

Universität Ulm
Arbeitsgruppe Klinische Ökonomik
Prof. Dr. Franz Porzsolt

Validität homöopathischer Studien

Dissertation
zur Erlangung des Doktorgrades der Medizin
der Medizinischen Fakultät
der Universität Ulm

vorgelegt von
Nicole Annette Nothardt
aus Ulm

2007

Amtierender Dekan: Prof. Dr. Klaus-Michael Debatin

1. Berichterstatter: Prof. Dr. Franz Porzsolt

2. Berichterstatter: Prof. Dr. Marko Kornmann

Tag der Promotion: 20. Juni 2008

Meiner Familie

**In Erinnerung
an
Charlotte Luise Nothardt**

Issues of design always trump issues of analysis.

Gerard E. Dallal, Ph.D.

Inhaltsverzeichnis

Begriffe und Abkürzungen.....	III
1 Einleitung	1
2 Methodik.....	2
2.1 Auswahl der Studien	2
2.2 Ein- und Ausschlusskriterien homöopathischer Studien	2
2.3 Bewertungskriterien zur Evaluation der Validität	
homöopathischer Studien	4
2.4 Darstellung des Kriterienkataloges	4
2.5 Anwendung des Kriterienkataloges.....	9
2.6 Inhalte des Kriterienkataloges.....	10
2.7 Statistische Angaben	24
3 Ergebnisse	24
3.1 Studienselektion.....	24
3.2 Anzahl der teilnehmenden Institutionen im Validitätsvergleich.....	29
3.3 Herstellung der Arzneimittel	30
3.4 Informed Consent.....	31
3.5 Einschlusskriterien	31
3.6 Rekrutierung und Randomisierung von Patienten.....	32
3.7 Ergebnisse zur Durchführung der Randomisierung	38
3.8 Auswertung der Studienpatienten	38
3.9 Intent to Treat.....	43
3.10 Verblindung der Studienteilnehmer	43
3.11 Vergleich der Ausgangsrisiken mittels „Tabelle 1“	44
3.12 Gleichbehandlung der Studienteilnehmer	45
3.13 Gibt es Ansätze, den speziellen Anforderungen der Homöopathie gerecht zu werden?.....	46
3.14 Formulierung des Zieles einer Studie.....	48
3.15 Evaluation der Ergebnisangaben	48
3.16 Durchführung einer Eigenevaluation	50
3.17 Vergleich der subjektiven Validität mit der internen Validität.....	50
3.18 Häufigkeitsverteilung der Validitätskriterien	51

4	Diskussion.....	56
4.1	Einfluss von Studien mit geringer Validität auf die Qualität von Übersichtsarbeiten	56
4.2	Grundsätzliche validitätseinschränkende Faktoren.....	61
4.3	Kriterien mit unklarem Einfluss auf die Validität	64
4.4	Spezielle Problematik der Homöopathie	66
4.5	Ist eine Bearbeitung von Studien in dargestellter Weise zur adäquaten Validitätsbewertung notwendig?.....	71
4.6	Ist eine standardisierte Bewertung von Studienmaterial für systematische Reviews und Meta-Analysen erforderlich?	71
5	Zusammenfassung.....	74
6	Literaturverzeichnis	76
7	Anhang.....	81
8	Tabellen-und Abbildungsverzeichnis.....	85
9	Danksagung.....	86

Begriffe und Abkürzungen

bzw.	beziehungsweise
Critical Appraisal	kritische Auswertung
Critical Appraisal Tools	Mechanismen/Instrumente zur kritischen Bewertung
Guidelines	Richtlinien
IHK	Industrie- und Handelskammer
Informed Consent	Einwilligung des Patienten auf Basis ausführlicher Information und Aufklärung
Intent to Treat	Auswertung von Studienteilnehmern in der Gruppe, in welche sie ursprünglich hinein randomisiert wurden, unabhängig von den Ereignissen im Studienverlauf
Int. Validität	Interne Validität
IQR	Interquartiler Range
Items	gelistete Inhalte/Posten
n	Nummer/Anzahl
n.a.	nicht angegeben
n.e.	nicht evaluiert
NHS	National Health Service
n.v.	nicht valide
Outcome	Ergebnis
Pat.	Patient
Protocols	Protokolle
rand.	randomisiert
rekrut.	rekrutiert
Reviews	Übersichtsarbeiten
Statement	Aussage
Subj. Validität	Subjektive Validität
u.	und
v.	valide
vgl.	vergleiche
vs.	Versus

1 Einleitung

Grundlage der Bewertung jeder Gesundheitsleistung sind wissenschaftliche Daten, die belegen, dass die in diesen Publikationen beschriebenen Effekte auch tatsächlich erreicht werden können. Diese Art der Übereinstimmung wird als Validität bezeichnet. Deshalb ist die Validität wissenschaftlicher Daten als eine unabdingbare Voraussetzung für den Einschluss einer wissenschaftlichen Studie in eine systematische Übersichtsarbeit und als Grundlage für die Akzeptanz von Gesundheitsleistungen zu fordern.

Diese Forderung ist nur umsetzbar, wenn anerkannte Verfahren zur Prüfung der Validität klinischer Studien existieren. Solche Verfahren wurden von verschiedenen Institutionen oder Wissenschaftlern veröffentlicht. Beispiele sind die Empfehlungen des Consort Statements ([www.consort-statement.org/Statement/ revisionedStatement.htm](http://www.consort-statement.org/Statement/revisionedStatement.htm)), der Evidence-Based Medicine Working Group des Mc Master University Health Science Centers, Hamilton, Ontario (10), des Buchs Evidence-Based Medicine von D.L.Sackett (2. Auflage) (29), der Deutschen Cochrane Collaboration, Freiburg (www.cochrane.de), des Deutschen Netzwerks Evidenzbasierte Medizin e.V., Berlin (www.ebm-netzwerk.de) und des Scottish Intercollegiate Guidelines Network (SIGN), part of NHS Quality Improvement Scotland (www.sign.ac.uk).

Da zahlreiche Verfahren zur Prüfung der Validität wissenschaftlicher Studien verfügbar sind, sollten Übersichtsarbeiten zum gleichen Thema, die von verschiedenen Autoren angefertigt wurden, zu ähnlichen Ergebnissen kommen. Am Beispiel homöopathischer Studien haben Linde und Willich (2003) allerdings beobachtet, dass die Bewertung homöopathischer Therapien in verschiedenen systematischen Übersichtsarbeiten nicht zu einheitlichen Ergebnissen führte (22). Die wahrscheinlichste Ursache ist die Anwendung unterschiedlicher Selektionskriterien der eingeschlossenen Studien. Diese Unterschiede können auf unterschiedlichen Suchstrategien oder auf einer unterschiedlich stringenten Anwendung allgemein akzeptierter Validitätskriterien beruhen.

Da die letztgenannte Ursache, die Anwendung von Validitätskriterien, einfach geprüft werden kann, haben wir als Aufgabe unserer Studie definiert, die Validität homöopathischer Studien zu prüfen, welche die Hürde zum Einschluss in eine systematische Übersichtsarbeit überwunden haben.

