontinuierliche Zielkriterien (z.B. Schmerz gemessen auf einer visuellen Analogskala)

1.1. Standardisierte Mittelwertsdifferenzen (SMD)

Sowohl in Meta-Analysen zu Placeboeffekten wie zur Placeboresponse werden diese Maße häufig berichtet, wenn die Primärstudien kontinuierliche Zielkriterien mit unterschiedlichen Methoden messen (z.B. die Intensität einer Depression verschiedenen Instrumenten oder auch völlig unterschiedliche Symptome wie Schmerz, Lebensqualität, Funktion etc.). Dabei wird die Differenz der Mittelwerte zwischen Placebogruppe und unbehandelter Kontrollgruppe bzw. in der Placebogruppe vor und nach Behandlung ermittelt und dann durch ein Maß für die Variabilität (häufig eine aus beiden Gruppen gepoolte Standardabweichung) geteilt. Vorteil der SMD ist, dass sie breit anwendbar ist, der wichtigste Nachteil, dass sie klinisch schwer interpretierbar ist. Als Daumenregel kann gelten: eine SMD von 0 bedeutet, dass keine Unterschiede vorhanden sind, eine SMD bis 0,40 ist ein kleiner Effekt, 0,40 bis 0,70 ein mittlerer, und $>0,70$ ein großer Effekt (Cochrane Collaboration, 2008).

1.2. Prozentuale Verbesserung

Eine Reihe von Meta-Analysen zur Placeboresponse berichtet die mittlere prozentuale Verbesserung. Dabei wird die Veränderung über die Zeit (change from baseline) in einer Gruppe durch den Baselinewert dividiert und mit 100 multipliziert, z.B. mittlerer Rückgang des HAMD Scores bei Depression = 10 geteilt durch Baselinewert = 20 ergibt 0,5 (entsprechend 50%).

2. Für dichotome Zielkriterien (z.B. klinisch relevante Besserung ja/nein)

2.1. Relatives Risiko (Risk ratio = RR) – nur für Studien mit unbehandelter Kontrollgruppe

Der Anteil der Ereignisse (z.B. Response = klinische relevante Besserung ja) in beiden Gruppen wird durcheinander geteilt, z.B. 20% Responder unter Placebo vs. 10% ohne Behandlung ergibt eine RR von 2.

2.2. Responderanteil

Die Zahl der Personen mit einem definierten Maß einer Verbesserung (= Responder; häufig mindestens 50% Symptomrückgang) wird durch die Zahl der Patienten in der jeweiligen Gruppe dividiert, z.B. 50 Responder bei 100 Patienten ergibt einen Responderanteil von 0,5 entsprechend 50%.

4. Zu Frage 1: Gibt es Hinweise, dass der Placeboeffekt in Behandlungsstudien verschiedener Krankheiten bzw. Diagnosegruppen differiert?

4.1. Höchste Evidenzstufe: Meta-Analysen zu dreiarmligen Studien

Die wichtigsten Ergebnisse zur Untersuchung krankheitsspezifischer Placeboeffekte stammen von Hróbjartsson und Gøtzsche. Diese legten zum ersten Mal 2001 im New England Journal of Medicine eine Meta-Analyse aller randomisierten klinischen Studien vor, in denen neben einer Placebogruppe auch eine unbehandelte Placebogruppe mitgeführt wurde (Hróbjartsson and Gøtzsche, 2001). Aktualisierungen dieser Arbeit mit Einbezug neuer Studien und zusätzlichen Analysen wurden 2004 (Hróbjartsson and Gøtzsche, 2004a, b) und 2010 publiziert (Hróbjartsson and Gøtzsche, 2010). Die im Folgenden präsentierten Daten stammen aus der aktuellsten und umfassendsten Publikation in der Cochrane Library aus dem Jahre 2010.

Abbildung 5 zeigt die Ergebnisse der krankheitsbezogenen Analyse von Hróbjartsson und Gøtzsche für kontinuierliche Parameter, Abbildung 6 für dichotome. Zunächst ist hervorzuheben, dass lediglich bei den Indikationen Schmerz, Übelkeit, Phobie und Asthma überhaupt ein signifikanter Unterschied zwischen Placebo und unbehandelter Kontrollgruppe für kontinuierliche Zielparameter gefunden wurde. Gemäß der in Kapitel 3 genannten Daumenregel ist nur der Placeboeffekt bei Phobien von mittlerer Größe, bei den anderen Erkrankungen ist er $< 0,40$ und damit klein. In allen anderen Indikationen sowie bei Analysen zu dichotomen Zielgrößen umfasst das 95%-Konfidenzintervall den Wert 0 (kein Gruppenunterschied bei kontinuierlichen Parametern) bzw. 1 (kein Unterschied bei relativen Risiken bzw. Risk Ratios). Dies ist vermutlich primär darauf zurückzuführen, dass die Zahl der Studien pro Indikation relativ gering ist, zum Teil aber auch darauf, dass auch innerhalb einer Indikation die Ergebnisse in einem erheblichen Maße streuten. Optisch scheint der Placeboeffekt bei Phobien und bei der Raucherentwöhnung besonders ausgeprägt zu sein (Abb. 5 und 6). Die sehr breiten Konfidenzintervalle zeigen jedoch, wie unsicher der sogenannte Punktschätzer (also der berechnete mittlere Effekt) ist – der wahre Wert könnte sowohl deutlich größer wie auch deutlich kleiner sein. Leider berichteten Hróbjartsson und Gøtzsche keinen Test auf Subgruppenunterschiede. Ein solcher Test würde Aufschluss darüber geben, ob die zwischen den Indikationen beobachteten Unterschiede noch mit dem Zufall vereinbar sind.

Neben der Arbeit von Hróbjartsson & Gøtzsche gibt es nur wenige weitere Meta-Analysen, die auf der Basis direkter Vergleiche krankheitsbezogenen Placeboeffekte untersucht haben.

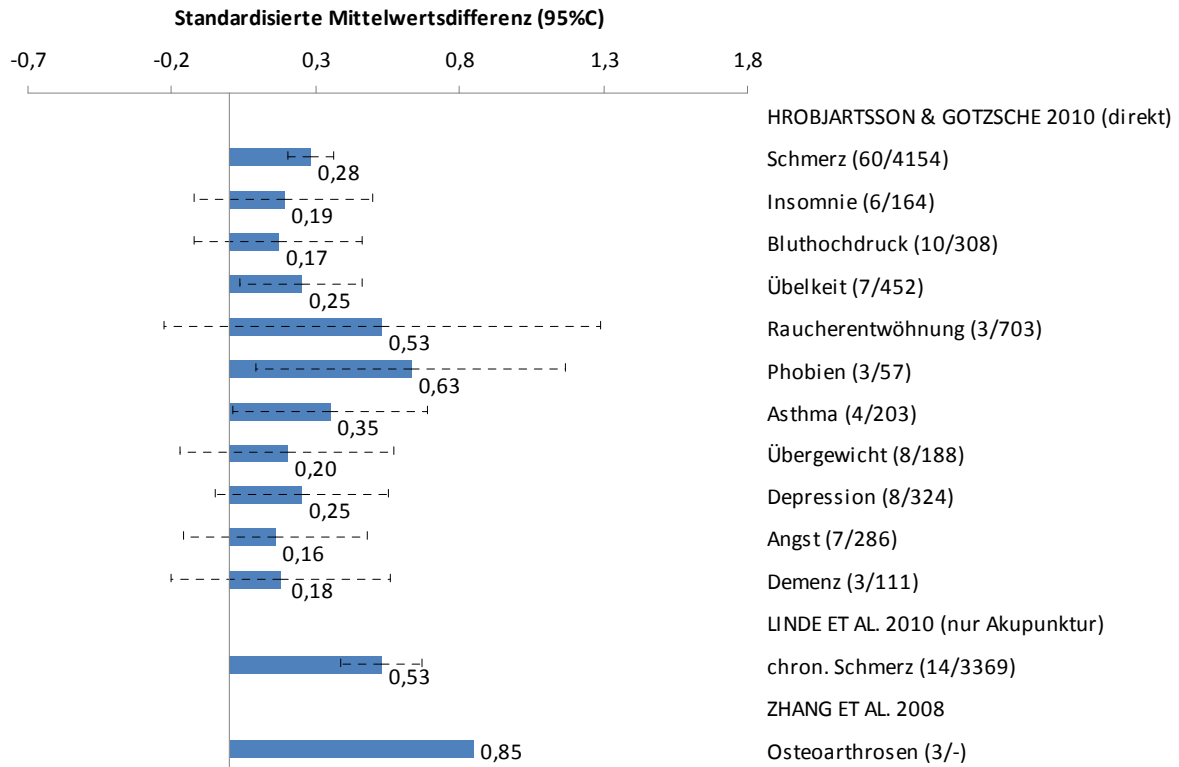


Abbildung 5

Größe der Placeboeffekte (kontinuierliche Maße) bei unterschiedlichen Erkrankungen – Ergebnisse der Meta-Analysen von Hróbjartsson & Gøtzsche 2010, Linde et al. 2010 und Zhang et al. 2008. Dargestellt sind die gepoolten standardisierte Mittelwertsdifferenzen mit den zugehörigen 95%-Konfidenzintervallen (95%CI) zu den einzelnen Indikationen (die erste Zahl nach der Indikation zeigt die Anzahl der vorhandenen Studien, die zweite Zahl die Anzahl der in diesen Studien untersuchten Patienten). Werte > 0 zeigen einen Überlegenheit von Placebo gegenüber Nichtbehandlung

Unsere eigene Arbeitsgruppe (Linde et al., 2010b) hat die vorliegenden dreiarmigen Studien zur Akupunktur zusammengestellt. Für 14 Studien zu chronischen Schmerzen konnte dabei eine standardisierte Mittelwertsdifferenz berechnet werden (siehe Abb. 5), die mit 0,53 und einem relativ engen 95%-Konfidenzintervall einen mittelgroßen Effekt anzeigten, der deutlich über dem von Hróbjartsson & Gøtzsche für Schmerz berechneten liegt. Das ist dadurch zu erklären, dass in unsere Analyse einige neuere Akupunkturstudien mit großen Effekten mit einbezogen wurden, die Hróbjartsson & Gøtzsche noch nicht eingeschlossen hatten. Im Rahmen eines umfangreichen Reviews zu Placeboresponse und Placeboeffekten in Studien zu Arthroseschmerzen fanden Zhang et al. (2008) auch drei Studien mit Placebo- und unbehandelter Kontrollgruppe. Während in den unbehandelten Gruppen eine leichte Verschlechterung (SMD 0,08) beobachtet wurde, ergab sich in den Placebogruppen im Mittel

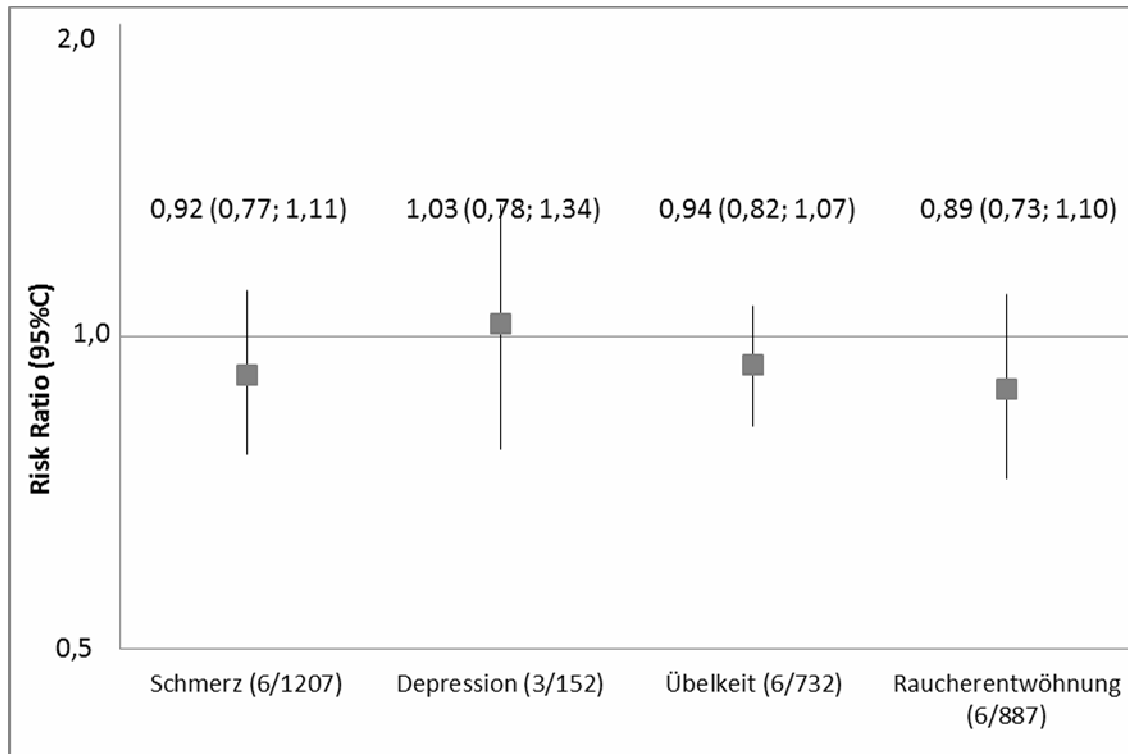


Abbildung 6

Größe der Placeboeffekte (dichotome Maße) bei unterschiedlichen Erkrankungen – Ergebnisse der Meta-Analyse von Hróbjartsson & Gøtzsche 2010. Dargestellt sind die Risk Ratios mit den zugehörigen Konfidenzintervallen zu den einzelnen Indikationen (die erste Zahl nach der Indikation zeigt die Anzahl der vorhandenen Studien, die zweite Zahl die Anzahl der in diesen Studien untersuchten Patienten). Werte < 1 zeigen eine Überlegenheit von Placebo vs. Nichtbehandlung, Werte > 1 eine Unterlegenheit

ein großer Effekt (0,77). Aufgrund der Darstellung von Zhang et al. kann für den Unterschied der SMDs zwischen Placebo- und unbehandelter Kontrollgruppe in Abbildung 5 kein Konfidenzintervall angezeigt werden. Nicht in Abbildung 5 aufgenommen wurden die Ergebnisse einer Meta-Analyse von Furukawa et al. (2007) zu drei Studien, in denen untersucht wurde, ob die Gabe eines Placebos zusätzlich zu einer kognitiven Verhaltenstherapie bei Patienten mit Panikstörungen Zusatzeffekte im Vergleich zu einer kognitiven Verhaltenstherapie allein auslöst. Tatsächlich wurde am Ende der Akutbehandlungsphase ein Zusatzeffekt von 26% (95%CI 2 bis 55%) gefunden, in der Nachbeobachtungsphase war der Effekt jedoch nicht mehr statistisch signifikant (22%, 95%CI -10 bis 64%).

Die Arbeitsgruppe von Hróbjartsson & Gøtzsche hat unter der Federführung von L. Krogsbøll (2009) eine zusätzliche Analyse ihrer Daten vorgelegt, in die ausschließlich Studien mit Verum-, Placebo- und unbehandeltem Arm mit einem kontinuierlichem Zielkriterium

eingeschlossen wurden, wenn für eine Indikation mindestens 3 Studien vorhanden waren. Diese Analyse erlaubt eine Abschätzung der Anteile von Verumeffekten, Placeboeffekten und sonstigen Veränderungen dargestellt als prozentualer Anteil an der Gesamtresponse (siehe Abbildung 7). Über alle Indikationen hinweg war der Anteil der Verumeffekte an der Gesamtresponse 56%, der von Placeboeffekten 20% und der der sonstigen Veränderungen 24%. Die ausgeprägtesten sonstigen Veränderungen waren bei Übelkeit (45%), Raucherentwöhnung (40%), Depression (35%), Phobien (34% und akuten Schmerzen (25%) zu beobachten, die ausgeprägtesten Placeboeffekte bei Phobien (40%), Schlafstörungen (29%), Übelkeit (28%), Übergewicht (28%) und chronischem Schmerz (27%). Diese Ergebnisse sind jedoch mit großer Zurückhaltung zu interpretieren. Die Zahl der Studien pro Indikation ist sehr gering und die (für die Prozentanalyse nicht verfügbaren, für die SMD-Analysen in der Publikation jedoch berichteten) Konfidenzintervalle sind sehr breit.