2 Methodik

2.1 Auswahl der Studien

Um aus der Vielzahl der Studien zum Thema „Homöopathie“ eine überschaubare Zahl der qualitativ besten Arbeiten zu erhalten, haben wir uns auf Studien beschränkt, die in der Cochrane Datenbank (www.cochrane.de) registriert waren und unter dem Suchbegriff „homeopath*“ erfasst wurden. Auf weitere Einschränkungen des Suchbegriffes wurde verzichtet, um ein möglichst breites Suchergebnis zu erhalten. Die in der Cochrane Library verfügbaren Kurzbeschreibungen des Studiendesigns einschließlich der tabellarisch erfassten Daten wurden benutzt, um die prospektiv definierten Einschluss- und Ausschlusskriterien (siehe 2.2) für unsere Studie zu prüfen.

Die in Punkt 2.2.1 und 2.2.2 näher erläuterten Ein- und Ausschlusskriterien dienten zur Einschränkung des Suchergebnisses auf Studien, welche das Kriterium homöopathischer Studien erfüllten und in einer Publikationsform vorlagen, die eine ausführliche Analyse dieser Studien erlaubten.

2.2 Ein- und Ausschlusskriterien homöopathischer Studien

2.2.1 Einschlusskriterien:

- 1) Es wurden nur Studien eingeschlossen, in welchen ein Arzneimittel im klassisch homöopathischen Sinn nach Samuel Hahnemann, definiert in der 6. Fassung von Samuel Hahnemann's Organon der Heilkunst verwendet wurde (11).
- 2) Die Studie musste in deutscher oder englischer Sprache publiziert sein.

2.2.2 Ausschlusskriterien:

- 1) Mischungen von Einzelmitteln (genannt „Komplexe“), weil diese nicht nach homöopathischen Wirkprinzipien geprüft sind und sich die einzelnen verwendeten Komponenten in ihrer Wirkung angeblich gegenseitig aufheben können.

- 2) Arzneimittel, welche nicht nach Hahnemann potenziert wurden (z.B. Extrakte, einfache Verdünnungen) (11).
- 3) Potenzierte Allergene (gehören in den Bereich der Isopathie) sowie potenzierte Bakterien, Viren oder Pilze.
- 4) Phytotherapeutische Arzneimittel und rein prophylaktisch verabreichte Arzneimittel (in der klassischen Homöopathie wird keine Prophylaxe betrieben sondern symptomorientiert behandelt).
- 5) Zusätzlich wurden nicht publizierte Arbeiten und Veröffentlichungen ausgeschlossen, welche nur als Kongressberichte, Protokolle oder Dissertationsarbeiten zur Verfügung standen.

Die Ausschlusskriterien 1.) bis 4.) wurden gemeinsam mit Frau Cornelia Mantel-Arnold (München-Eching), Mitglied des Homöopathie Forums, einer Organisation für klassisch homöopathisch arbeitende Heilpraktiker, erarbeitet.

2.3 Bewertungskriterien zur Evaluation der Validität homöopathischer Studien

Um die homöopathischen Einzelstudien auf ihre Validität hin zu überprüfen, wurde ein Kriterienkatalog entwickelt.

2.3.1 Bestandteile des Kriterienkataloges

Die Validitätskriterien orientierten sich an den etablierten Leitlinien zur Bewertung von Therapiestudien nach David L. Sackett (28) und der Consort-Gruppe (24).

Der Katalog besteht aus insgesamt 26 Punkten.

Diese beinhalten

- 1) Punkt 1-22: 22 differenzierte Fragestellungen zu Studiendesign und -durchführung sowie zur Beschreibung der Studienergebnisse.
- 2) Punkt 23: eine subjektive Bewertung der Validität.
- 3) Punkt 24: eine interne Validitätskontrolle.
- 4) Punkt 25: Freitext-Anmerkungen zu Problemen, die im Rahmen der 22 Fragestellungen nicht erfasst wurden.
- 5) Punkt 26: eine Graphik des Studiendesigns.

Die Fragestellungen wurden so angelegt, dass sie entweder mit „zutreffend/ja“ oder „nicht zutreffend/nein“ beantwortet werden konnten.

In Punkt 2.4 werden die Bewertungskriterien im Detail dargestellt.

2.3.2 David Sackett und Consort-Gruppe

Prof. Dr. David L. Sackett ist Gründer der Abteilung für klinische Epidemiologie der McMaster Universität in Hamilton, Ontario, Kanada und Gründer und ehemaliger Direktor des Center for Evidence-Based Medicine des John Radcliffe Hospitals in Headington, Oxford, England. Prof. Sackett gilt als einer der Pioniere der evidenzbasierten Medizin.

Die Consort-Gruppe besteht seit 1996 und ist ein Zusammenschluss internationaler Wissenschaftler (Kliniker, Statistiker, Epidemiologen) und Herausgeber medizinischer Fachjournale. Ziel der Consort Gruppe ist es, den Lesern wie auch den Autoren von randomisierten kontrollierten Studien einen Zugang zu mehr Transparenz zu verschaffen. Zu diesem Zweck wurde eine Checkliste und ein Flussdiagramm entwickelt, welche als Werkzeuge zu einer adäquaten Bewertung respektive Beschreibung solcher Studien benutzt werden können. Die Checkliste und das Flussdiagramm sind mit ausführlicher Erläuterung im Internet unter www.consort-statement.org frei verfügbar und wurden im Jahre 2001 neu überarbeitet. Das Consort Statement wird mittlerweile von über 150 internationalen Fachzeitschriften unterstützt und angewendet, unter Anderem vom British Medical Journal, dem Lancet und dem Journal of the American Medical Association.

2.4 Darstellung des Kriterienkataloges

Der Aufbau des Kriterienkataloges stellt sich wie folgt dar:

1. Wie viele Institutionen haben an der Studie teilgenommen?

1.1 Anzahl der Institutionen: $n=x$

a) 1-2 Institutionen (monozentrisch/bizentrisch)

b) 3 oder mehr Institutionen

c) nicht beschrieben

2. Herstellung der Arzneimittel:

2.1: Experimentelles Arzneimittel

- a) intern
- b) extern
- c) nicht beschrieben

2.2: Kontrollarzneimittel:

- a) intern
- b) extern
- c) nicht beschrieben

2.3: Übereinstimmende Herstellung:

- a) beide Arzneimittel extern
- b) beide Arzneimittel intern
- c) nicht beschrieben/verschiedene Herstellungsorte

3. Informed Consent eingeholt?

- a) ja, Inhalt klar
- b) ja, Inhalt unklar/nicht angegeben
- c) nein
- d) nicht beschrieben

4. Einschlusskriterien:

- a) differenziert formuliert
- b) unpräzise formuliert
- c) nicht beschrieben

5. Wie viele Patienten wurden rekrutiert?

- a) Anzahl
- b) nicht beschrieben

6. Wurden die Patienten randomisiert?

- a) Anzahl
- b) Anzahl klar
- c) Anzahl unklar
- d) nein/nicht beschrieben

7. Bei Differenz zwischen Anzahl der rekrutierten und randomisierten Patienten (Vergleiche Nr. 5. und 6.): Sind Gründe angegeben?

- a) ja
- b) nein
- c) Welche Gründe wurden angegeben?
- d) Beeinflussen die Gründe die Validität der Studie?
- e) Warum beeinflussen die Gründe die Validität der Studie?

8. Wer hat die Randomisierung durchgeführt?

- a) zentrale (unabhängige) Institution
- b) rekrutierende Institution
- c) nicht beschrieben

9. Wurde der Randomisierungsplan geheim gehalten?

- a) ja
- b) nein
- c) nicht beschrieben

10. Wie viele Patienten waren in der Auswertung einbezogen? (Vergleiche Nr. 5. und 6.):

- a) Anzahl
- b) Anzahl klar
- c) Anzahl unklar/nicht angegeben

11. Wie viele Patienten wurden von der Auswertung ausgeschlossen?

- a) Anzahl
- b) kein Patient ausgeschlossen
- c) Anzahl klar, Ausschlussgründe valide
- d) Anzahl klar, Ausschlussgründe nicht valide
- e) Anzahl klar, Ausschlussgründe nicht angegeben/unklar
- f) Anzahl der ausgeschlossenen Patienten unklar/nicht angegeben
- g) Welche Gründe wurden angegeben?

12. Intent to Treat:

- a) realisiert
- b) nicht realisiert/missachtet
- c) nicht beschrieben

13. Waren die Teilnehmer der Studie verblindet?

- a) Arzt
- b) Patient
- c) Auswerter
- d) niemand
- e) einfach verblindet
- f) doppelt verblindet

14. Gibt es eine Beschreibung des Verblindungsprozesses?

- a) ja, nachvollziehbare Methodik
- b) ja, unklare Methodik
- c) nein

15. Gibt es eine Tabelle 1 (Gegenüberstellung der Ausgangsrisiken)?

- a) ja
- b) nein

16. Werden in Tabelle 1 Items beschrieben, die bekannte Risiken für die Zielerreichung darstellen?

- a) bekannte Risiken
- b) nicht bekannte Risiken
- c) keine Tabelle 1 vorhanden

17. Waren die Risiko-relevanten Items gleich verteilt?

- a) ja
- b) nein
- c) nicht vorhanden

18. Wurden die Studiengruppen bis auf die experimentelle Therapie gleich behandelt?

- a) ja
- b) nein
- c) nicht beschrieben

19. Versucht der Behandlungsplan den speziellen Anforderungen der Homöopathie gerecht zu werden?

- a) ja
- b) nein
- c) Bemerkung:

20. Ist die Hauptfragestellung der Studie klar formuliert?

- a) ja
- b) nein
- c) Bemerkung:

21. Dimension, in welcher das Ergebnis gemessen wird:

21.1: Ist die Dimension angegeben?

- a) ja
- b) nein

21.2 Ist die Skalierung diskriminativ?

- a) ja
- b) nein

21.3 Entsprechen die Bewertungskriterien der Variabilität der eingeschlossenen Krankheitsbilder?

- a) ja
- b) nein
- c) Bemerkung:

22. Eigener „Critical Appraisal“ vorhanden?

- a) ja
- b) nein

23. Subjektive Einschätzung der Validität in Schulnoten von 1-6

24. Anmerkungen

25. Interne Validitätskontrolle in Schulnoten von 1-6

26. Graphische Skizze des Studiendesigns

2.5 Anwendung des Kriterienkataloges

Zur Bearbeitung der Studien anhand des Kataloges wurden die Einzelstudien in 3 Durchgängen bewertet:

- 1) Während der ersten Durchsicht erfolgte eine komplette Beantwortung der Punkte 1. bis 22. Die Skizze des Studiendesigns wurde ebenfalls in der ersten Runde angefertigt. Kritische Faktoren wie beispielsweise unvollständige oder unklare Angaben von Daten, welche nicht im Rahmen der 22 Items erfasst werden konnten, wurden markiert. Abschließend erfolgte für jede Studie eine subjektive Validitätsbewertung im Sinne der Vergabe einer Schulnote von 1,0-6,0 (siehe 2.6 Punkt 23).
- 2) In einer zweiten Durchsicht wurden die Markierungen der unklaren Angaben stichwortartig unter Punkt 24 „Anmerkungen“ beschrieben.
- 3) Die dritte Durchsicht diente zur Überprüfung der internen Validität (siehe 2.6 Punkt 25).

2.6 Inhalte des Kriterienkataloges

Die einzelnen Punkte des Kataloges werden im Folgenden in der Reihenfolge ihrer Gliederung näher erläutert.

1. Teilnehmende Institutionen

Als teilnehmende Institutionen wurden alle Einrichtungen verstanden, welche an der Rekrutierung und Behandlung von Studienpatienten teilgenommen haben. Die Gesamtzahl der teilnehmenden Institutionen wurde vermerkt. Unter „monozentrisch/bizentrisch“ wurden alle Studien mit maximal zwei teilnehmenden Institutionen aufgenommen.

2. Herstellung der Arzneimittel

Das zu untersuchende Arzneimittel wurde als experimentelles Arzneimittel bezeichnet, das etablierte Vergleichspräparat beziehungsweise das Placebo als Kontroll-Arzneimittel. Die Herstellung der Präparate wurde als intern definiert, wenn diese in der hauseigenen Apotheke von einer der teilnehmenden Institutionen stattgefunden hatte. Als extern wurden die Herstellung definiert, wenn die Präparate von einer Pharmafirma bezogen wurden oder von einer nicht an der Studiendurchführung beteiligten Institution bereitgestellt wurden.

3. Informed Consent

Unter dem “Informed Consent“ wird die Einwilligung eines aufgeklärten Patienten verstanden, welcher über die Teilnahme in einer Studie und die Therapiemodalitäten informiert ist. Um den Inhalt des Informed Consent als „klar“ zu werten, musste im Text zumindest eine weiterführende Erläuterung zum Inhalt vermerkt sein. Wurde die Einholung des Informed Consent lediglich erwähnt, aber nicht durch einen Kommentar erläutert, wurde diese Aussage der Kategorie „Informed Consent eingeholt, Inhalt unklar/nicht angegeben“ zugeordnet.

Wenn kein Informed Consent erwähnt war, galt die Frage als „nicht beschrieben“. Als „nicht eingeholt“ wurde die Frage nur bewertet, wenn im Text explizit vermerkt wurde, dass der Informed Consent nicht eingeholt wurde oder ein direkter Hinweis darauf bestand, dass die Patienten nicht über ihre Studienteilnahme informiert waren.

4. Einschlusskriterien

Die Einschlusskriterien müssen stringent definiert sein. Nur so ist eine Übertragung der Studienergebnisse im klinischen Alltag auf die entsprechende Zielgruppe möglich. Als „differenziert formuliert“ wurden die Einschlusskriterien nur dann bewertet, wenn die Auswahl der Studienpopulation nachvollziehbar war und die beschriebenen Einschlusskriterien das zu therapierende Krankheitsbild eindeutig definierten.