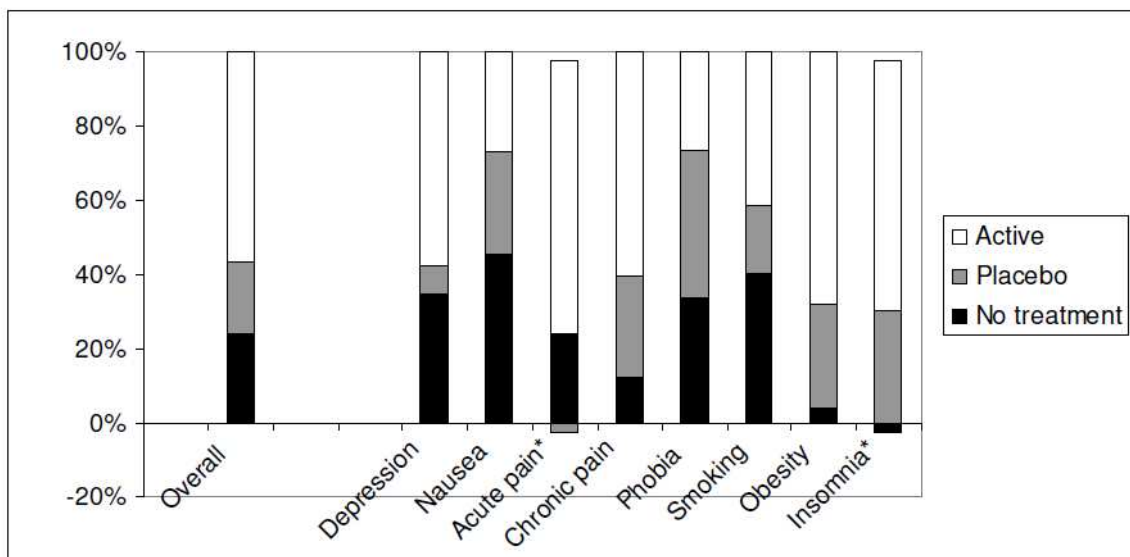


Abbildung 7 (aus Krogsbøll et al. 2009)

Relativer Anteil von Spontanbesserung, Placebo und spezifischem Effekt zur Veränderung gegenüber Baseline (*bei akutem Schmerz und Insomnie wurden in der unbehandelten Gruppen Verschlechterungen beobachtet)

Es gibt einige weitere Meta-Analysen auf der Basis direkter Vergleiche. Zwei enthalten jedoch keine krankheitsbezogenen Analysen (Meissner et al., 2007; Wampold et al., 2005). Zwei weitere (Vase et al., 2009; Vase et al., 2002) verglichen die Größe von Placeboeffekten in klinischen und rein experimentellen Schmerzstudien. Der Placeboeffekt war in den rein experimentellen Studien deutlich größer als in den klinischen Studien.

4.2. Mittlere Evidenzstufe: Meta-Analysen zu indirekten Vergleichen

Auch die Zahl systematischer Reviews, die die Größe von Placeboeffekten krankheitsbezogen anhand indirekter Vergleiche mit unbehandelten Kontrollgruppen untersucht, ist gering. Wir gehen davon aus, dass es eine Reihe weiterer systematischer Übersichtsarbeiten gibt, in denen sowohl placebo-kontrollierte Studien als auch solche mit unbehandelter Kontrolle eingeschlossen und als „unbehandelte Kontrollen“ gemeinsam analysiert wurden, die aber das Thema Placeboeffekte nicht formal untersuchen, obwohl sie dazu prinzipiell geeignet wären. Den Autoren lagen insgesamt 7 Meta-Analysen bzw. systematische Reviews zu indirekten Vergleichen vor: 2 zur Depression (Khan et al., 2012; Kirsch and Sapirstein, 1998) und jeweils eine zu Schmerzen bei Osteoarthrosen (Zhang et al., 2008, die auch 3 direkte Vergleiche einbezogen hatte, siehe oben), in der Onkologie (Chvetzoff and Tannock, 2003), bei Schlafstörungen (Belanger et al., 2007), bei Kopfschmerzen (de Groot et al., 2011) und bei multipler Sklerose (La Mantia et al., 1996).

Nur aus den Arbeiten von Kirsch & Sapirstein (1998) und von Zhang et al. (2008) lassen sich SMDs ermitteln, die eine unmittelbare Gegenüberstellung der Daten mit den Ergebnissen der direkten Analysen erlauben (siehe Abbildung 8). Dabei wird deutlich, dass Zhang et al. (2008) in der umfangreicheren Analyse mit Einschluss der indirekten Vergleiche deutlich kleinere Placeboeffekte fanden als in den direkten Vergleichen allein. Die SMD von 0,48 ist jedoch immer noch ein bedeutsamer Effekt, zumal in einer Zusatzpublikation zu der Meta-Analyse von einem Teil des Autorenteam (Doherty and Dieppe, 2009) darauf hingewiesen wird (über eine Referenz auf eine NICE-Leitlinie), dass bei Osteoarthrosen Verumeffekte in der Regel zwischen 0,2 und 0,3 liegen. Das würde bedeuten, dass die Größe der Placeboeffekte die der Verumeffekte für diese Indikation deutlich übersteigt. Auch die Ergebnisse von Kirsch und Sapirstein suggerieren bei der Depression einen deutlich größeren Placeboeffekt (SMD 0,89) als die (auf der Basis anderer Studien ermittelten) Daten von Hróbjartsson und Gøtzsche. Hier lassen sich auch Prozentanteile an der Gesamtresponse berechnen, die ebenfalls den Ergebnissen von Hróbjartsson und Gøtzsche widersprechen (siehe Abbildung 8). Diese Widersprüche sind vermutlich dadurch begründet, dass sich Studien mit Placebo- und unbehandelter Kontrollgruppe von Studien, die jeweils nur eine dieser Kontrollarten mitführen, systematisch unterscheiden. Die Meta-Analyse von Khan berichtet keine SMDs sondern ausschließlich prozentuale Rückgänge der Depressionssymptomatik. In dieser Übersicht wurden publizierte Studien und unpublizierte Studien, die der US-amerikanischen Food and Drug Administration vorgelegt, separat analysiert. Der prozentuale Symptomrückgang betrug in den beiden Subgruppen unter Placebo 38% bzw. 32%, ohne Behandlung 11% bzw. 12% und unter Verum 51% bzw. 42%. Der Anteil der Placebowirkung am Gesamteffekt ist ähnlich wie in der Arbeit von Kirsch und Sapirstein (Abbildung 9).

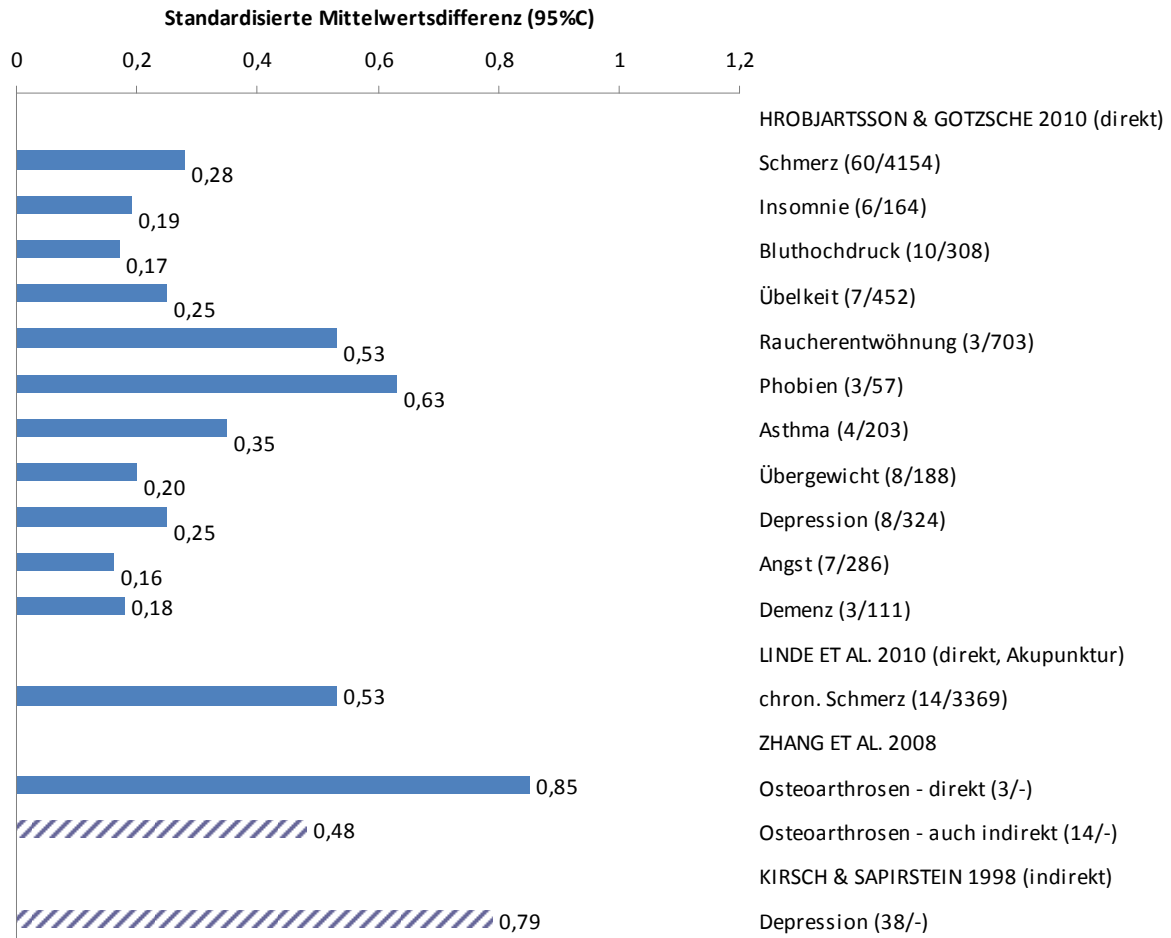


Abbildung 8

Größe der Placeboeffekte (kontinuierliche Maße) bei unterschiedlichen Erkrankungen – Ergebnisse auf der Basis direkter (ausgefüllte Balken) und indirekter Vergleiche (schraffierte Balken) mit unbehandelten Kontrollgruppen. Dargestellt sind die gepoolten standardisierten Mittelwertsdifferenzen zu den einzelnen Indikationen (die erste Zahl nach der Indikation zeigt die Anzahl der vorhandenen Studien, die zweite Zahl die Anzahl der in diesen Studien untersuchten Patienten).

Die übrigen vier Übersichtsarbeiten lassen sich schwer in die Darstellung einarbeiten, da sie sehr unterschiedlich vorgehen und zum Teil stark deskriptiven Charakter haben (Details zu Ergebnissen siehe Tabelle A1 im Anhang). Chvetzoff & Tannock (2003) verglichen die Ergebnisse von 37 Studien bei Krebspatienten mit einer Placebokontrolle mit denen von 10 Studien mit einer Best Supportive Care Kontrolle unter der Annahme, dass in den Placebokontrollgruppen ebenfalls alle Patienten Best Supportive Care erhalten hatten. Die Studien wurden nach Ergebnissen zu den Zielkriterien Schmerz, Appetit, Gewicht, körperliche Leistungsfähigkeit, Lebensqualität und Tumorresponse durchgesehen. Nur für einen Teil der Studien wurden dabei konkrete Zahlen extrahiert und nicht für alle Parameter

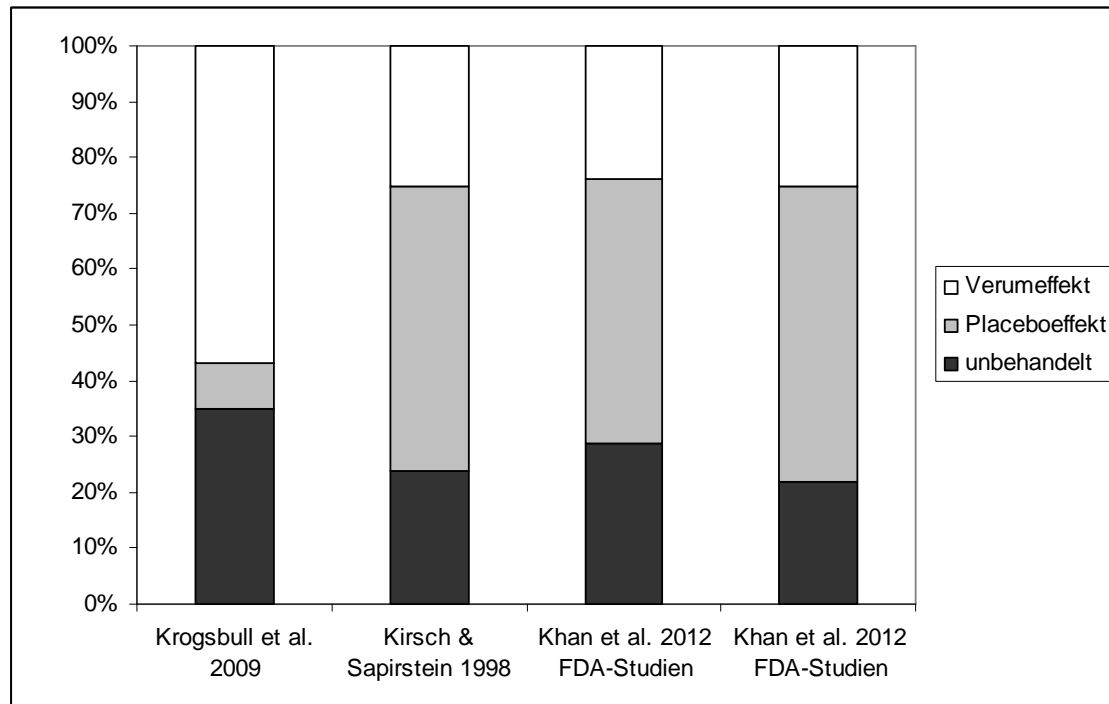


Abbildung 9

Relativer Anteil von Spontanbesserung, Placebo und spezifischem Effekt zur Veränderung gegenüber Baseline in Studien zu Depression in Analysen basierend auf direkten Vergleichen (Krøgsbull et al.) und indirekten Vergleichen

konnte ein Vergleich mit Best Supportive Care allein erfolgen. Die Autoren schließen, dass bei der Symptombehandlung zu einem begrenzten Grad Placeboeffekte auftreten, bei der Tumorresponse eher nicht.

In die Übersicht von La Mantia et al. wurden insgesamt 15 placebokontrollierte Studien bei Patienten mit multipler Sklerose eingeschlossen. Wurde der Verlauf mit dem Jahr verglichen, das der Studie vorausging, ergab sich unter Placebothherapie eine verlangsamte Progression und Verschlechterung der Funktion. Für 6 Studien konnte die Verschlechterung nach ein, zwei oder drei Jahren mit den Daten unbehandelter Patienten aus 3 unabhängigen Studien verglichen werden. Die Ergebnisse waren sehr variabel, legen insgesamt aber nahe, dass der Verlauf unter Placebothherapie etwas günstiger wer.

Bélanger et al. (2007) verglichen die Ergebnisse von 23 Gruppen mit arzneilichen Placebos, 4 Gruppen mit psychologischen Placebos und 7 Wartelistenkontrollen aus 34 unabhängigen Studien zu Schlafstörungen in Bezug auf einzelne Schlafparameter. Signifikante Unterschiede zwischen pharmakologischen Placebos und Warteliste ergaben sich für zwei von sechs subjektiven Schlafparametern und keinen von vier objektiven, zwischen psychologischen Placebos und Wartelisten für einen subjektiven Parameter (siehe auch 5.2).