5. Zahl der rekrutierten Patienten

Die Anzahl der für die Studie rekrutierten Patienten bezieht sich auf die Patienten, welche für die Studie primär gewonnen wurden, also die Anzahl der Patienten vor Randomisierung. War die Zahl der Patienten nicht vermerkt oder aus dem Text nicht klar zu verstehen, wurde die Frage mit „nicht beschrieben“ beantwortet.

6. Randomisierung der Patienten

Randomisierung von Patienten bedeutet, dass die Patienten gemäß dem Zufallsprinzip den einzelnen Behandlungsgruppen zugeteilt wurden, zum Beispiel per Los oder nach einem im Vorhinein festgelegten Nummerncode. Die Randomisierung von Patienten soll eine homogene Verteilung der Patienten hinsichtlich Ihrer Ausgangsrisiken gewährleisten.

7. Differenz zwischen der Anzahl der rekrutierten und der Anzahl der randomisierten Patienten

Ein Missverhältnis zwischen der Anzahl der rekrutierten und der randomisierten Patienten sollte stets nachvollziehbar sein. Dazu ist die Angabe der Gründe erforderlich. In Teilfrage 7d) wurde nachfolgend evaluiert, ob die angegebenen Gründe die Validität der Studie beeinträchtigen.

8. Durchführung der Randomisierung

Bedeutend ist bei dieser Frage, ob die Randomisierung von einer unabhängigen, also externen Stelle („zentrale Institution“) durchgeführt wurde oder ob die Randomisierung im eigenen Haus statt fand („rekrutierende Institution“).

9. Geheimhaltung des Randomisierungsplans

Eine Randomisierung ist nur glaubwürdig, wenn der Randomisierungsplan bis zur Randomisierung geheim gehalten wurde. Dies hat den Zweck, einer Voreingenommenheit des Behandlers gegenüber Patienten in einer Behandlungsgruppe vorzubeugen. Wurde der Randomisierungsplan entsprechend detailliert beschrieben oder die Zuordnung von extern durchgeführt, so wurde die Frage nach der Geheimhaltung des Randomisierungsplans mit „ja“ bewertet. Ebenso mit „ja“ beantwortet wurden die Studien, welche selbst angaben, dass der Randomisierungscode bis zur Auswertung geheim gehalten wurde. Mit „nein“ wurde die Frage bewertet, wenn in der Studie explizit angegeben wurde, dass der Code vor der Randomisierung gebrochen wurde oder ein eindeutiger Hinweis darauf bestand.

10. Anzahl der ausgewerteten Patienten

Um die Ergebnisse einer Studie verwerten zu können, muss die Anzahl der ausgewerteten Patienten als Zahlenwert angegeben werden. Nur so sind Differenzen zwischen der Anzahl der randomisierten Patienten und der Anzahl der ausgewerteten Patienten nachvollziehbar (siehe auch nachfolgenden Punkt 11). Relative Angaben (z.B. „75%“ oder „die Hälfte aller Patienten.....“) wurden unter „Anzahl unklar/nicht angegeben“ vermerkt.

11. Anzahl der ausgeschlossenen Patienten

Ebenso sollte die Anzahl der ausgeschlossenen Patienten eindeutig dargestellt sein. Die Angabe der Gründe für den Ausschluss von Studienpatienten ist erforderlich, um bewerten zu können, ob das Studienergebnis durch den Ausschluss verzerrt wurde.

War die genannte Anzahl der ausgeschlossenen Patienten rechnerisch nicht mit der Anzahl der ausgewerteten Patienten in Einklang zu bringen oder bestand eine Diskrepanz zwischen Zahlenangaben im Text und den Auswertungstabellen, so wurde die Teilfrage „Anzahl unklar/nicht angegeben“ mit „ja“ beantwortet.

12. Intent to Treat

Das „Intent to Treat“-Prinzip beschreibt die Tatsache, dass Patienten in der Gruppe ausgewertet werden, in die sie zu Studienbeginn randomisiert wurden.

Eine Verletzung des „Intent to Treat“-Prinzips besteht also beispielsweise, wenn Patienten, welche in die Verum-Gruppe randomisiert wurden und ihr Arzneimittel unvollständig einnahmen, anschließend in der Placebo-Gruppe ausgewertet wurden.

Die Frage nach der Realisierung des „Intent to Treat“ wurde mit „nein“ bewertet, wenn:

- 1) dies in der Studie selbst angegeben wurde.
- 2) die Begründung für Patienten, welche nicht mehr in ihrer Studiengruppe ausgewertet wurden, eindeutig auf eine Verletzung des „Intent to Treat“ schließen ließ.

Als „nicht beschrieben“ wurde die Frage beantwortet, wenn

- 1) ein Missverhältnis in der Anzahl der randomisierten Patienten pro Studiengruppe
- 2) und der Anzahl der darin ausgewerteten Patienten ohne nachvollziehbaren oder angegebenen Grund bestand.
- 3) keinerlei Anmerkung zum „Intent to Treat“ vorlagen.

13. Verblindung der Studienteilnehmer

Durch die Verblindung soll erreicht werden, dass das Outcome nicht von den Erwartungen des Patienten, Behandlers oder Auswerter beeinflusst wird, welche mit einer bestimmten Behandlungsmethode verknüpft sind. Dazu konnten die Kategorien „einfach, doppelt oder nicht verblindet („niemand“)" vermerkt werden.

Als „einfach verblindet“ wurden folgende Kombinationen gewertet:

- 1) die Verblindung von nur Patient oder nur Arzt beziehungsweise Auswerter.
- 2) die Verblindung von Patient und Auswerter bei gleichzeitiger Nicht-Verblindung des Arztes/Behandlers.

Als „doppelt verblindet“ wurde gewertet, wenn

- 1) eine Verblindung von Patient und Arzt oder
- 2) von Patient und Arzt und Auswerter vorhanden war.

14. Beschreibung des Verblindungsprozesses

Eine nachvollziehbare Beschreibung des Verblindungsprozesses beinhaltet Informationen zu den verwendeten Substanzen (Verum, Placebo) und deren

Übereinstimmung in Form, Größe, Farbe und Geschmack (Beispiel: „das Placebo bestand aus der selben Grundsubstanz wie das Verum und war identisch in Form, Größe, Farbe und Geschmack“) sowie zur Verabreichung (Beispiel: „in zahlenkodierten, identischen Dispensern“). Aussagen, welche sich auf die stattgefundene Verblindung von Patient und/oder Arzt und/oder Auswerter beschränkten wurde als „Verblindung ja“, jedoch mit „unklarer Methodik“ gewertet.

15. Tabelle 1 (Gegenüberstellung der Ausgangsrisiken)

Die Tabelle 1 findet sich in der Regel eingangs des Methodik-Teils und beschreibt die Charakteristika der Behandlungsgruppen. Die Tabelle 1 dient dem Zweck, die Verteilung der Ausgangsrisiken auf die durch die Randomisierung gebildeten Behandlungsgruppen darzustellen.

16. Beschreibung risikorelevanter Items in Tabelle 1

Die Tabelle 1 soll Items beinhalten, welche auch tatsächlich relevant für das Gleichgewicht der Ausgangspopulation sind. Dazu können Vorbehandlungen mit anderen Arzneimitteln oder Behandlungsformen, Angaben zur körperlichen Konstitution oder Schweregrad der Symptomatik zählen. Alter und Geschlecht der Patienten sind nicht immer relevante Einflussgrößen. Bei der Behandlung der Otitis media bei Kindern spielt das Geschlecht sicher keine bedeutsame Rolle. Die Frage wurde nur mit „ja“ für die Beschreibung bekannter Risiken bewertet, wenn in der Tabelle mindestens zwei risiko-relevante Items vorhanden waren. Die ausschließliche Angabe von Alter und Geschlecht der Studienteilnehmer wurde dabei nicht als Erfüllung des Anforderungskriteriums betrachtet!

17. Gleichverteilung der risiko-relevanten Items

Eine Gleichverteilung risiko-relevanter Items bedeutet nicht, dass eine exakte Zahlengleichheit in den jeweiligen Behandlungsgruppen für alle Items zu finden ist. Jedoch sollte insgesamt eine gewisse Homogenität der Studiengruppen erreicht werden. Befindet sich bei einer Studienpopulation von Herzinfarkt-Patienten beispielsweise eine größere Anzahl von Rauchern in der einen Gruppe und eine ähnliche Überzahl adipöser Patienten in der anderen Gruppe, so wurde das als Gleichverteilung angesehen. Eine angemessene Gleichverteilung spiegelt eine effektive Randomisierung wieder.

18. Gleichbehandlung der Gruppen abgesehen von der experimentellen Therapie

Eine Gleichbehandlung ist notwendig, um den Einfluss anderer Faktoren auf das Outcome zu minimieren. Dazu gehört ein standardisiertes Studienprotokoll, wonach jeder Patient dieselbe Verhaltensinstruktion, Betreuung und Evaluation erhält. Die Beantwortung dieser Fragen mit „ja“ setzte eine kurze Beschreibung des Studienprotokolls, ein Statement, dass die Patienten gleich behandelt wurden oder die Nachvollziehbarkeit dessen aus einer detailliert beschriebenen Studiendurchführung voraus.

19. Übereinstimmung des Behandlungsplans mit den speziellen Anforderungen der Homöopathie

Diese Frage wurde mit „ja“ behandelt, wenn ein Ansatz verwendet wurde, bei dem Studienpatienten eine für sie individuell ausgewählte homöopathische Substanz oder das entsprechende Placebo verabreicht bekamen. Dies beinhaltete Studienkonzepte, bei denen eine gewisse Anzahl an homöopathischen Substanzen präselektiert wurde und jedem Studienpatienten das für ihn am besten geeignete Verum oder Placebo aus dieser Auswahl zugeordnet wurde. Ferner wurden darunter Studienkonzepte verstanden, bei welchen den Patienten sowohl eine Substanz als Basismedikation als auch eine Substanz als „Akutpräparat“ verordnet wurde.

20. Klare Formulierung der Hauptfragestellung

Die Hauptfragestellung muss so verfasst sein, dass für den Leser die Zielsetzung der Studie klar ersichtlich ist. Die Fragestellung wird in der Regel am Ende der Einleitung formuliert.

21. Dimension, in welcher das Ergebnis gemessen wird

Diese Frage wurde in drei Punkte unterteilt:

- 1) Ist die Dimension angegeben? Damit soll zunächst überprüft werden, ob die Zielgrößen mit den entsprechenden Maßeinheiten angegeben sind.
- 2) Kann die Skalierung Unterschiede feststellen? Die Meßmethoden müssen geeignet sein, auch tatsächlich die Größen zu messen, die für die Fragestellung relevant sind.

Die Skalierung, also die Maßeinheiten müssen zudem so gestaltet sein, dass die Unterschiede im Outcome tatsächlich erfasst werden. Bei der Evaluation von Schmerzen ist beispielsweise die Verwendung einer Visual-Analog-Scale erheblich präziser als die Aussagen stark-mittel-wenig.

- 3) Entsprechen die Bewertungskriterien der Variabilität der eingeschlossenen Krankheitsbilder? Krankheitsbilder und Studienpopulationen, welche eine große Bandbreite an Symptomen bieten, benötigen besonders sorgfältig ausgewählte Evaluationskriterien, welche diese Bandbreite auch abdecken. Bei der Beurteilung von Patienten mit chronisch entzündlicher Darmerkrankung ist eine Evaluation der Stuhlfrequenz und der Schmerzen zur Beurteilung des Krankheitsverlaufes sicher nicht ausreichend.

22. Eigener Critical Appraisal

Mit der Frage soll beantwortet werden, ob eine kritische Auseinandersetzung des Autors mit seiner Arbeit stattgefunden hat. Die Frage geht nicht auf den Umfang oder die Qualität dieser Eigenevaluation ein.

23. Subjektive Bewertung der Validität anhand von Schulnoten

Diese subjektive Beurteilung spiegelt den Gesamteindruck der jeweiligen Studie hinsichtlich ihrer Glaubwürdigkeit und Verwertbarkeit wieder. Der Benotung wurden folgende Maßstäbe zu Grunde gelegt:

Note eins: Die im Katalog evaluierten Kriterien konnten zum Großteil positiv beantwortet werden, Fragestellung, Zielsetzung, Ein- und Ausschlusskriterien wurden klar definiert. Das Studiendesign und die Durchführung sind transparent, die Ergebnisse nachvollziehbar, eventuelle Mängel wurden angesprochen.

Note zwei: Die im Katalog evaluierten Kriterien konnten in der Mehrzahl positiv beantwortet werden, die Studie wurde sorgfältig durchgeführt und entspricht dem Standard einer guten, validen Studie.

Note drei: Das Studiendesign und die Durchführung sind zufriedenstellend strukturiert und nachvollziehbar, eine ausführlichere Beschreibung und Einhaltung gewisser Validitätskriterien wäre wünschenswert gewesen.

Note vier: Basiskriterien wie eine durchgeführte Randomisierung und Verblindung wurden beachtet, es bestehen jedoch deutliche Mängel in der Durchführung oder/und in der Formulierung, die Ergebnisse sind nur eingeschränkt beurteilbar.

Note fünf: Die Studie weist erhebliche Mängel auf, Validitätskriterien wurden missachtet, die Studienbeschreibung ist in diversen Abschnitten so unzulänglich, dass über das Vorgehen nur gemutmaßt werden kann.

Note sechs: Die Studie kann die Minimalanforderungen an eine wissenschaftliche Studie nicht erfüllen, da das Studiendesign und die Durchführung nicht mehr nachvollziehbar sind. Ergebnisse sind unvollständig dargestellt und nicht verwertbar.

24. Anmerkungen

Unter Anmerkungen wurden Probleme zur jeweiligen Studie vermerkt, welche im Rahmen der vorangehenden 22 Evaluationsfragen nicht oder nicht in ausreichendem Rahmen bewertet werden konnten. Diese finden in der Ergebnisdarstellung und Diskussion besondere Berücksichtigung.