Aufgrund der geringen Fallzahlen bei psychologischen Placebos und Warteliste sind die Ergebnisse jedoch zurückhaltend zu interpretieren.

De Groot et al. (2011) stellten die Anteile von (unterschiedlich berechneten) Respondern aus Kontrollgruppen sehr unterschiedlicher Kopfschmerzstudien (Spannungskopfschmerzen, Migräne akut und prophylaktisch) einander gegenüber. Unter pharmakologischem Placebo lag der Responderanteil bei 38,5% (95% CI, 37,8-39,3; n=78), bei pseudobehavioralen Placebos bei 8,4% (95% CI, 6,3-10,5%; n=13), bei Aufmerksamkeitskontrolle bei 18,3% (95% CI, 15,7-20,9%; n=11), bei Wartelistekontrolle bei 17,9% (95% CI, 15,3-20,6%; n=20) und in sonstigen Gruppen ohne Behandlung bei 6,4% (95% CI, 3,8-9,0%; n=10).

4.3. Untere Evidenzstufe: Meta-Analysen zur Placeboresponse und zum Anteil der Placebo- and der Gesamtresponse

Es liegt eine große Zahl von Meta-Analysen und systematischen Reviews zur Größe von Veränderungen über die Zeit unter Placebothherapie (Placeboresponse) vor. In einem Teil dieser Arbeiten wurden auch die entsprechenden Veränderungen in der Verumgruppe berechnet, sodass der Anteil der Placebo- an der Gesamtresponse berechnet werden kann. An einigen Beispielen soll demonstriert werden, wie problematisch es ist, auf der Basis solcher Vorher-Nachher Veränderungen Aussagen zu krankheitsspezifischen Placeboeffekten zu machen.

Abbildung 10 zeigt die Placeboresponse aus den vorliegenden Meta-Analysen zu Kopfschmerzstudien. Je nach Art des Zielkriteriums, nach Art des Placebos oder Alter (Erwachsene oder Kinder) unterscheidet sich die Placeboresponse zum Teil deutlich. Besonders drastisch ist der Unterschied bei der Akutbehandlung von Migräneattacken: Die Größe der Placeboresponse unterscheidet sich zwischen dem Kriterium Response (deutliche Schmerzlinderung) und dem Kriterium Schmerzfreiheit nach 2 Stunden bis um den Faktor 8.

Ein weiteres Problem ist die häufig große Variabilität der Placeboresponse in den einzelnen Studien auch in Analysen, in denen in allen Studien die gleiche Erkrankung mit dem gleichen Zielkriterium untersucht wurde. Abbildung 11 stammt aus einer Meta-Analyse von Fernandes et al. (2008). Bei Studien zur Behandlung akuter Migräneattacken wird typischerweise der Anteil der Patienten erfasst, der innerhalb von 2 Stunden eine deutliche Schmerzlinderung hat (linke Hälfte), sowie der Anteil der Patienten, der in diesem Zeitraum schmerzfrei geworden ist (rechte Hälfte). Die Variabilität in dieser Meta-Analyse kann als durchschnittlich bezeichnet werden. In einer großen Übersicht von Moerman (2000) zu Studien zur

Behandlung von Magenerkrankungen mit Cimetidin und Ranitidin variierte die endoskopisch verifizierte Heilungsrate in den Placebogruppen zwischen 0 und 100%.

Ein Beispiel für die Probleme, die bei der Bewertung der Placeboresponse bei progredienten Erkrankungen auftreten, gibt die Übersicht von Shetty et al. (1999) zur Symptomatik bei Morbus Parkinson. In dieser Arbeit wurden 36 placebokontrollierte Studien mit Beobachtungszeiträumen von 1 bis 18 Monaten zusammengefasst. Die mediane prozentuale Veränderung der Symptomatik lag in den Placebogruppen bei 3%, in den Verumgruppen bei 16%. Man könnte daraus berechnen, dass der Anteil der Placeboresponse an der Gesamtresponse unter Verum 19% $((3\%/16\%) \times 100)$ war. Bei M. Parkinson ist jedoch davon auszugehen, dass sich die Symptomatik im Zeitverlauf ohne

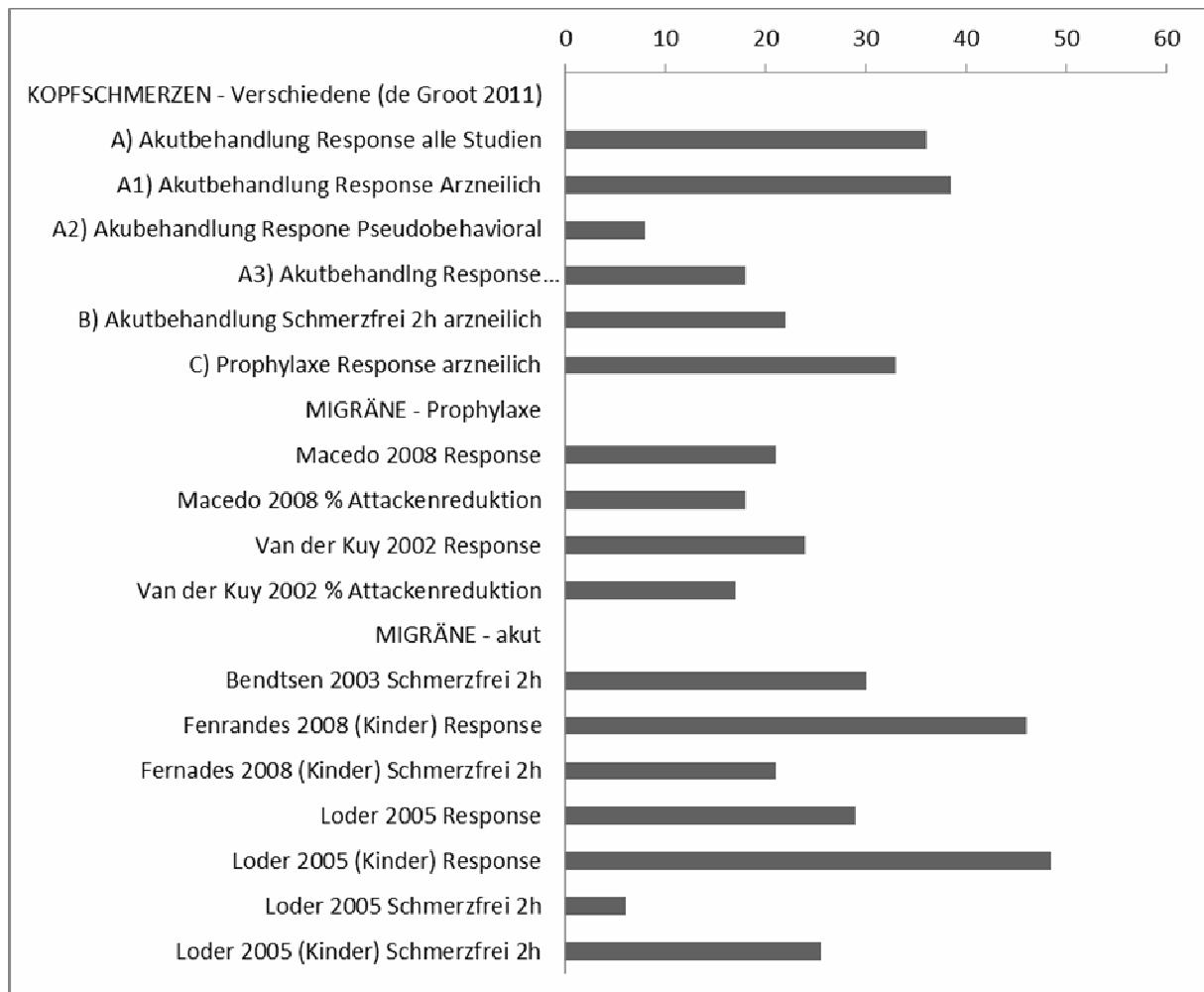


Abbildung 10

Variabilität der Punktschätzer der Placeboresponse in Meta-Analysen je nach Zielparameter, Kopfschmerzdiagnose und Alter. Dargestellt Placeboresponse in % in Studien zur Behandlung von Kopfschmerzen

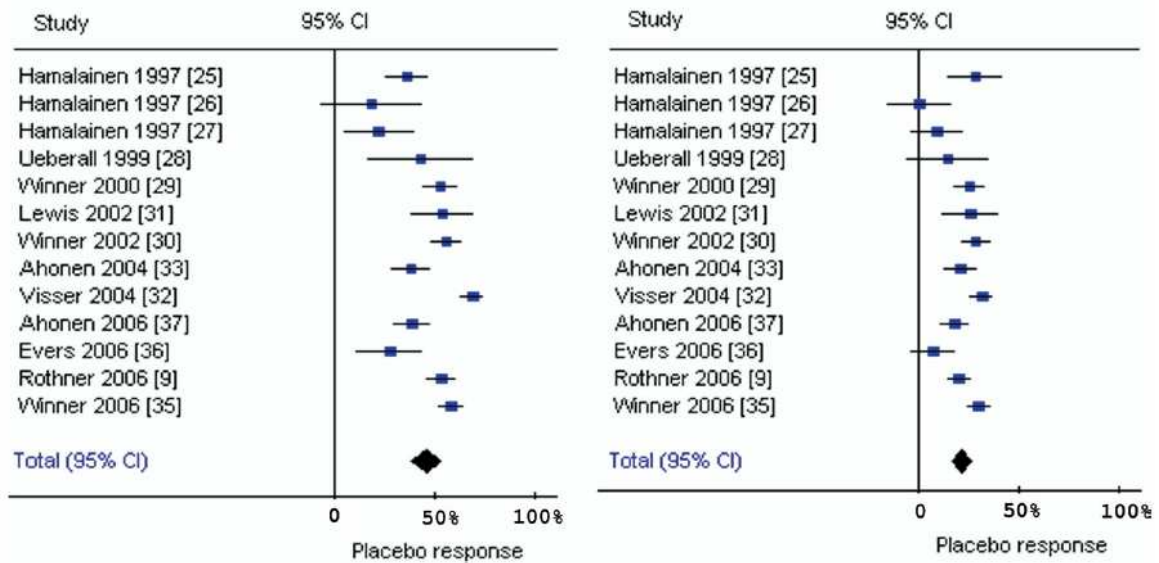


Abbildung 11 (aus Fernandes et al. (2008))

Variabilität der Placeboresponse zwischen einzelnen Studien - Anteil der Patienten, der nach Placebogabe innerhalb von 2 Stunden eine deutliche Schmerzlinderung erfuhr (linke Hälfte) bzw. schmerzfrei wurde (rechte Hälfte). Dargestellt sind Prozentanteile (05%-Konfidenzintervalle)

Behandlung verschlechtert. In 12 Placebo- und 7 Verumgruppen wurde auch tatsächlich eine Verschlechterung beobachtet. Man könnte dies zwar als Noceboresponse interpretieren, es ist jedoch offensichtlich, dass zumindest für diese Übersichtsarbeit eine Bewertung von Placeboeffekten ohne unbehandelte Kontrollgruppe und auch eine Abschätzung des Anteils der Placebo- an der Gesamtresponse nicht sinnvoll möglich sind.

Angesichts dieser Probleme sind die in der folgenden Tabelle 2 zusammengestellten Angaben zur krankheitsbezogenen Placeboresponse und dem entsprechenden Anteil der Placeboresponse an der Gesamtresponse mit größter Zurückhaltung zu interpretieren. Die Tabelle zeigt dass Analysen zur Placeboresponse insbesondere für neurologische, psychiatrische und gastroenterologische Erkrankungen durchgeführt wurden. Bemerkenswert erscheint den Autoren, dass abgesehen von Kopfschmerzen kaum Analysen zu anderen gängigen Schmerzerkrankungen (insbesondere Rückenschmerzen) vorliegen. Besonders häufig wurde die Placeboresponse bei medikamentöser Therapie der Depression untersucht (für eine ausführliche Zusammenstellung der einzelnen Übersichtsarbeiten s. Tabelle A2 im Anhang).

Tabelle 2

Placeboresponse und der Anteil an Gesamtresponse (Placeboresponse/Verumresponse) in verfügbaren Meta-Analysen. Angegeben sind nach Möglichkeit die Ergebnisse für Responsekriterien (möglichst $\geq 50\%$ Symptomverbesserung) oder, wenn keine Response angegeben wurde, der prozentuale Symptomrückgang (Kürzel V). Mehrere Zahlen pro Spalte beziehen sich auf die Punktschätzer der vorliegenden Meta-Analysen

Indikation	Placeboresponse	Anteil an Gesamtresponse (=Verumresponse)
NEUROLOGISCH		
Kopfschmerzen	36%	
Migräneprophylaxe	12%, 24%	50%, 52%
Migräneattackenbehandlung	29%, 30%, Kinder: 46%, 48%	
Neuropathische Schmerzen	17%; 24%%	
Diabetische Neuropathie		V 63%
Postherpetische Neuralgien		42%
Restless Leg Syndrom	40%	59%
Fibromyalgie	19%	70%, V 45%
Epilepsie	9%, 13%	
M. Parkinson		V 56%/62%
PSYCHIATRISCH		
Depression	27% (ITT Analyse), 28%, 28%, 28%, 30%, 31%, 35%, 35%, 37%, 37%, 38%, 46% (PP Analyse)	51%, 59% (ITT-Analyse), 60%, 62% (PP-Analyse), 67%, 68%, 69%, 69%, 70%, 72%, 73%
Manie	31%	64%
Bipolare Manie	33%	69%
Zwangserkrankungen	10%	38% Medik., 24% psych. Ther.
Zwangserkrankungen (Kinder)	31%	
Angst generalisiert	37%	
Angst (Kinder)	40%	
Panikstörung	48%, 64%	
Soziale Phobie	39%, 24%	
Schizophrenie	21%	55%, V21%
Nikotinsucht	8%	

Tabelle 2 (Fortsetzung)

Indikation	Placeboresponse	Anteil an Gesamtresponse (=Verumresponse)
Alkoholsucht	18%	
Kokainabusus	47%	
Opioidabusus	44%	
ADHS	21%	
GASTROENTEROLOGISCH		
Gastroösophagealer Reflux	12%, 19%, endoskopisch 27%	38%, 43%, endoskopisch 57%
Funktionelle Dyspepsie	24%, 35%	59%
Funktionelle Magen- Darmstörungen	43%	67%
Magenulkus	endoskopisch 34% (4 Wochen), 52% (8 Wochen)	54% (4 Wochen), 60% (8 Wochen)
Chronische Pankreatitis	20%	
Funktionelle Darmbeschwerden	38%, 40%	
Reizdarmsyndrom	36%, 40% 36%	74%
M. Crohn (intervall)	56%	
M. Crohn (akut)	19%	
Colitis ulcerosa	27%, 28%, 32%, 35%	
VERSCHIEDENE		
Sexuelle Dysfunktion (weibl.)	35%	
Benigne Prostatahyperplasie		V 95%
Saisonale allerg. Rhinitis		V 66%
Chronische Müdigkeitssyndrom	20%	
Müdigkeit bei Krebserkr.	56%	
Schmerz bei Krebserkr.	9%	27%
Husten		85%
Pulmon. arterielle Hypertonie	12%	37%
Hypertonie		V 25% (systolisch), 32% (diastolisch)
Asthma	15%	38%
Psoriasis	4%	8%

4.4. Zusammenfassende Bewertung zur Evidenzlage bezüglich der Frage 1

Auf der Basis theoretischer Überlegungen unter Berücksichtigung des aktuellen Kenntnisstandes zu Placeboeffekten erscheint es nahe liegend anzunehmen, dass Placeboeffekte bei bestimmten Erkrankungen eine größere Rolle spielen als bei anderen (siehe 2.1.). Die vorliegende Evidenz aus in Meta-Analysen und systematischen Übersichtsarbeiten zusammengefassten klinischen Studien ist zwar grundsätzlich mit diesen Überlegungen gut vereinbar, gleichzeitig macht sie aber deutlich, dass 1) unser Wissenstand sehr begrenzt ist, da Studien und Meta-Analysen mit direkten und indirekten Vergleichen von Placebogruppen und unbehandelten Kontrollgruppen selten sind, und 2) ganz offensichtlich so viele Faktoren Placeboeffekte und Placeboresponse beeinflussen, dass die Art der Erkrankung nur ein Einflussfaktor unter vielen ist.

Die Menge der Evidenz aus direkten und indirekten Vergleichen ist nicht nur begrenzt, die Ergebnisse sind zum Teil auch widersprüchlich. So scheinen z.B. die vorliegenden Ergebnisse zur Größe von Placeboeffekten und dem Anteil von Placeboeffekten im Vergleich zum spezifischen Effekt kaum miteinander vereinbar. Diese Widersprüche könnten dadurch erklärbar sein, dass dreiarmige Studien und Studien, die in indirekte Vergleiche eingehen deutlich unterschiedlich sind. In die indirekten Vergleiche gehen viele „typische“ placebo-kontrollierte Studien ein, unbehandelte Kontrollgruppen finden sich häufiger in Studien mit psychologischen Prüfinerventionen. Dreiarmige Studien werden meist nur bei bestimmten Fragestellungen durchgeführt und könnten daher mit anderen Kontextbedingungen einhergehen.

Unter kritischer Berücksichtigung von Umfang und Qualität der Evidenz könnte man sagen, dass in *klinischen* Studien Placeboeffekte mit ausreichender Sicherheit nur bei Schmerz wirklich eindeutig nachgewiesen sind. Bei vielen anderen Indikationen scheint das Vorliegen von Placeboeffekten zwar plausibel und wahrscheinlich, es bestehen aber erhebliche Unsicherheiten. Eine zuverlässige Quantifizierung der Größe in Bezug auf einzelne Erkrankungen erscheint daher zum jetzigen Zeitpunkt unmöglich. Zu beachten ist auch, dass unklar ist, in welchem Ausmaß Placeboeffekte und Placeboresponse in klinischen Studien die Realität in der ärztlichen Praxis abbilden.

5. Zu den Fragen 2 und 3: Gibt es Unterschiede im Placeboeffekt hinsichtlich der gewählten Messmethoden, insbesondere im Vergleich biologischer versus nicht-biologischer Verfahren? Wirkt Placebo verschieden auf PRO (Patient-Reported Outcomes, Patientenzufriedenheit) und objektive Messgrößen?

5.1 Höchste Evidenzstufe: Metaanalysen zu dreiarmligen Studien

In der Meta-Analyse von Hróbjartsson & Gøtzsche (2010) war der gepoolte Placeboeffekt für dichotome „Patient-reported Outcomes“ (PRO; RR 0,93, 95% CI 0,86 bis 1,00; n=31) und für dichotome „Observer-reported Outcomes“ (RR 0,93, 95% CI 0,85 bis 1,02; n=13) fast identisch. Für kontinuierliche PROs war der gepoolte Placeboeffekt (SMD -0,26, 95% CI 0,32 bis -0,19; n=109) signifikant größer ($p=0,045$) als für kontinuierliche „Observer-reported Outcomes“ (SMD -0,13, 95% CI -0,24 bis -0,02; n=49). Dieser Unterschied beruhte hauptsächlich auf dem leicht negativen Placeboeffekt für Labormaße (SMD 0,16, 95% CI 0,01 bis 0,30) sowie einem kleinen Effekt auf „Observer-reported Outcomes“, die keine Kooperation des Patienten erforderten (SMD -0,12, 95% CI -0,29 bis 0,05). Der Effekt auf „Observer-reported Outcomes“, die eine Kooperation des Patienten erfordern, war ähnlich groß (SMD -0,26, 95% CI -0,41 bis -0,12) wie der Effekt auf PROs.

Im Rahmen dieses Gutachtens verglichen wir zudem die gepoolten Placeboeffekte für subjektive (PROs) und objektive Maße (OROs/BIOs)¹ aus der Meta-Analyse von Hróbjartsson & Gøtzsche (2010) auch innerhalb einzelner Indikationen, und zwar dann, wenn mindestens zwei Studien mit subjektiven und objektiven Zielparametern für eine Indikation vorlagen. Dies war für die kontinuierlichen Parameter der Indikationen Schmerz und Depression der Fall.

Von 60 Schmerzstudien erfassten vier Studien den Schmerz mit objektiven Maßen. Sieht man von einer Studie an Neugeborenen ab, für die man a priori keinen Placeboeffekt erwarten würde (Carbajal, 1999; SMD 0,10, 95% CI -0,45 bis 0,60), so zeigten drei weitere Studien mit „Schmerzmittelverbrauch“ als objektivem Maß einen gepoolten Placeboeffekt von -0,39 SMD (95% CI -0,74 bis -0,03). Dieser Effekt unterschied sich nicht signifikant vom Gesamtplaceboeffekt in dieser Subgruppe (SMD -0,28, 95% CI -0,36 bis -0,19). In der Gruppe der Depressionsstudien mit kontinuierlichen Zielparametern befanden sich je vier

¹ Im Folgenden werden als ‚OROs‘ ‚Observer-reported Outcomes‘ bezeichnet, die auf einer subjektiven Einschätzung durch einen Beobachter (in der Regel eines Arztes) beruhen, und als ‚BIOs‘ solche, die biologische Maße darstellen (in der Regel biochemische, physiologische oder maschinell gemessene Parameter).

Studien mit subjektiven bzw. objektiven Maßen. Auch hier waren die gepoolten Placeboeffekte für beide Maße vergleichbar groß (subjektive Maße: SMD -0,22, 95% CI -0,58 bis 0,14; objektive Maße: SMD -0,21, 95% CI -0,77 bis 0,35).

Wampold et al. (2005) griffen aus dem Datensatz von Hróbjartsson & Gotzsche (2001) 19 Studien heraus, die sowohl Patientenmaße (PROs) als auch biologische Maße (BIOs) berichteten. Der Unterschied des gepoolten Placeboeffekts zwischen PROs und BIOs war mit einer Effektstärke von 0,11 gering und nicht signifikant. Um welche Indikationen und welche PROs und BIOs es sich bei den 19 Studien genau handelte, wurde nicht berichtet.

Dass auch innerhalb der Gruppe der BIOs die Größe der Placeboeffekte möglicherweise stark variiert, legen die Ergebnisse einer indikationsübergreifenden Meta-Analyse nahe (Meissner et al., 2007). In dieser Meta-Analyse wurde nach systematischen Unterschieden in der Größe von Placeboeffekten auf verschiedene BIOs gesucht. Die Ergebnisse zeigten, dass Placeboeffekte auf BIOs vor allem bei funktionellen, physikalisch gemessenen Zielgrößen auftraten (z. B. Blutdruck, Lungenfunktion; SMD 0,22, 95% CI 0,07 bis 0,36; n=20), während biochemische Laborparameter (z. B. Cholesterin, Glucose) eine leichte Verschlechterung unter Placebo im Vergleich zu Nichtbehandlung zeigten (SMD -0,17, 95% CI -0,31 bis -0,02; n=6).

Nicht zuletzt sollen hier auch die Ergebnisse einer neueren, schon eingangs erwähnten Placebostudie zu Asthma erwähnt werden, die einen direkten Vergleich von Placeboeffekten auf PROs und BIOs ermöglichten (Wechsler et al., 2011). Sechsendvierzig Asthmapatienten erhielten an vier verschiedenen Tagen je vier Interventionen gegen Asthma, nämlich pharmakologische Inhalationen, Placeboinhalationen, Placeboakupunktur, oder keine Behandlung (in randomisierter Reihenfolge). Während die Patienten nach beiden Placebobehandlungen eine signifikante Besserung (von 45-50%) ihrer Asthmasymptome (PRO) gegenüber Nichtbehandlung (21%) berichteten, ergab die Auswertung der objektiv erfassten Lungenfunktion (FEV-1; BIO) nach keiner der Placebobehandlungen (je 7%) eine signifikante Besserung im Vergleich zu Nichtbehandlung (7%). Diese Ergebnisse zeigen, dass subjektive Placeboeffekte bei Asthma auch ganz unabhängig von objektiven Verbesserungen der Lungenfunktion auftreten können. Die Ergebnisse beruhen jedoch auf einer Einzelstudie an 46 Patienten und lassen sich zum gegenwärtigen Zeitpunkt nicht verallgemeinern.

5.2 Mittlere Evidenzstufe: Meta-Analysen zu indirekten Vergleichen

Chvetzoff et al. (2003) analysierten Placeboeffekte in der Onkologie in 37 Studien mit Placebokontrollen und 10 Studien mit ‚Best supportive care‘ als Kontrollen. Ein Placeboeffekt auf Schmerzen (Zusatzeffekt über ‚Best supportive care‘) trat bei etwa 9% der Patienten auf,

ein Placeboeffekt auf den Appetit bei 2%, und ein Placeboeffekt auf den Tumor bei 1% der Patienten. Der relative Anteil des Placeboeffekts am Gesamteffekt lag für die PROs ‚Schmerz‘ und ‚Appetit‘ bei 21% und 27% und erreichte auch für den BIO ‚Tumorresponse‘ 19% (s. Abb. 13). Auch die Placeboresponse für den BIO ‚Gewicht‘ und den ORO ‚Allgemeinzustand‘ (absolute Verbesserungen, 10% und 11%) wich mit jeweils 40% Anteil am Gesamteffekt nicht wesentlich von der Placeboresponse der übrigen Parameter ab. Somit lassen sich bei Krebserkrankungen bisher keine konsistenten Unterschiede hinsichtlich des Placeboeffekts auf PROs, OROs und BIOs erkennen. Jedoch muss einschränkend gesagt werden, dass die Besserungen in der Placebogruppe insgesamt gering waren, was auch den Vergleich zwischen den einzelnen Zielparametergruppen erschwert.

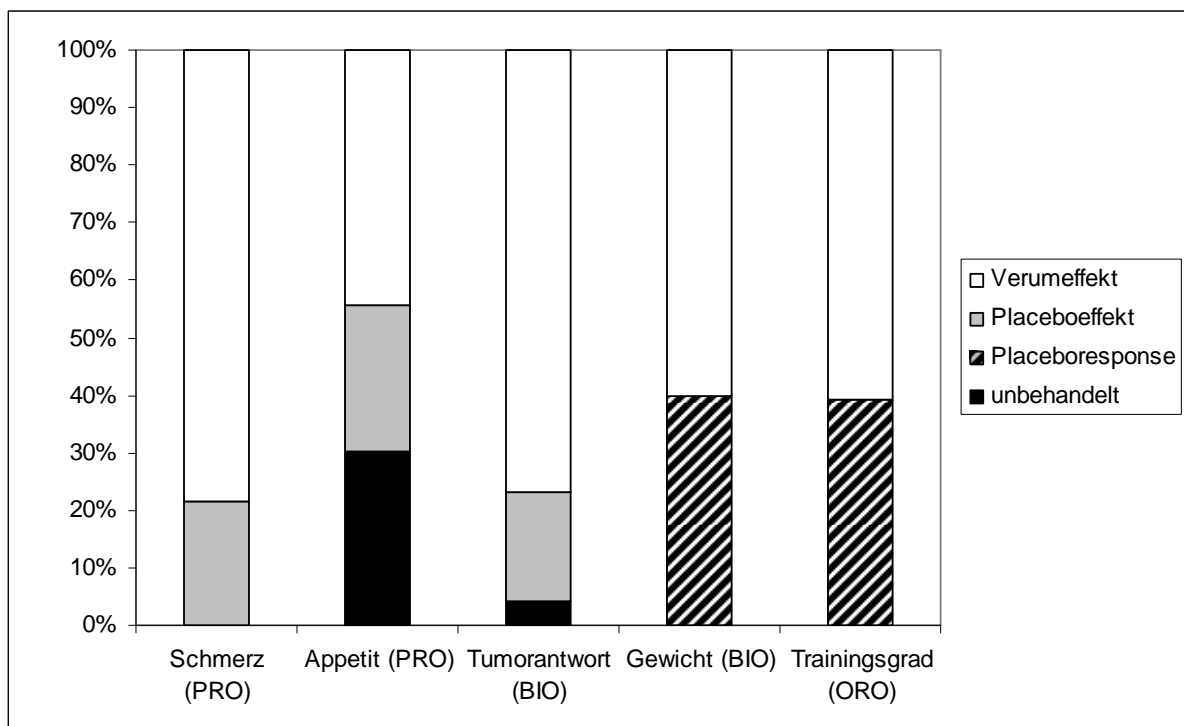


Abbildung 13

Relativer Anteil von Spontanbesserung (‚Best supportive care‘), Placeboeffekt bzw. Placeboresponse und spezifischem Effekt an Vorher-Nachher-Änderungen in 37 onkologischen Studien, basierend auf indirekten Vergleichen (Chvetzoff et al., 2003). PRO = patient-reported Outcome, ORO = observer-reported Outcome, BIO = biologischer Outcome.

Belanger et al. (2007) untersuchten die Placeboresponse bei Schlafstörungen und berichteten hierbei sowohl Effektstärken von PROs als auch von BIOs für dieselben Kriterien. So wurde z. B. die Einschlafzeit subjektiv vom Patienten eingeschätzt sowie

objektiv mit Hilfe des EEG's erfasst. Die Autoren unterschieden zudem zwischen zwei Arten von Placebokontrollen, nämlich pharmakologischen und psychologischen Placebointerventionen, und berichteten auch die Ergebnisse von mehreren Wartelistekontrollen. In Abb. 14 sind die Ergebnisse für pharmakologische Placebos abzüglich der Effekte der Wartelistekontrollen zu sehen. Durchgehend lagen die Effektstärken für PROs über denen der BIOs für dieselben Parameter; signifikant war der Unterschied eigenen Berechnungen zufolge allerdings nur für die Schlaflatenz. Bei den psychologischen Placebointerventionen (Abb. 15) zeigten alle PROs Verbesserungen, während sich die BIOs derselben Parameter verschlechterten; die Unterschiede wurden (mutmaßlich aufgrund der sehr kleinen Fallzahlen) nicht signifikant. Die Ergebnisse dieser Meta-Analyse legen nahe, dass BIOs bei Schlafstörungen kleinere Placeboeffekte zeigen als PROs. Zudem machen die Ergebnisse aber auch deutlich, dass die Größe des Placeboeffekts auch von der Art der Placebointervention mitbestimmt wird: In der Behandlung von Schlafstörungen scheinen pharmakologische Placebos wirksamer zu sein als psychologische Placebointerventionen.