25. Interne Validitätskontrolle

Die interne Validitätskontrolle soll eine möglichst objektive Gesamtbewertung der Studien ermöglichen und einen Vergleich zu der in Frage 23 getroffenen „subjektiven Einschätzung“ erlauben. Dazu wurden 18 der im Bewertungskatalog erhobenen Kriterien jeweils eine Punktezahl von null bis zwei zugeteilt.

Schema der Punktevergabe:

- 1) Null Punkte: die Beantwortung des Kriteriums bedeutete einen Verstoß gegen die Regeln der Validität
- 2) Ein Punkt: die Beantwortung des Kriteriums ist als Einschränkung der Validität zu sehen

- 3) Zwei Punkte: das Kriterium wurde so beantwortet, dass daraus keine Beeinträchtigung der Validität entstand.

Definition der Validitätskriterien und Punktebewertung:

Als Kriterium zur Beurteilung der internen Validität wurden folgende Aussagen definiert und bewertet:

1) Herstellung der Arzneimittel:

- 2 Punkte: Bei externer, also unabhängiger Herstellung von Test- und Kontroll-Arzneimitteln
- 1 Punkt: Bei interner Herstellung beider Arzneimittelgruppen (über die „Hausapotheke“)
- 0 Punkte: Bei fehlender Angabe zu der Herstellung oder bei verschiedenen Herstellungsorten von experimentellen und Kontroll-Arzneimitteln

2) Informed Consent:

- 2 Punkte: Einholung eines inhaltlich beschriebenen „Informed Consent“
- 1 Punkt: Einholung eines „Informed Consent“, jedoch mit unklarem Inhalt
- 0 Punkte: Es wurde kein „Informed Consent“ eingeholt oder beschrieben

3) Einschlusskriterien:

- 2 Punkte: Bei differenzierter Formulierung der Einschlusskriterien
- 1 Punkt: Bei unpräziser Formulierung
- 0 Punkte: Bei fehlenden Einschlusskriterien

4) Randomisierung:

- 2 Punkte: Stattgefundene Randomisierung der Patienten mit Angabe der Patientenzahl
- 1 Punkt: Beschreibung einer Randomisierung mit unklarer Zahlenangabe
- 0 Punkte: Es lag keine Randomisierung vor

5) Begründung zur Differenz zwischen Anzahl der rekrutierten und randomisierten Patienten:

- 2 Punkte: Vorliegen einer validen Begründung bei einer Differenz zwischen der Anzahl der rekrutierten und randomisierten Patienten
- 1 Punkt: Eine Begründung für die differierende Anzahl der Patienten liegt vor, diese schränkt jedoch die Validität ein
- 0 Punkte: Zur vorhandenen Differenz wurde kein Kommentar angeführt

6) Durchführung der Randomisierung:

- 2 Punkte: Durchführung der Randomisierung von einer unabhängigen Stelle
- 1 Punkt: Randomisierung durch die Institution, an welcher die Studie durchgeführt wurde
- 0 Punkte: Bei fehlender Angabe

7) Geheimhaltung des Randomisierungsplanes:

- 2 Punkte: Bei beschriebener Geheimhaltung des Randomisierungsplans
- 1 Punkt: Bei nicht vorhandenen näheren Angaben. Dabei wurde von einer wahrscheinlichen Geheimhaltung ausgegangen.
- 0 Punkte: Bei beschriebener vorzeitiger Offenbarung des Randomisierungsplans

8) Angaben zur Anzahl ausgewerteter Patienten:

- 2 Punkte: Bei klarer Beschreibung der Anzahl an Patienten, welche in der Auswertung einbezogen waren
- 0 Punkte: Bei unklarer oder fehlender Information

9) Angaben zur Anzahl ausgeschlossener Patienten:

- 2 Punkte: Klare Beschreibung der Anzahl der ausgeschlossenen Patienten mit valider Begründung oder kein Ausschluss von Patienten (das heißt, alle randomisierten Patienten wurden ausgewertet).
- 0 Punkte: Bei angegebener Patientenzahl, jedoch mit fehlenden, unklaren oder nicht validen Gründen sowie bei komplett fehlender Zahlenangabe

10) Intent to Treat:

- 2 Punkte: Bei realisiertem „Intent to Treat“
- 0 Punkte: Bei nicht beschriebenem oder missachtetem „Intent to Treat“

11) Art der Verblindung:

- 2 Punkte: Bei doppelter Verblindung
- 1 Punkt: Bei einfacher Verblindung
- 0 Punkte: Bei nicht-Verblindung

12) Beschreibung des Verblindungsprozesses:

- 2 Punkte: Nachvollziehbare Beschreibung der Verblindung
- 1 Punkt: Eine Beschreibung der Verblindung existiert, der Prozess ist jedoch nicht eindeutig nachvollziehbar
- 0 Punkte: Keine Beschreibung des Verblindungsprozesses

13) Vorhandensein der Tabelle 1:

- 2 Punkte: Eine Tabelle 1 ist vorhanden, darin sind bekannte Risiken für die Zielerreichung enthalten
- 0 Punkte: Eine Tabelle 1 ist nicht vorhanden oder enthält nur irrelevante Items

14) Verteilung der risiko-relevanten Items:

- 2 Punkte: Gleichverteilung der risiko-relevanten Items
- 1 Punkt: die risiko-relevanten Items waren nicht gleich verteilt
- 0 Punkte: es gab keine risiko-relevanten Items (siehe 2.6 16)

15) Formulierung der Hauptfragestellung:

- 2 Punkte: die Hauptfragestellung der Studie ist eindeutig formuliert
- 0 Punkte: die Hauptfragestellung ist nicht klar ersichtlich

16) Angabe der Zielgrößen:

- 2 Punkte: die Zielgrößen, mit denen das Ergebnis gemessen wurde, sind angegeben
- 0 Punkte: keine Angabe der Zielgrößen und Maßeinheiten

17) Skalierung der Zielgrößen:

- 2 Punkte: die Skalierung ist diskriminativ
- 0 Punkte: die Skalierung ist ungenau und vage

18) Ausreichende Erfassung des Krankheitsbildes mittels der Bewertungskriterien:

- 2 Punkte: die Bewertungskriterien entsprechen der Bandbreite des zu messenden Resultates
- 0 Punkte: die Bewertungskriterien erfassen das Resultat nicht ausreichend

Bei acht der 18 Kriterien konnte aus der Beantwortung der Frage nur zwischen Akzeptanz bzw. „kein Verstoß“ oder Nicht- Akzeptanz bzw. „eindeutiger Verstoß“ gegen die Validität unterschieden werden. Deshalb wurden entweder zwei oder null Punkte vergeben.