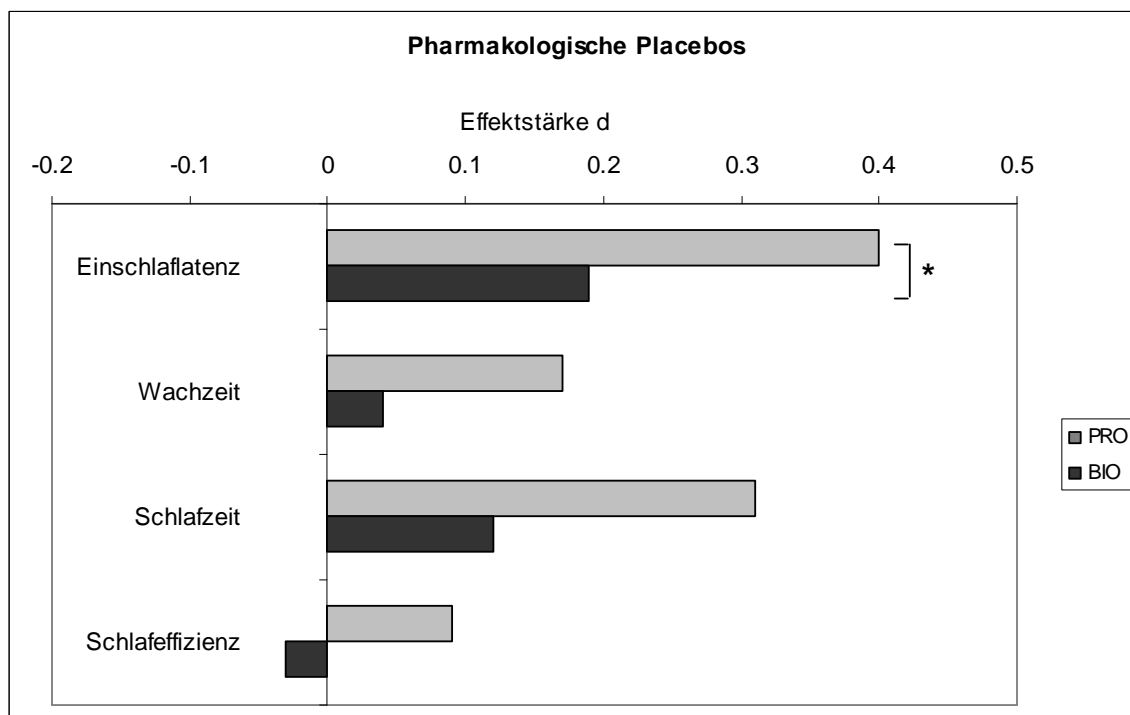


Abbildung 14

PRO vs. BIO (Effektstärke/pharm. Placebos minus Effektstärke/Warteliste) in 34 Studien zu Schlafstörungen (Belanger et al., 2007). * Unterschied statistisch signifikant ($p < 0,05$)

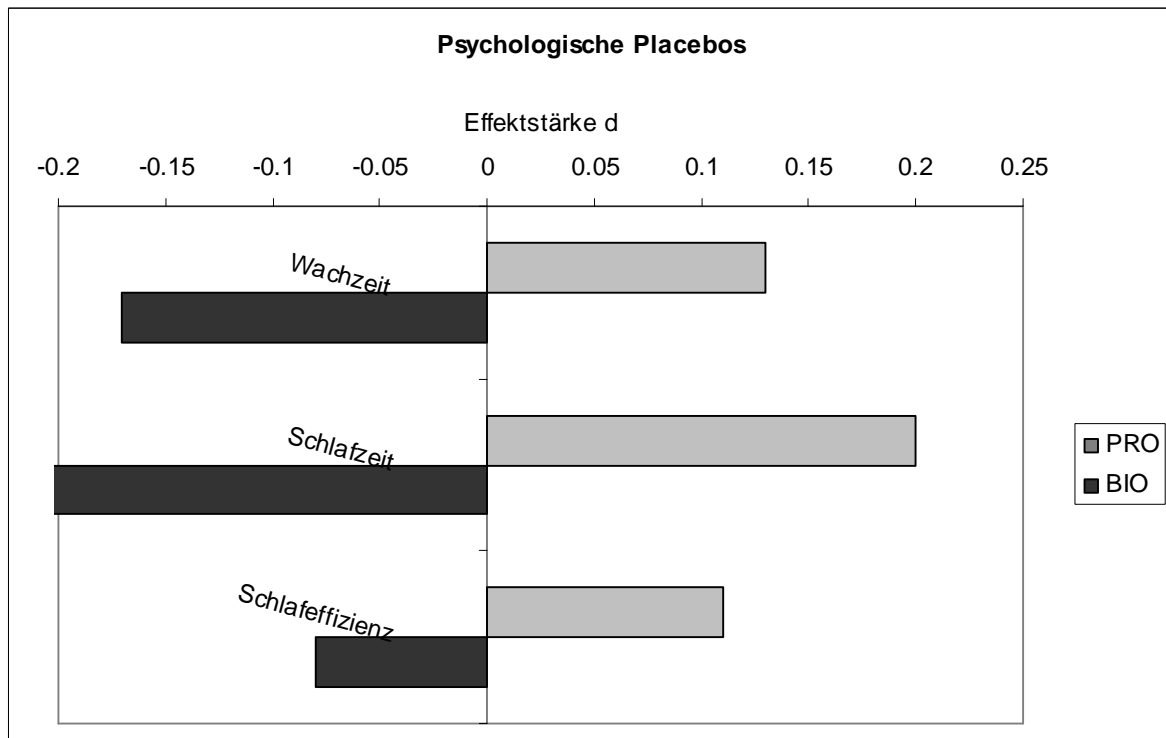


Abbildung 15

PRO vs. BIO (Effektstärke/psychologisches Placebo minus Effektstärke/Warteliste) in 34 Studien zu Schlafstörungen (Belanger et al., 2007).

5.3 Untere Evidenzstufe: Meta-Analysen zur Placeboresponse und zum Anteil der Placeboresponse an der Gesamtresponse

In einer Metaanalyse von 198 Studien zu Osteoarthrosen berichteten Zhang et al. (2008) und Doherty and Dieppe (2009) die Punktschätzer für verschiedene Zielparameter, darunter jeweils mindestens einen PRO, ORO und BIO (s. Tabelle A2 im Anhang). Alle PROs (Schmerz, Steifigkeit, Funktion) wiesen auf signifikante Besserungen unter Placebobehandlung hin (Effektstärken zwischen 0,43 und 0,51). In 14 unbehandelten Kontrollgruppen fand sich hingegen keine Besserung bezüglich Schmerz, was auf einen genuine Placeboeffekt hinweist. Der einzige ORO, nämlich das globale Arzturteil, wies ebenfalls eine signifikante Verbesserung unter Placebo auf (ES 0,66). Die Ergebnisse der BIOs waren gemischer: Während der Parameter Gehzeit/Gehdistanz eine signifikante Verbesserung zeigte (ES 0,22), deuteten die Ergebnisse für Gelenkspaltenbreite auf eine signifikante Verschlechterung hin (ES -0,32). Keine signifikanten Veränderungen zeigten die BIOs Quadrizepsstärke, Knieumfang und Bewegungsradius. Die Ergebnisse dieser Metaanalyse sind mit der Annahme konsistent, dass die Placeboresponse auf BIOs kleiner ist als die Placeboresponse auf PROs und OROs, wobei es zudem starke Unterschiede

zwischen verschiedenen BIOs zu geben scheint. Die Änderungen in den Placebogruppen für OROs und PROs hingegen unterschieden sich nicht wesentlich voneinander.

Fulda and Wetter (2008) untersuchten die Placeboresponse beim Restless-Legs Syndrom und analysierten dabei PROs (Symptomreduktion, subj. Schlafqualität, subj. Schlafdauer, Tagesschläfrigkeit, Lebensqualität), BIOs (periodische Beinbewegungen, Schlafeffizienz, Gesamtschlafdauer) und einen ORO (klinische Response). Bis auf den BIO Schlafeffizienz besserten sich alle Zielkriterien in der Placebogruppe signifikant. Der Placeboanteil an der Gesamtresponse variierte zwischen 13% und 64% (s. Abb. 16). Dabei war der Anteil der klinischen Response (ORO; 59%) vergleichbar mit den Anteilen, die sich für die überwiegende Anzahl der PROs ergab (56-60%). Die BIOs zeigten entweder hohe (Gesamtschlafdauer; 65%) oder niedrige Placeboanteile (periodische Beinbewegungen, 13%; Schlafeffizienz, 19%) an der Gesamtresponse. Auch hier weisen die Ergebnisse also darauf hin, dass sich PROs und OROs nicht wesentlich hinsichtlich der Placeboresponse unterscheiden. Die Placeboansprechrates verschiedener BIOs wiederum scheint parameterspezifisch stark zu variieren.

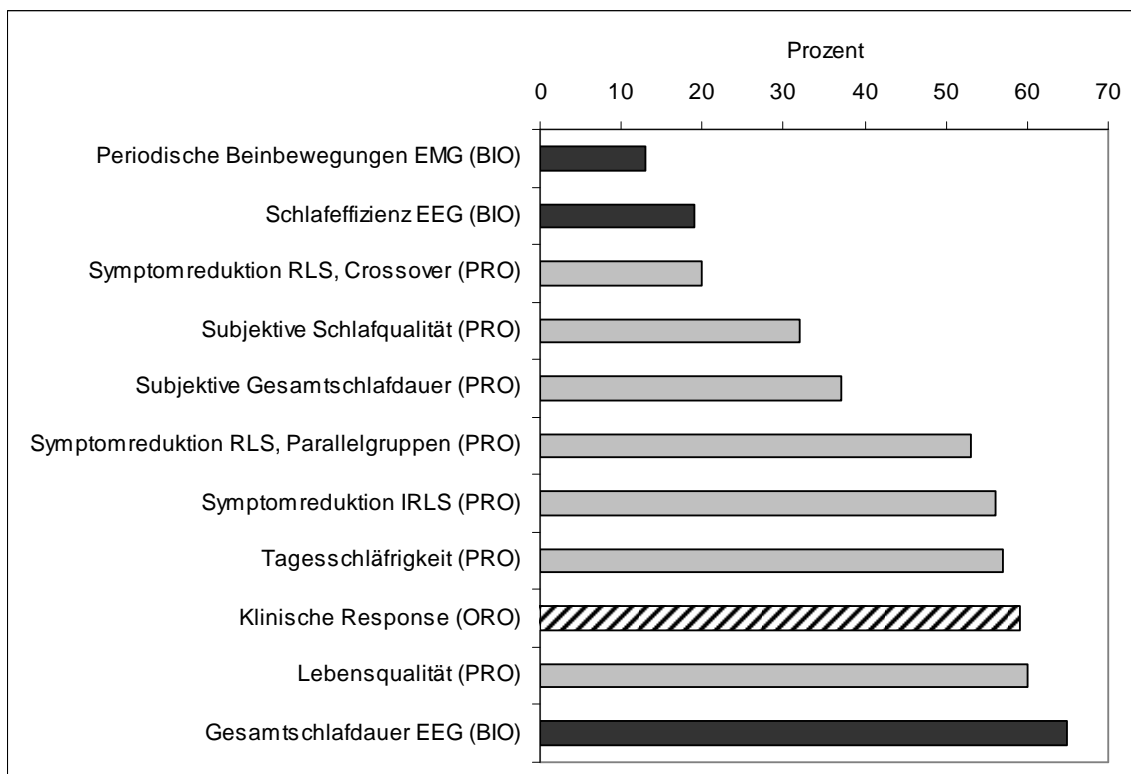


Abbildung 16

Anteil der Placeboresponse an der Gesamtresponse in einer Meta-Analyse von 60 Studien zur Insomnie (Fulda and Wetter, 2008). PROs in grau, OROs schraffiert, BIOs in schwarz.

McCall et al. (2003) untersuchten die Placeboresponse bei Schlaflosigkeit und analysierten dabei sowohl die subjektive Einschlafzeit (PRO) als auch die (anhand des EEGs ermittelte) objektive Einschlafzeit (BIO). Nur die subjektive Einschlafzeit zeigte eine signifikante Besserung unter Placebogabe (PRO: Reduktion um 13,1 min, 95% CI 9,2, bis 17,0, n=5; BIO: 2,5 min, 95% CI -5,9 bis 10,9; n=4), was wiederum auf einen Unterschied zwischen PROs und BIOs hinweist.

In der Meta-Analyse von Rief et al. (2009) Depression wurden in 28 Studien Depressionsscores nach Einschätzung der Patienten (PRO), und in 93 Studien Depressionsscores nach Einschätzung der Ärzte (ORO) berichtet. Der Punktschätzer für die Reduktionen der Depressionsscores lagen für PROs bei 0,67 (95% CI 0,49 bis 0,85), und damit signifikant unter denen für OROs (1,85; 95% CI 1,60 bis 2,01). Der Anteil an der Gesamtresponse zeigte interessanter Weise jedoch keinen Unterschied zwischen PROs und OROs (67% für beide). In einer Meta-Analyse zu sozialer Phobie (Oosterbaan et al. 2001; 15 Studien) lag der Placeboanteil am Gesamteffekt für die vom Arzt beurteilte Response (ORO; 44%) leicht höher als der Placeboanteil für die von Patienten beurteilten Symptomreduktion (PRO; 37%). Zusammengenommen legen diese Ergebnisse nahe, dass die Placeboresponse bei psychiatrischen Erkrankungen für OROs etwas höher liegen könnte als für PROs.

Bei gastroenterologischen Erkrankungen wurden in mehreren Meta-Analysen sowohl subjektive Symptombesserungen als auch objektive Veränderungen (endoskopische und histologische Abheilungen) analysiert (s. Abb. 17). Pace et al. (1995) analysierten die Placeboresponse in 22 Studien zur Refluxösophagitis. Während 27% der placebobehandelten Patienten das Kriterium für eine endoskopische Abheilung erfüllten, zeigten sich nur 12% symptomfrei. Auch der Placeboanteil an der Gesamtresponse lag für die endoskopische Abheilung höher (57%) als für die subjektive Response (38%). Des Weiteren befassten sich drei Meta-Analysen (Garud et al., 2008; Ilnyckij et al., 1997; Meyers and Janowitz, 1989) mit der Placeboresponse bei Colitis ulcerosa (s. Abb. 17). Es fanden sich nur geringe Unterschiede innerhalb der einzelnen Meta-Analysen zwischen klinischen (PROs/OROs) und biologischen Maßen (BIOs) für Response und Remission. Die Ergebnisse zu gastroenterologischen Erkrankungen ergaben demnach teilweise für BIOs sogar stärkere Besserungsraten unter Placebo als für PROs. Wie hoch der Anteil der Spontanremission an diesen Besserungen ist, kann jedoch ohne unbehandelte Kontrollgruppen wiederum nicht beurteilt werden.

In einer Meta-Analyse zum Reizdarmsyndrom (IBS) fanden Ford et al. (2010) heraus, dass die Placeboresponse für Zielkriterien, die vom Arzt eingeschätzt wurden (OROs), größer war als für Maße, die vom Patienten eingeschätzt wurden (PROs).

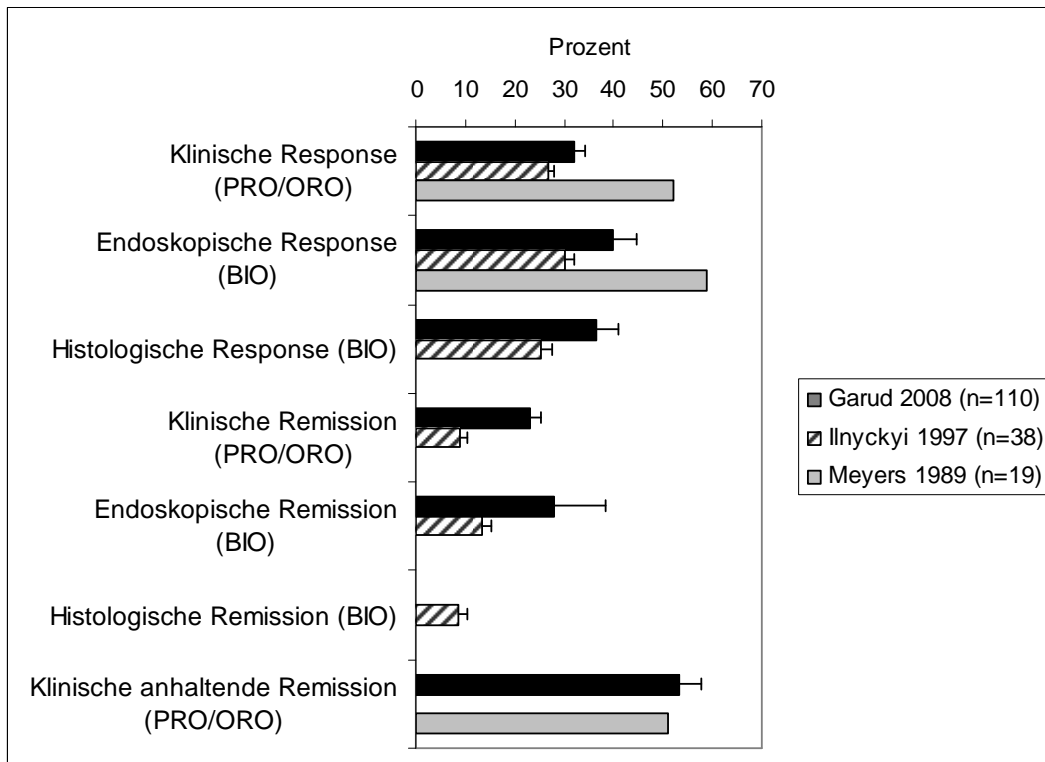


Abbildung 17

Placeboresponse in drei Meta-Analysen zur Colitis ulcerosa (\pm SE, sofern verfügbar).