Folgende Fragen wurden nicht als Validitätskriterium betrachtet:

- 1) Wie viele Institutionen nahmen an der Studie teil?
Ob die Teilnahme möglichst vieler Institutionen eine positive Auswirkung auf die Validität hat, soll bei dieser Arbeit erst geprüft werden.
- 2) Wurden die Gruppen bis auf die experimentelle Therapie gleich behandelt?
Die Frage ist bei vielen Studien schwierig zu beantworten, da häufig nicht genug Informationen vorliegen, um eine eindeutige Aussage zu treffen. Die Beantwortung dieses Kriteriums kann deshalb nicht zur Validitätskontrolle herangezogen werden.
- 3) Versucht der Behandlungsplan den speziellen Anforderungen der Homöopathie gerecht zu werden?
Im Rahmen dieser Arbeit soll untersucht werden, ob Studien, in welchen man sich um einen entsprechenden Behandlungsplan bemühte, ein besseres Ergebnis hinsichtlich ihrer Validität erzielten oder nicht.
- 4) Gibt es einen eigenen „Critical Appraisal“?
Das Vorliegen einer kritischen Stellungnahme der Autoren ist zwar als positiv zu vermerken, es kann jedoch nicht als Validitätskriterium hinsichtlich der Studiendurchführung bewertet werden.

Umrechnung der Punktezahl in Schulnoten:

Die maximal erreichbare Punktzahl pro Studie betrug 36 Punkte. Diese wurden wieder in das Schulnotensystem transformiert um einen Vergleich mit der subjektiven Beurteilung zu ermöglichen. Als Grundlage zur Umrechnung der Punktwerte in ein Schulnotensystem sollte ein anerkanntes und etabliertes Verfahren verwendet werden. Zu diesem Zweck wurde der Punkteschlüssel der Industrie- und Handelskammer (IHK) Baden-Württemberg benutzt, der Umrechnungsfaktor betrug 0,36.

Tabelle 1: Umrechnungstabelle für die Transformierung der Punktzahlen der internen Validität in Schulnoten durch Applikation der IHK- Punkteskala

Abkürzung:

IHK: Industrie- und Handelskammer

IHK- Punkte	Umrechnungsfaktor	Punktzahl der internen Validität	Schulnote
100-92	x 0,36	36-33	1
unter 92-81	x 0,36	unter 33-29	2
unter 81-67	x 0,36	unter 29-24	3
unter 67-50	x 0,36	unter 24-18	4
unter 50-30	x 0,36	unter 18-11	5
unter 30-0	x 0,36	unter 11-0	6

26. Graphik des Studiendesigns

Die Wiedergabe des Studiendesigns als Flussdiagramm dient dem Zweck, eine schnelle Übersicht von der Grundstruktur einer Studie zu erhalten. Zudem erlaubt diese Darstellung einen Überblick über die Entwicklung der Patientenzahlen im Verlauf der Rekrutierung, Randomisierung und Auswertung.

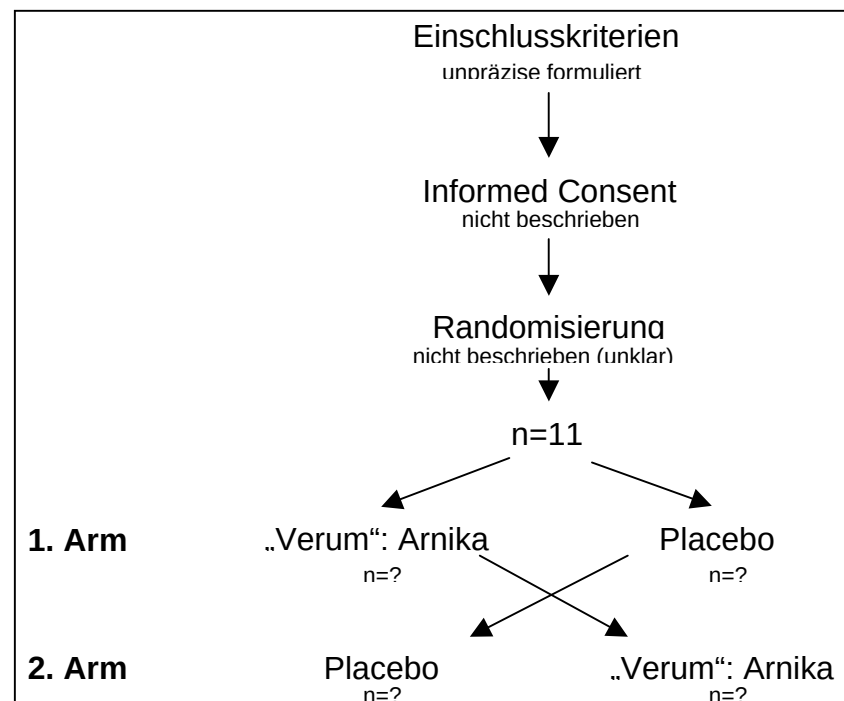


Abbildung 1: Flussdiagramm zur Studie von Campbell A (1976): Two pilot controlled trials of Arnica montana

2.7 Statistische Angaben

Als statistische Maßzahl wurde der Median und als Streuungsmaß das zugehörige erste und dritte Quartil (Q1 und Q3) bestimmt. Angegeben wurde das Streuungsmaß als interquartiler Range „IQR (Q1;Q3)“, welcher den Bereich zwischen dem ersten und dritten Quartil beschreibt und in dem sich 50% der Daten bewegen. Bei größeren Datenbereichen oder Datenbereichen mit großer Streuungsbreite wurden zusätzlich zum interquartilen Range Minimum und Maximum angegeben. Prozentangaben wurden grundsätzlich auf die ganze Zahl gerundet.

3 Ergebnisse

3.1 Studienselektion

3.1.1 Durchführung des Ausschlussverfahrens

Unter dem Suchbegriff „homeopath*“ in der Cochrane- Datenbank (www.cochrane.de) wurden sechs Protocols, 29 Reviews und drei Meta-Analysen angezeigt.

Aus diesem Kollektiv wurden die 29 Reviews und drei Meta-Analysen weiter bearbeitet, die sechs als „Protocols“ gelisteten Arbeiten entfielen entsprechend dem Ausschlusskriterium für die Publikationsform.

Zu allen in der Cochrane- Datenbank gelisteten Übersichtsarbeiten ist eine Kurzbeschreibung mit Angabe der verwendeten Intervention(en) vorhanden. Bei den 29 Reviews und drei Metaanalysen diente diese Kurzbeschreibung als Richtlinie zur Überprüfung der verwendeten Präparate hinsichtlich ihrer homöopathischen Grundlage. Die Literaturangaben zu diesen Übersichtsarbeiten sind im Anhang zu finden. Nach durchgeführter Überprüfung mussten 14 Reviews ausgeschlossen werden, bei welchen keine homöopathische Intervention gemäß unserer Einschlusskriterien beschrieben war. Dabei ist nicht völlig auszuschließen, dass möglicherweise in den Übersichtsarbeiten Einzelstudien enthalten waren, welche ein homöopathisches Präparat in unserem Sinne verwendeten!

Es verblieb eine vorläufige Auswahl von 15 Reviews mit insgesamt 147 Studien sowie drei Meta-Analysen mit insgesamt 111 Studien.