In einer Meta-Analyse von acht Studien zur benignen Prostatahyperplasie (Boyle et al., 2004) zeigten sich sowohl für PROs (Nykturie, Symptomenscore) als auch für den einzigen BIO (maximale Harnflussrate) signifikante Besserungen unter Placebo. Der Placeboanteil an der Gesamtresponse lag für den Symptomenscore bei 95%, für Nykturie bei 63% und erreichte auch für die maximale Harnflussrate noch 54%. Hier ist also die Placeboresponse auf BIOs niedriger als auf PROs. Die Ergebnisse legen jedoch auch nahe, dass eine symptomatische Besserung bei dieser Erkrankung tendenziell auch mit einer Besserung der physiologischen Harnblasenfunktion einhergehen könnte. Endgültige Aussagen darüber, ob die Besserungen auf einen Placeboeffekt zurückzuführen sind, können jedoch mangels Studien mit unbehandelten Kontrollen derzeit nicht getroffen werden.

In einer Meta-Analyse zur Placeboresponse bei Reizblase (Mangera et al., 2011) zeigten drei von vier PROs (Inkontinenzepisoden/tag, Miktionshäufigkeit/Tag, Urge-Inkontinenzepisoden/Tag) signifikante Besserungen unter Placebo. Der einzige biologische Parameter, nämlich die ‚maximale zystometrische Kapazität‘, zeigte hingegen eine signifikante Verschlechterung an. Ein Review von 30 Studien zur Placeboresponse bei Symptomen des unteren Harntrakts (van Leeuwen et al., 2006) wies auf beträchtliche Verbesserungen an PROs hin (Stressinkontinenz, 32-64%, Inkontinenzepisoden, 32-64%, Symptomenscore 9-

34%), während funktionelle BIOs deutlich kleinere Veränderungen unter Placebo zeigten (Entleerungsvolumen, 5-6%, Maximaler Harnfluss, 2-9%). Der einzige BIO, nämlich das Prostatavolumen, verschlechterte sich um 2-14%. Zusammengenommen weisen die Ergebnisse zu Symptomen des unteren Harntrakts darauf hin, dass BIOs eine niedrigere Placeresponse zeigen als PROs. Innerhalb der Gruppe der BIOs könnten funktionelle Parameter etwas besser auf Placebogabe ansprechen als strukturell-anatomische Maße.

In einem Review von 13 Studien zur pulmonalen arteriellen Hypertonie (Helman et al., 2007) zeigten alle BIOs (6-min Gehstrecke, pulmonärer vaskulärer Widerstand, mittlerer pulmonaler arterieller Druck, kardialer Index, gemischte venöse Sauerstoffsättigung), ebenso wie der Symptomenscore (Borg Dyspnoe Score) eine Verschlechterung unter Placebogabe an. Klinisch reagierten 12% der Patienten in der Placebogruppe mit einer Verbesserung und 9% mit einer Verschlechterung des NYHA-Scores um mindestens einen Grad. Die entsprechenden Werte für Verumtherapie lagen bei 31,2% für die klinische Response und bei 4,8% für den NYHA-Score. Die Placebothherapie konnte also das Fortschreiten der pulmonären arteriellen Hypertonie in jedem Fall nicht verhindern. Ob die Placebogabe das Fortschreiten der Erkrankung aber evtl. verlangsamt hat, ist mangels Studien mit unbehandelter Kontrollgruppe derzeit nicht zu beurteilen.

5.4. Zusammenfassende Bewertung zur Evidenzlage bezüglich der Fragen 2 und 3

Die Ergebnisse aller drei Evidenzstufen weisen konsistent darauf hin, dass Krankheitsmaße, die auf der subjektiven Einschätzung eines Arztes oder Beobachters beruhen (OROs), ähnlich große, oder sogar größere Placeboeffekte zeigen können (z. B. bei IBS und Depression) wie Krankheitsmaße, die auf der subjektiven Einschätzung des Patienten (PRO) beruhen. Zudem lassen sich Besserungen unter Placebo, die auf subjektiven Einschätzungen des Patienten beruhen, durchaus mit physiologischen Messungen (z. B. dem EEG bei Schlafstörungen) objektivieren. Es scheint aber auch biologische Zielkriterien zu geben, die wenig bis gar nicht auf Placebogabe ansprechen. So ergab sich für Laborparameter und krankheitsbedingte strukturelle Veränderungen bisher keine Evidenz für Placeboeffekte. Die Ergebnisse sind konsistent mit der Annahme, dass Placebointerventionen nicht das Fortschreiten der Erkrankung auf Organebene verhindern, jedoch die Symptomatik subjektiv und objektiv, und je nach Krankheitsbild in stärkerem oder geringerem Ausmaß, verbessern können. Weitere gut kontrollierte Studien sind jedoch nötig, um diese Ergebnisse zu untermauern und um zu klären, inwieweit nicht auch „härtere“ biologische Maße mancher Erkrankungen auf Placebogabe ansprechen können.

6. Zu Frage 4: Finden sich Hinweise auf die Bedeutung der initialen Krankheitsschwere bzw. auch der spezifischen Behandlungsvorgeschichte (z.B. bei der Behandlung sog. therapieresistenter Fälle)?

6.1. Untere Evidenzstufe: Meta-Analysen zur Placeboresponse und zum Anteil der Placebo- and der Gesamtresponse

In Meta-analysen von Placeboeffekten (d.h. Meta-analysen der höchsten Evidenzstufe) wurden bisher keine Einflüsse von initialer Krankheitsschwere oder Behandlungsvorgeschichte auf die Größe des Placeboeffekts untersucht (s. Tabelle A3 im Anhang). Mehrere Meta-analysen zur Placeboresponse (d.h. Meta-analysen der untersten Evidenzstufe) suchten jedoch nach Prädiktoren für die Besserungen unter Placebogabe, primär um Placeboeffekte in klinischen Studien besser kontrollieren zu können. Die Ergebnisse für Einflüsse der initialen Krankheitsschwere, Behandlungsvorgeschichte und weiterer Krankheitsmerkmale sind in Tabelle 3 zusammengestellt (für eine detaillierte Auflistung s. Tabelle A4 im Anhang).

Der Zusammenhang zwischen initialer Symptomschwere und Placeboresponse scheint den Ergebnissen zufolge stark von der Indikation abzuhängen. In sechs Schmerz-Reviews fanden sich Anhaltspunkte dafür, dass eine höhere initiale Schmerzintensität mit einer größeren Placeboresponse korreliert. Auch bei Schlaflosigkeit (1 Übersicht), funktioneller Dyspepsie (2), sexueller Dysfunktion (1), Reizblase (1), Allergie (1), Müdigkeit bei Krebserkrankungen (1) und M. Crohn (1) gab es Anhaltspunkte für eine größeren Placeboresponse bei stärkeren initialen Symptomen. Allerdings wies bei M. Crohn eine Meta-Analyse auch auf den umgekehrten Zusammenhang hin, also auf eine größere Placeboresponse bei geringeren initialen Symptomen. Bei Depression wiederum fanden sogar vier Studien Hinweise darauf, dass geringere initiale Symptome mit einer höheren Placeboresponse assoziiert waren, vier weitere Studien hingegen fanden keinen Zusammenhang. Keine Hinweise auf einen Zusammenhang zwischen initialer Symptomschwere und Placeboresponse zeigten sich bei sozialer Phobie (1), Angststörungen (1) Schizophrenie (1) und Epilepsie (2).

Bei M. Crohn war ein geringeres Ausmaß der Entzündung im Darm (1) mit einer größeren Placeboresponse assoziiert. Bei Epilepsie sagte eine kürzere Krankheitsdauer (1) eine größere Placeboresponse vorher. Eine Meta-Analyse zu Depressionen fand ebenfalls diesen Zusammenhang, eine andere hingegen nicht. Eine größere Placeboresponse bei Depressionen war zudem assoziiert mit einem höheren Patientenanteil ohne Antidepressiva-Vorerfahrung (1) und weniger behandlungsresistenten Patienten (1). Zudem stand eine geringere Ängstlichkeit zu Beginn (1) sowie das Auftreten der Depression im Rahmen einer

somatischen Erkrankung (1) mit einer besseren Placeboresponse im Zusammenhang. Bei ADHS war eine fehlende Vorerfahrung mit Stimulanzen (1) sowie die Komorbidität mit Tics (1) ein positiver Prädiktor für die Placeboresponse. Eine höhere Placeboresponse bei Manie war mit mehr Psychosezeichen zu Beginn assoziiert (1).

Tabelle 3: Einflüsse der initialen Krankheitsschwere und Behandlungsvorgeschichte auf die Placeboresponse bei verschiedenen Indikationen.

Erkrankung	Positive Assoziation mit Placeboresponse (Anzahl Studien)	Keine Assoziation mit Placeboresponse (Anzahl Studien)
Neuropathische Schmerzen	Höhere initiale Schmerzintensität (3)	
Fibromyalgie	Höhere initiale Schmerzintensität (1)	
Osteoarthrosen/-arthritis	Höhere initiale Schmerzintensität (2)	
Epilepsie	Kürzere Krankheitsdauer (1), höheres Alter bei Krankheitsbeginn (1)	Initiale Anfallshäufigkeit (2)
Depression	Geringere initiale Symptomschwere (4) Kürzere Krankheitsdauer (1) Kürzere Dauer der aktuellen Episode (1) Patienten ohne Antidepressiva-Vorerfahrung (1) Weniger behandlungsresistente Patienten (1) Komorbidität mit Axis-III Krankheiten (z. B. Diabetes) (1) Geringere initiale Ängstlichkeit (1)	Initiale Symptomschwere (4) Alter bei Krankheitsbeginn (1) Krankheitsdauer (1) Dauer der aktuellen Episode (1)
Manie	Mehr Psychosezeichen (1)	
Soziale Phobie		Initiale Symptomschwere (1)
Angststörungen		Initiale Symptomschwere (1) Krankheitsdauer (1) Alter bei Krankheitsbeginn (1)
Schizophrenie		Initiale Symptomschwere (1)
ADHS	Vortherapie (keine Stimulanzen) (1) Komorbidität (Tics) (1)	
Schlaflosigkeit	Höhere initiale Symptomschwere (1) größere Symptomvariation während Lead-in (1)	
Funktionelle Dyspepsie	Höhere initiale Symptomschwere (2) Schlechtere Magenentleerung zu Beginn (1) Größere Symptomverschlechterung während Run-in (1) Weniger konsistentes Symptomenmuster (1)	Krankheitsdauer (1)
Reizdarmsyndrom	Keine Rom-Diagnosekriterien (1)	
M. Crohn	Höhere initiale Symptomschwere (1) Geringere initiale Symptomschwere (1)	

	Vorherige Steroidtherapie (1) Normale rektale Mucosa (1) Beschränktes Ausmaß der Colitits (1) Kürzlich erreichte (medikamentös) oder aufrechterhaltene Remission (1)	
Chronische Pankreatitis		Alkoholische Genese (1)
Sexuelle Dysfunktion (weibl.)	Höhere initiale Symptomenschwere (1)	
Reizblase	Höhere initiale Symptomenschwere (1)	
Allergie	Höhere initiale Symptomenschwere (1)	
Müdigkeit bei Krebserkrankungen	Höhere initiale Symptomenschwere (1)	
Psoriasis	Mindestsymptomenschwere als Einschlusskriterium (1) Keine Psoriasisassoziierte Arthritis als Behandlungsindikation (1)	

Im Rahmen der Literaturanalyse wurden weitere Einflussfaktoren identifiziert, die teilweise wiederholt mit einer größeren Placeboresponse bzw. einem größeren Placeboeffekt korrelierten. Diese sind nach Kontextfaktoren und Krankheiten geordnet in Tabelle 4 zusammengefasst. Weitere, designspezifische Einflussfaktoren sind in den Tabellen 3 und 4 im Anhang zusammengefasst.

Ein Kontextfaktor zu Patienten-Charakteristika, der wiederholt mit der Placeboresponse korrelierte, war das Alter: Kinder und Jugendliche zeigten bei verschiedenen Erkrankungen eine größere Placeboresponse als Erwachsene, während in Erwachsenenpopulationen eher die Älteren eine größere Placeboresponse zeigten. Ein häufig wiederkehrender Einflussfaktor der Arzt-Patient-Interaktion auf die Placeboresponse war die Häufigkeit der Arzt-Patient-Kontakte, wobei in der Regel häufigere Kontakte mit einer größeren Placeboresponse einhergingen. Bezüglich des medizinischen Umfelds als Einflussfaktor fiel vor allem der Faktor des Studienlandes ins Auge, wobei die identifizierten Landesunterschiede krankheitsspezifisch unterschiedlich und heterogen waren. Charakteristika des Arztes wurden bisher in keiner Meta-Analyse als signifikanter Einflussfaktor identifiziert bzw. untersucht.

Zu den Kontextfaktoren, deren Einfluss bisher auch in Meta-Analysen der höchsten Evidenzstufe (direkter Vergleich zwischen Placebo und No treatment) verifiziert werden konnten, zählen neben der Art des Zielkriteriums (s. Kap. 5) auch die Art des Placebos und die Art der Aufklärung: Physikalische Placebointerventionen, und hier vor allem Scheinakupunktur, scheinen besser zu wirken als pharmakologische und psychologische Placebos (physikalische Placebos, und hier vor allem Scheinakupunktur wirken besser als pharmakologische und psychologische Placebos Hrobjartsson and Gotzsche, 2010; Linde et al., 2010a; Linde et al., 2010b), und eine wahrheitsgetreuere Patientenaufklärung ging mit

einem kleineren Placeboeffekt einher als eine eher beschönigende Aufklärung zur Natur der Placebokontrollen (Hrobjartsson and Gotzsche, 2010).

Tab. 4: Kontextfaktoren, die mit der Größe der Placeboresponse bzw. des Placeboeffekts in klinischen Studien in Zusammenhang standen (krankheitsspezifische Anordnung). Wenn nicht anders vermerkt, handelt es sich um Einflussfaktoren für die Placeboresponse.

Erkrankung	Charakteristika des Patienten	Arzt-Patient-Interaktion	Behandlungscharakteristika	Medizinisches Umfeld
Kopfschmerzen	Alter (Kinder)		Art des Placebos (pharm.)	
Migräne (Prophylaxe)				Land (Europa)
Migräne (Akuttherapie)	Alter (Kinder, Jugendliche)		Applikationsart (subkutan)	Land (Nordamerika)
Verschiedene Schmerzsyndrome	Alter (höher) Schmerzätiologie (v.a. HIV-Schmerz)		Art des Verums (Antikonvulsiva)	
Osteoarthritis/-arthritiden			invasives Placebo	
Epilepsie	Alter (Kinder)			
Depression	Ethnizität (Nicht-kaukasisch) Alter (Kinder) Alter (jünger)		Art der Behandlung (pharmakol.) Dosierungsschema (fix)	
Manie	Geschlecht (weibl.), Alter (höher)			
Angststörungen				Land (nicht Italien) Land (Europa)
Suchterkrankungen				Raucherentwöhnungsklinik; Gruppensetting
ADHS	Inattentiver Subtyp; Ethnische Zugehörigkeit (Nicht-kaukasisch)			
GERD			Art des Verums (H2-Rezeptor-Antagonisten)	
Funktionelle Magen-Darmstörungen	Nichtraucher höherer BMI	Höhere Zahl der Arztbesuche; Geringere Frequenz der Arzttermine	Häufigere Placebogabe	Land (Europa)
Entzündliche Darm-erkrankungen		Höhere Zahl der Studienkontakte Häufigere Arzttermine	Häufigere Dosen pro Tag (4 vs. 2),	Land (Europa, Deutschland)
Sexuelle Dysfunktion (weibl.)	Häufigkeit der zufriedenen sexuellen Kontakte während der Behandlung;			
Saisonale	Weibl. Geschlecht			

allergische Rhinitis				
Hauterkrankungen			Art der Applikation	
Hypertonie				Land (kleiner in Deutschland)
Hormonersatztherapie präventiv	Größere Adherence			
Alle		Art des Informed Consent (eingeschränkt) [Placeboeffekt]	Art des Placebos (Akupunktur, andere physikalische Placebos) [Placeboeffekt]	

5.2. Zusammenfassende Bewertung zur Evidenzlage bezüglich der Frage 4

Besserungen von Schmerzen und funktionellen Störungen unter Placebo zeigten eine starke Abhängigkeit vom Ausgangswert. Dieser Zusammenhang geht wahrscheinlich größtenteils auf Regression zur Mitte zurück: häufig werden Patienten nur dann in eine Studie eingeschlossen, wenn sie über ein Mindestmaß an Symptomen berichten, und diese Symptome gehen dann erwartungsgemäß in der Folge auch ohne Placebobehandlung etwas zurück. Einige Studien bei Depression und entzündlichen Darmerkrankungen weisen darüber hinaus darauf hin, dass ein geringeres Krankheitsausmaß sowie eine geringere Anzahl therapieresistenter Patienten mit größeren Verbesserungen unter Placebothherapie einhergeht. Ob das bedeutet, dass eine stärkere Krankheitsausprägung die Ausbildung von Placeboeffekten erschwert, muss in Studien mit unbehandelten Kontrollen geklärt werden. Dieser Schluss erscheint aber sehr plausibel und konsistent mit der Annahme, dass Krankheiten, die weit fortgeschritten sind und bereits zu strukturellen Veränderungen geführt haben, nicht mehr gut auf Placebogabe ansprechen.

Weitere Kontextfaktoren, die wiederholt mit einer größeren Placeboresponse assoziiert waren, beinhalteten das Alter bzw. die Altersgruppe der Patienten, die Häufigkeit der Arzt-Patient-Kontakte und den geographische Studienort. Diese Befunde sind wiederum noch nicht durch Vergleich mit unbehandelten Kontrollen abgesichert; eine weitere Untersuchung erscheint aufgrund der Konsistenz der Ergebnisse über viele Metaanalysen hinweg jedoch sehr viel versprechend zu sein.

7. Zu Frage 5: Wie wird der der Placebo-Wirkung zuzuschreibende Effekt interpretiert? Gibt es hier Hinweise auf Unterschiede nach Diagnosen bzw. Krankheiten?

Die Mehrheit der für diese Expertise gesichteten systematischen Übersichtsarbeiten klinischer Studien interpretiert Placeboeffekte als reale und klinisch relevante Phänomene. Implizit scheint bei einer Reihe von Autoren jedoch die Annahme vorzuherrschen, dass primär subjektive Parameter betroffen sind. In der Tat sprechen die Ergebnisse dafür, dass Placeboeffekte insbesondere bei subjektiven Symptomen und eher globaleren physiologischen Zielgrößen (z.B. Schlafdauer, verifiziert im EEG) auftreten. Dennoch werden Placeboeffekte meist nicht als Artefakte oder Verzerrungen interpretiert, da Bewertungen von Arzt und Patient bzw. auch Ergebnisse bei physiologischen Parametern konsistent erscheinen. Bei biochemischen Parametern werden dagegen kaum signifikante Effekte berichtet. Bei einigen Erkrankungen (z.B. M. Parkinson) sind die Belege – insbesondere aus der Grundlagenforschung - für „objektive“ Effekte besser als bei anderen. Hier spielt jedoch eine wichtige Rolle, ob bei den entsprechenden Erkrankungen überhaupt Untersuchungen erfolgten, bei denen eine „Objektivierung“ angestrebt war. Diese Untersuchungen haben meist einen mehr experimentellen als klinischen Charakter.

Einige Erkrankungen wurden deutlich intensiver untersucht als andere. Insbesondere bei Depressionen, Kopfschmerzen und gastroenterologischen Erkrankungen liegt eine beträchtliche Zahl von Meta-Analysen vor. Die Gutachter haben den Eindruck, dass dies weniger auf ein primäres Interesse an Placeboeffekten in den zugrunde liegenden klinischen Studien selbst zurückzuführen ist, sondern darauf, dass bei diesen Indikationen der Nachweis eines spezifischen Verumeffektes – möglicherweise aufgrund einer hoher Placeboresponse bzw. eines großen Placeboeffektes – häufig scheitert. Dementsprechend scheint bei vielen Analysen letztlich das Ziel zu sein, Ansätze zu erarbeiten, wie Placeboeffekte minimiert und damit placebokontrollierte Studien „optimiert“ werden können (siehe auch Enck et al., 2013 zur Minimierung von Placeboeffekten in klinischen Studien).

Aus den vorliegenden Übersichtsarbeiten ergeben sich aus Sicht der Gutachter ansonsten kaum Hinweise auf krankheitsspezifische Interpretationen von Placeboeffekten. Das Phänomen wird eher allgemein interpretiert. In den vergangenen Jahren gewinnt die lange vernachlässigte Untersuchung von Noceboresponse und Noceboeffekten (also negativen Auswirkungen von Placebogaben) zunehmend an Bedeutung. Diese wurden jedoch für die vorliegende Expertise nicht eingehend bearbeitet. Nach Kenntnis der Gutachter spielen jedoch auch hier krankheitsspezifische Aspekte keine vordringliche Rolle.

8. Zu Frage 6: Welche Forschungsfragen sollten noch gelöst werden? Welche haben eine besondere Priorität?

Die in dieser Expertise zusammengestellten Ergebnisse machen deutlich, dass die gestellten Fragen beim jetzigen Kenntnisstand nur sehr eingeschränkt beantwortet werden können. Für eine bessere Abschätzung der Größe von Placeboeffekten in klinischen Studien wären deutlich mehr dreiarmlige Studien bzw. Studien, die neben einer Placebokontrollgruppe auch eine unbehandelte Kontrollgruppe mitführen, wesentlich. Da solche Studien aufwändig (aufgrund eher kleiner Effektgrößen hohe Fallzahlen notwendig), problematisch (ethische Probleme oder mangelnde Akzeptanz von unbehandelten Kontrollgruppen) und nicht ohne Verzerrungsrisiko (Patienten nicht verblindbar) sind, erscheinen auch Netzwerk-Meta-Analysen sinnvoll, um die Evidenzbasis zu verbreitern. Netzwerk-Meta-Analysen können direkte (Studien mit Placebo- und unbehandelter Kontrolle) und indirekte Vergleiche (Studien Verum vs. Placebo und Studien Intervention vs. unbehandelte Kontrollen) miteinander verknüpfen und sind aus methodischer Sicht rein indirekten Vergleichen deutlich überlegen. Weitere Meta-Analysen von dreiarmligen Studien und zur Placeboresponse scheinen nur sinnvoll, wenn tatsächlich neue Fragestellungen bearbeitet werden. Aus Sicht der Gutachter können aber auch dreiarmlige Studien und Netzwerk-Meta-Analysen nur eine grobe Abschätzung der Größe von Placeboeffekten in der klinischen Forschung erlauben. Die Vorstellung, dass vielfältige Kontextfaktoren Placeboeffekte auslösen und modulieren können, ist theoretisch und empirisch inzwischen gut belegt. Im Rahmen von klinischen Studien sind allenfalls einzelne Kontextfaktoren kontrollierbar; dementsprechend ist eine hohe Variabilität in den Studienergebnissen zu erwarten. Für die Beurteilung der Wirksamkeit von Therapien sind solche Ergebnisse jedoch wichtig. Beispielsweise erscheint es notwendig zu untersuchen, ob unterschiedliche Placebointerventionen (z.B. die Gabe einer Placebopille oder einer Scheinakupunktur) mit verschiedenen großen Placeboeffekten einher gehen.

Ebenso erstrebenswert erscheint den Gutachtern eine weitere Ausdifferenzierung der Befunde bezüglich der Größe von Placeboeffekten auf PROs, OROs und BIOs zu sein. Es sollte geklärt werden, ob die Tatsache, dass die Effekte auf PROs und OROs häufig in der Größe vergleichbar waren, tatsächlich bedeutet, dass die Besserungen auch klinisch bedeutsam sind. So besteht hier immer noch die Möglichkeit, dass sowohl Arzt als auch Patient dazu geneigt sind, in zweideutigen Situationen eine Besserung zu konstatieren, auch wenn tatsächlich keine Besserung stattgefunden hat (der sog. "signal detection error"; siehe hierzu Allan and Siegel, 2002). Eine Möglichkeit zur Beurteilung der klinischen Relevanz der Besserungen wäre hier der Einbezug auch globalerer Gesundheitsmaße in die klinische

Forschung, wie z. B. Fragebogen zur Lebensqualität und Lebenszufriedenheit. Auch weitere Studien zur Frage, welche BIOs durch Placebogabe beeinflusst werden können und welche nicht, erscheint den Gutachtern von großer Bedeutung zu sein. Während strukturell-anatomische und biochemische Maße eher weniger auf Placebogabe anzusprechen scheinen, könnten Placeboeffekte auf globalere physiologische Maße und vegetativ gesteuerte Organfunktionen dennoch von klinischer Relevanz sein. Keine Meta-Analyse untersuchte bisher Placeboeffekte auf immunologische und endokrinologische Maße; angesichts der Ergebnisse aus der psychoneuroimmunologischen Forschung könnten hier jedoch durchaus klinisch relevante Placeboeffekte zu erwarten sein. Weitere Studien sollten auch klären, ob das Krankheitsausmaß eine Rolle für die Ausprägung von Placeboeffekten spielt. Hieraus könnten sich wichtige Anhaltspunkte für den Einsatz von Placebos in der Praxis ergeben.

Unklar erscheint den Gutachtern, in wie weit die Ergebnis solcher Forschung auf praktischer Ebene interpretiert werden können. Zum einen ist unklar, ob die unter Studienbedingungen beobachteten Ergebnisse auf den klinischen Alltag übertragen werden können. Hier gibt es sowohl Argumente, warum Placeboeffekte in der klinischen Routine noch ausgeprägter sein könnten (Enck et al., 2011a), als auch Argumente für das Gegenteil (z. B. Colagiuri, 2010; Hegerl and Mergl, 2010). Zum anderen stellt sich die Frage, ob im Falle relevanter Placeboeffekte, Placebointerventionen eine medizinisch und ethisch vertretbare Therapieoption sein könnten. Eine Placebointervention kann nur erwogen werden, wenn keine wirksame Therapie zur Verfügung steht (Bundesärztekammer, 2011). Zeigt eine dreiarmlige Studie eine Überlegenheit des Verums, ist diese Voraussetzung nicht gegeben. Zeigt sich ein Placeboeffekt, aber kein Unterschied zwischen Verum und Placebo, ist zunächst abzuwägen, ob eine andere wirksame Therapie verfügbar ist. Ist auch dies nicht der Fall, könnte, wenn der Patient ausdrücklich eine Behandlung wünscht, bei entsprechender Information des Patienten eine Placebointervention erwogen werden. Aus Sicht der Autoren existieren aber für alle oder zumindest fast alle Indikationen, die in den dieser Expertise zugrund liegenden Übersichtsarbeiten untersucht wurden, Interventionen, bei denen eine spezifische Wirksamkeit zumindest wahrscheinlich ist. Damit eine Placebointervention eine ethisch und professionell vertretbare Option ist, müssten daher weitere Bedingungen gegeben sein (z.B. der Patient lehnt die spezifisch wirksamen Optionen ausdrücklich ab oder das mit der Therapie verbundene Risiko ist unverträglich).

In aller Regel verwenden Ärzte in der Praxis Therapien, die wirksam sind oder zumindest vom Arzt für wirksam gehalten werden. Nach heutiger Vorstellung gehen auch Veruminterventionen mit Kontexteffekten einher. Aus klinischer Sicht erscheint daher die Frage zentral, wie Kontexteffekte grundsätzlich zum Nutzen des Patienten optimiert werden können. Studien zum Effekt von Placebointerventionen können hier Beiträge leisten, es

scheint jedoch wissenschaftlich, ärztlich und ethisch vorrangig, den Einfluss von Kontextvariablen vor allem im Zusammenhang mit Veruminterventionen zu untersuchen.

Aus Sicht der Autoren ist bisher zu wenig erforscht, warum und unter welchen Umständen Ärzte in der Praxis Placebointerventionen tatsächlich verwenden. Befragungen von niedergelassenen Ärzten zeigen eindeutig, dass diese so genannte unreine Placebos (potenziell aktive Therapien, die jedoch bei der vorliegenden Erkrankung keine spezifische Wirksamkeit haben) deutlich häufiger als reine Placebos (z.B. Lactosetabletten oder NaCl-Injektionen) verwenden (Fässler et al., 2010; Meissner et al., 2012). Die Befragungen, die standardisierte Fragebogen mit vorgegebenen Antwortoptionen verwenden, lassen aber nur sehr begrenzt Einblicke in die tiefer liegenden Beweggründe zu. Es erscheint den Autoren fraglich, ob die in der Stellungnahme der Bundesärztekammer genannten Voraussetzungen für eine vertretbare Placebogabe tatsächlich erfüllt sind. Qualitative Untersuchungen könnten hier weiter gehende Erkenntnisse erbringen.

9. Zusammenfassung

Placeboeffekte werden heute als reale psychobiologische Phänomene gewertet, die ihren Ursprung im therapeutischen Kontext haben. Über die Größe von Placeboeffekten bei verschiedenen Erkrankungen ist bislang noch wenig bekannt. Dieses Gutachten hatte zum Ziel, das bisher verfügbare Wissen zur krankheitsspezifischen Ausprägung von Placeboeffekten zusammenzutragen und zu bewerten.

Im ersten Teil des Gutachtens wurde auf methodische Schwierigkeiten bei der Beurteilung der Größe von Placeboeffekten in klinischen Studien bzw. Metaanalysen von klinischen Studien eingegangen: So ist der Anteil der Besserungen in den Placebogruppen, der kausal mit der Placebointervention in Verbindung steht, häufig nicht von anderen Einflüssen (z. B. Regression zur Mitte, Spontanverlauf der Erkrankung) zu trennen. Studien, die durch einen Vergleich mit unbehandelten Kontrollgruppen eine Abschätzung des wahren Placeboeffekts erlauben, sind vergleichsweise selten. Die Evidenzlage wurde deshalb im Gutachten in drei Abstufungen beurteilt: Als höchste Evidenzstufe wurden Meta-Analysen eingestuft, die einen direkten Vergleich von Placebo- und unbehandelten Kontrollgruppen ermöglichten, als mittlere Evidenzstufe solche, die einen indirekten Vergleich dieser Gruppen über verschiedene Studien hinweg erlaubten, und als unterste Evidenzstufe solche Meta-Analysen, die lediglich eine rein deskriptive Analyse der Veränderungen in den Placebogruppen und ggf. einen Vergleich mit den Effekten in den Verumgruppen zuließen.

Die vorliegende Evidenz aus Meta-Analysen und systematischen Übersichtsarbeiten zur krankheitsspezifischen Ausprägung von Placeboeffekten ist zwar grundsätzlich mit der Annahme gut vereinbar, dass Placeboeffekte bei manchen Erkrankungen größer sind als bei anderen, gleichzeitig macht sie aber deutlich, dass 1) unser Wissenstand sehr begrenzt ist, da Studien und Meta-Analysen mit direkten und indirekten Vergleichen von Placebogruppen und unbehandelten Kontrollgruppen selten sind, und 2) ganz offensichtlich so viele Faktoren Placeboeffekte und Placeboresponse beeinflussen, dass die Art der Erkrankung nur ein Einflussfaktor unter vielen ist. Die Menge der Evidenz aus direkten und indirekten Vergleichen ist nicht nur begrenzt, die Ergebnisse sind zum Teil auch widersprüchlich. Diese Widersprüche könnten dadurch erklärbar sein, dass dreiarmlige Studien und Studien, die in indirekte Vergleiche eingehen, deutlich unterschiedlich sind. Unter kritischer Berücksichtigung von Umfang und Qualität der Evidenz könnte man sagen, dass in *klinischen* Studien Placeboeffekte mit ausreichender Sicherheit nur bei Schmerz wirklich eindeutig nachgewiesen sind. Bei vielen anderen Indikationen scheint das Vorliegen von Placeboeffekten zwar plausibel und wahrscheinlich, es bestehen aber erhebliche Unsicherheiten. Eine zuverlässige Quantifizierung der Größe in Bezug auf einzelne Erkrankungen erscheint daher zum jetzigen Zeitpunkt unmöglich.

Interessanter Weise weisen die Ergebnisse aller drei Evidenzstufen konsistent darauf hin, dass Krankheitsmaße, die auf der subjektiven Einschätzung eines Arztes oder Beobachters beruhen (OROs), ähnlich große, teilweise sogar größere Placeboeffekte zeigen können (z.B. bei funktionellen Darmerkrankungen und Depression) wie Krankheitsmaße, die „nur“ auf der subjektiven Einschätzung des Patienten (PROs) beruhen. Zudem lassen sich Besserungen unter Placebo, die auf subjektiven Einschätzungen des Patienten beruhen, durchaus auch mit physiologischen Messungen (z. B. dem EEG bei Schlafstörungen) objektivieren. Somit scheint das landläufige Argument, dass Placeboeffekte rein subjektive Phänomene sind, widerlegt. „Harte“ biologische Maße (BIOs) könnten aber tatsächlich eher weniger auf Placebogabe ansprechen. So ergaben die Analysen von Laborparametern und krankheitsbedingten strukturellen Veränderungen bisher keine Evidenz für Placeboeffekte. Die bisherigen Ergebnisse zu Placeboeffekten legen den Schluss nahe, dass Placebointerventionen zwar nicht in den pathogenetischen Prozess einer Erkrankung eingreifen, jedoch sehr wohl funktionelle, objektivierbare Besserungen hervorrufen können.

Besserungen von Schmerzen und funktionellen Störungen unter Placebo zeigten eine starke Abhängigkeit vom Ausgangswert. Dieser Zusammenhang geht wahrscheinlich größtenteils auf Regression zur Mitte zurück. Studien zu Depression und entzündlichen Darmerkrankungen weisen hingegen darauf hin, dass ein geringeres Krankheitsausmaß sowie eine geringere Anzahl therapieresistenter Patienten mit größeren Verbesserungen unter Placebothherapie einhergehen könnten. Weitere Studien sollten klären, ob ein geringeres Krankheitsausmaß prinzipiell ein Prädiktor für eine höhere Placeboansprechrate ist. Weitere Kontextfaktoren, die in den Analysen wiederholt mit einer größeren Placeboansprechrate assoziiert waren, beinhalteten das Alter bzw. die Altersgruppe der Patienten (einerseits bei Kindern, andererseits bei älteren Erwachsenen), die Häufigkeit der Arzt-Patient-Kontakte und den geographische Studienort.

Weitere Studien sind nötig, um die aus den Analysen der gegenwärtigen Studien gewonnenen Ergebnisse zu überprüfen und zu untermauern. Während zusätzliche Meta-Analysen aufgrund der begrenzten Studienanzahl mit unbehandelten Kontrollen nur bedingt zielführend erscheinen, wären vor allem weitere gut kontrollierte, klinische Studien erstrebenswert, um die Größe von Placeboeffekten an PROs, OROs und BIOs verlässlich bestimmen zu können. Da nach heutigem Wissen Placeboeffekte nicht nur bei einer Placebogabe auftreten, sondern auch bei der Verabreichung von aktiven Therapien, sollte vermehrt auch an Möglichkeiten der Maximierung des Gesamttherapieeffekts durch Maximierung der Placebokomponente geforscht werden. Hierzu sollten verschiedene Kontextfaktoren als Einflussvariablen von Therapieeffekten systematisch untersucht werden. Zu beachten ist bei all diesen Studien aber, dass bislang noch völlig unklar ist, in welchem Ausmaß Placeboeffekte in klinischen Studien die Realität in der ärztlichen Praxis abbilden.

Ein weiterer Bereich, der aus Sicht der Gutachter noch zu wenig erforscht ist, betrifft die Häufigkeiten und Gründe für die Verabreichung von Placebos in der ärztlichen Praxis. Bisherige Umfragen lassen vermuten, dass die Situationen, in denen Placebos in der Praxis eingesetzt werden, deutlich von den Empfehlungen der Bundesärztekammer abweichen. Qualitative Umfragen könnten hier helfen, die komplexen Gründe der Ärzte für den Einsatz von Placebos in der Praxis besser zu verstehen, nicht zuletzt, um dem möglichen Missbrauch von Placebointerventionen in der Praxis gezielt und effizient vorbeugen zu können.

10. Literatur

- Allan, L.G., Siegel, S., 2002. A signal detection theory analysis of the placebo effect. *Eval Health Prof* 25, 410-420.
- Belanger, L., Vallieres, A., Ivers, H., Moreau, V., Lavigne, G., Morin, C.M., 2007. Meta-analysis of sleep changes in control groups of insomnia treatment trials. *J Sleep Res* 16, 77-84.
- Benedetti, F., 2009. *Placebo effects: understanding the mechanisms in health and disease*. Oxford University Press, USA.
- Bingel, U., Wanigasekera, V., Wiech, K., Ni Mhuircheartaigh, R., Lee, M.C., Ploner, M., Tracey, I., 2011. The effect of treatment expectation on drug efficacy: imaging the analgesic benefit of the opioid remifentanyl. *Sci Transl Med* 3, 70ra14.
- Boyle, P., Robertson, C., Lowe, F., Roehrborn, C., 2004. Updated meta-analysis of clinical trials of *Serenoa repens* extract in the treatment of symptomatic benign prostatic hyperplasia. *BJU Int* 93, 751-756.
- Bundesärztekammer auf Empfehlung ihres Wissenschaftlichen Beirates (Ed.), 2011. *Placebo in der Medizin*. Köln: Deutscher Ärzteverlag.
- Caldwell, D.M., Ades, A.E., Higgins, J.P., 2005. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 331, 897-900.
- Carbajal, R., 1999. [Analgesia with sugar nipples in the newborn]. *Soins Pediatr Pueric*, 22-23.
- Chvetzoff, G., Tannock, I.F., 2003. Placebo effects in oncology. *J Natl Cancer Inst* 95, 19-29.
- Cochrane Collaboration, 2008. *Cochrane Handbook for Systematic Reviews of Interventions*.
- Colagiuri, B., 2010. Participant expectancies in double-blind randomized placebo-controlled trials: potential limitations to trial validity. *Clin Trials* 7, 246-255.
- Colloca, L., Lopiano, L., Lanotte, M., Benedetti, F., 2004. Overt versus covert treatment for pain, anxiety, and Parkinson's disease. *Lancet Neurol* 3, 679-684.
- de Groot, F.M., Voogt-Bode, A., Passchier, J., Berger, M.Y., Koes, B.W., Verhagen, A.P., 2011. Headache: the placebo effects in the control groups in randomized clinical trials; an analysis of systematic reviews. *J Manipulative Physiol Ther* 34, 297-305.
- Di Blasi, Z., Harkness, E., Ernst, E., Georgiou, A., Kleijnen, J., 2001. Influence of context effects on health outcomes: a systematic review. *Lancet* 357, 757-762.
- Doherty, M., Dieppe, P., 2009. The "placebo" response in osteoarthritis and its implications for clinical practice. *Osteoarthritis Cartilage*.
- Enck, P., Bingel, U., Schedlowski, M., Rief, W., 2013. The placebo response in medicine: minimize, maximize or personalize? *Nat Rev Drug Discov* 12, 191-204.
- Enck, P., Klosterhalfen, S., Weimer, K., Horing, B., Zipfel, S., 2011a. The placebo response in clinical trials: more questions than answers. *Philos Trans R Soc Lond B Biol Sci* 366, 1889-1895.

- Enck, P., Klosterhalfen, S., Zipfel, S., 2011b. Novel study designs to investigate the placebo response. *BMC Med Res Methodol* 11, 90.
- Fässler, M., Meissner, K., Schneider, A., Linde, K., 2010. Frequency and circumstances of placebo use in clinical practice--a systematic review of empirical studies. *BMC Med* 8, 15.
- Fernandes, R., Ferreira, J.J., Sampaio, C., 2008. The placebo response in studies of acute migraine. *J Pediatr* 152, 527-533.
- Finniss, D.G., Kaptchuk, T.J., Miller, F., Benedetti, F., 2010. Biological, clinical, and ethical advances of placebo effects. *Lancet* 375, 686-695.
- Ford, A.C., Moayyedi, P., 2010. Meta-analysis: factors affecting placebo response rate in the irritable bowel syndrome. *Aliment Pharmacol Ther* 32, 144-158.
- Fulda, S., Wetter, T.C., 2008. Where dopamine meets opioids: a meta-analysis of the placebo effect in restless legs syndrome treatment studies. *Brain* 131, 902-917.
- Furukawa, T.A., Watanabe, N., Omori, I.M., Churchill, R., 2007. Can pill placebo augment cognitive-behavior therapy for panic disorder? *BMC Psychiatry* 7, 73.
- Garud, S., Brown, A., Cheifetz, A., Levitan, E.B., Kelly, C.P., 2008. Meta-analysis of the placebo response in ulcerative colitis. *Dig Dis Sci* 53, 875-891.
- Hegerl, U., Mergl, R., 2010. The clinical significance of antidepressant treatment effects cannot be derived from placebo-verum response differences. *J Psychopharmacol* 24, 445-448.
- Helman, D.L., Jr., Brown, A.W., Jackson, J.L., Shorr, A.F., 2007. Analyzing the short-term effect of placebo therapy in pulmonary arterial hypertension: potential implications for the design of future clinical trials. *Chest* 132, 764-772.
- Hrobjartsson, A., Gotzsche, P.C., 2001. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *N Engl J Med* 344, 1594-1602.
- Hrobjartsson, A., Gotzsche, P.C., 2004a. Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. *J Intern Med* 256, 91-100.
- Hrobjartsson, A., Gotzsche, P.C., 2004b. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev*, CD003974.
- Hrobjartsson, A., Gotzsche, P.C., 2010. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev*, CD003974.
- Hrobjartsson, A., Kaptchuk, T.J., Miller, F.G., 2011. Placebo effect studies are susceptible to response bias and to other types of biases. *J Clin Epidemiol* 64, 1223-1229.
- Ilnyckyj, A., Shanahan, F., Anton, P.A., Cheang, M., Bernstein, C.N., 1997. Quantification of the placebo response in ulcerative colitis. *Gastroenterology* 112, 1854-1858.
- Khan, A., Faucett, J., Lichtenberg, P., Kirsch, I., Brown, W.A., 2012. A systematic review of comparative efficacy of treatments and controls for depression. *PLoS ONE [Electronic Resource]* 7, e41778.
- Kirsch, I., Sapirstein, G., 1998. Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. *Prevention & Treatment*.

- Krogsboll, L.T., Hrobjartsson, A., Gotzsche, P.C., 2009. Spontaneous improvement in randomised clinical trials: meta-analysis of three-armed trials comparing no treatment, placebo and active intervention. *BMC Med Res Methodol* 9, 1.
- La Mantia, L., Eoli, M., Salmaggi, A., Milanese, C., 1996. Does a placebo-effect exist in clinical trials on multiple sclerosis? Review of the literature. *Ital J Neurol Sci* 17, 135-139.
- Linde, K., Niemann, K., Meissner, K., 2010a. Are sham acupuncture interventions more effective than (other) placebos? A re-analysis of data from the Cochrane review on placebo effects. *Forsch Komplementarmed* 17, 259-264.
- Linde, K., Niemann, K., Schneider, A., Meissner, K., 2010b. How large are the nonspecific effects of acupuncture? A meta-analysis of randomized controlled trials. *BMC Med* 8, 75.
- Mangera, A., Chapple, C.R., Kopp, Z.S., Plested, M., 2011. The placebo effect in overactive bladder syndrome. *Nat Rev Urol* 8, 495-503.
- McCall, W.V., D'Agostino, R., Jr., Dunn, A., 2003. A meta-analysis of sleep changes associated with placebo in hypnotic clinical trials. *Sleep Med* 4, 57-62.
- Meissner, K., Distel, H., Mitzdorf, U., 2007. Evidence for placebo effects on physical but not on biochemical outcome parameters: a review of clinical trials. *BMC Medicine* 5, 3.
- Meissner, K., Hofner, L., Fassler, M., Linde, K., 2012. Widespread use of pure and impure placebo interventions by GPs in Germany. *Fam Pract* 29, 79-85.
- Meyers, S., Janowitz, H.D., 1989. The "natural history" of ulcerative colitis: an analysis of the placebo response. *J Clin Gastroenterol* 11, 33-37.
- Moerman, D.E., 2000. Cultural variations in the placebo effect: ulcers, anxiety, and blood pressure. *Med Anthropol Q* 14, 51-72.
- Pace, F., Maconi, G., Molteni, P., Minguzzi, M., Bianchi Porro, G., 1995. Meta-analysis of the effect of placebo on the outcome of medically treated reflux esophagitis. *Scand J Gastroenterol* 30, 101-105.
- Rief, W., Nestoriuc, Y., Weiss, S., Welzel, E., Barsky, A.J., Hofmann, S.G., 2009. Meta-analysis of the placebo response in antidepressant trials. *J Affect Disord* 118, 1-8.
- Shetty, N., Friedman, J.H., Kieburtz, K., Marshall, F.J., Oakes, D., 1999. The placebo response in Parkinson's disease. *Parkinson Study Group. Clin Neuropharmacol* 22, 207-212.
- van Leeuwen, J.H.S., Castro, R., Busse, M., Bemelmans, B.L.H., 2006. The placebo effect in the pharmacologic treatment of patients with lower urinary tract symptoms. *Eur Urol* 50, 440-452; discussion 453.
- Vase, L., Petersen, G.L., Riley, J.L., 3rd, Price, D.D., 2009. Factors contributing to large analgesic effects in placebo mechanism studies conducted between 2002 and 2007. *Pain* 145, 36-44.
- Vase, L., Riley, J.L., 3rd, Price, D.D., 2002. A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. *Pain* 99, 443-452.
- Wampold, B.E., Minami, T., Tierney, S.C., Baskin, T.W., Bhati, K.S., 2005. The placebo is powerful: estimating placebo effects in medicine and psychotherapy from randomized clinical trials. *J Clin Psychol* 61, 835-854.

Wechsler, M.E., Kelley, J.M., Boyd, I.O., Dutilleul, S., Marigowda, G., Kirsch, I., Israel, E., Kaptchuk, T.J., 2011. Active albuterol or placebo, sham acupuncture, or no intervention in asthma. *N Engl J Med* 365, 119-126.

Zhang, W., Robertson, J., Jones, A.C., Dieppe, P.A., Doherty, M., 2008. The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis* 67, 1716-1723.