Bei diesen 258 Studien wurde als Basis zur erneuten Durchsicht die jeweilige Einzelstudienbeschreibung herangezogen. Die Überprüfung auf die Verwendung von homöopathischen Arzneimitteln im Sinne von Hahnemann (11) in Zusammenarbeit mit der Heilpraktikerin, Frau Cornelia Mantel- Arnold, ergab folgendes

Ergebnis:

Von den drei Meta-Analysen und 15 Reviews enthielten neun komplette Reviews mit insgesamt 92 Einzelstudien keine einzige Studie, welche ein homöopathisches Präparat nach Hahnemann entsprechend unserer Kriterien benutzte. Dabei erwies sich als Ausschlussgrund die Verwendung von:

- 1) Kombinationspräparaten wie z.B. „Esberitox®“ (Kombination unter anderem aus Echinacea, Baptisia tinctoria, Thuja occidentalis)
- 2) nicht potenzierten Phytotherapeutika wie z.B. „Jarsin®“ (Johanneskraut-Extrakt)
- 3) potenzierten Allergenen, Bakterien und Viren
- 4) diversen Vitaminen (B6) und Mineralien (Magnesium, Calcium)
- 5) Ölen, wie z.B. Soja- oder Avocadoöl.
- 6) prophylaktisch verabreichten Substanzen

Bei den restlichen sechs Reviews und drei Meta-Analysen mit insgesamt 166 Studien kamen 33 Studien doppelt und zwei Studien dreifach vor. Demnach verblieben 129 Studien.

Von diesen entfielen weitere 64 Studien, welche nicht mit den Ein- und Ausschlusskriterien bezüglich des verwendeten Arzneimittels, der Publikationsform oder der Sprache (englisch/deutsch) konform waren. Dies betraf auch alle sechs Studien einer der drei Meta-Analysen, welche in französisch publiziert waren und zudem in drei Fällen als Dissertationen vorlagen.

Von den übrigen 65 Studien konnten sieben Studien nicht bestellt werden, da sie aufgrund fehlender Publikationsangaben nicht auffindbar waren.

Es lagen schließlich 58 Studien im Original vor, von denen erneut insgesamt 14 Studien ausgeschlossen werden mussten. Sieben Studien waren nicht auswertbar: bei einer Studien fehlte der gesamte Ergebnisteil, sechs Studien waren Teil einer Publikation, welche über laufende Projekte berichtete. Zwei Studien waren Vorläuferstudien und in ihrem Aufbau einer Folgestudie analog, welche ausgewertet wurde. Bei vier Studien stellte sich heraus, dass sie einem Dissertationsbericht beziehungsweise einem Editorial (jeweils zwei Studien) entsprachen. Eine weitere Studie wurde ausgeschlossen, da das verwendete Arzneimittel als Extrakt vorlag (Asa foetida und Nux vomica in 0,1% alkoholischer Lösung) und somit nicht dem Kriterium eines homöopathischen Arzneimittels entsprach.

Von den verbliebenen 44 Studien wurden in willkürlicher Reihenfolge 40 Studien ausgewertet. Nach der Auswertung dieser Studien war ein stabiles Ergebnis erreicht. Auf die Auswertung der letzten vier Arbeiten wurde verzichtet, nachdem daraus kein Informationsverlust, aber eine deutliche Erleichterung des mit jeder

Auswertung wachsenden inneren Widerstandes gegen eine sehr sorgfältige Analyse sich ständig wiederholender Mängel resultierte.

Tabelle 2: Summation des Ausschlussverfahrens

	Protokolle	Reviews	Meta-Analysen	Einzelstudien
Suchergebnis "homeopath**"	6	29	3	nicht erfasst
Ausschlussgründe 1.Durchsicht:				
Publikationsform Kein homöopathisches Arzneimittel	6	14		
Verbleibend	0	15	3	258
Ausschlussgründe 2.Durchsicht:				
Kein homöopathisches Arzneimittel		9		92
Verbleibend	0	6	3	166
Doppelt /dreifach vorhandene Studien				37
Verbleibend	0	6	3	129
Ausschlussgründe 3.Durchsicht:				
Kein homöopathisches Arzneimittel				25
Sprache (weder Englisch noch Deutsch)				9
Publikationsform				8
Kein homöopathisches Arzneimittel und Sprache				10
Kein homöopathisches Arzneimittel und Publikation				0
Kein homöopathisches Arzneimittel und Sprache und Publikation				3
Sprache und Publikation				9
Gesamt				64
Verbleibend	0	6	2	65
Davon nicht auffindbar/bestellbar				7
Verbleibend				58
Ausschlussgründe bei Durchsicht der bestellten Studien:				
Nicht auswertbar				7
Folgestudie, Vorläuferstudie wurde ausgewertet				2
Publikation: Dissertation, Editorial				4
Arzneimittel nicht homöopathisch				1
Summe verbleibend	0	5	2	44
Ausgewertet				40

3.1.2 Ergebnis des Ausschlussverfahrens

Die Literaturangaben zu den 29 Reviews und drei Meta-Analysen, welche die Grundlage für die Selektion darstellten, sowie zu den vier nicht bearbeiteten Einzelstudien sind im Anhang zu finden.

In Tabelle drei sind die ausgewerteten 40 Einzelstudien aufgelistet. Die Reihenfolge ist zufällig und entspricht dem Eingang in die Bearbeitung.

Tabelle 3: Bearbeitete Studien

Studien-Nr.	Literaturangabe
1	Campbell A: Two pilot controlled trials of Arnica montana. Br Homeopath J 65: 154-158 (1976)
2	Campbell A: Two pilot controlled trials of Arnica montana. Br Homeopath J 65: 154-158 (1976)
3	Gibson RG, Gibson S, MacNeill AD, Gray GH, Dick WC, Watson Buchanan W: Salicylates and homeopathy in rheumatoid arthritis: preliminary observations. Br J Clin Pharmacol 6: 391- 395 (1978)
4	Hildebrandt G, Eltze C: Über die Wirksamkeit verschiedener Potenzen von Arnica beim experimentell erzeugten Muskelkater. Erfahrungsheilkunde 7: 430-435 (1984)
5	Kainz JT, Kozel G, Haidvogel M, Smolle J: Homeopathic versus placebo therapy of children with warts on the hands: a randomized, double-blind clinical trial. Dermatology 193: 318-320 (1996)
6	Walach H, Haeusler W, Lowes T, Mussbach D, Schamell U, Springer W, Stritzl G, Gaus W, Haag G: Classical homeopathic treatment of chronic headache. Cephalalgia 17: 119-126 (1997)
7	Kaziro GSN: Metronidazole (Flagyl) and Arnica montana in the prevention of postsurgical complications: a comparative placebo controlled clinical trial. Br J Oral Maxillofac Surg 22: 42-49 (1984)
8	Kennedy CO: A controlled trial. Br Homeopath J 60: 120-127 (1971)
9	Bignami M, Saruggia M, Sansonetti G: Homeopathic treatment of anal fissures using Nitricum acidum. The Berlin Journal on Research in Homeopathy* 1: 286-287 (1991)
10	Gibson J, Haslam Y, Lurneson L, Newmann P, Pitt R, Robins M: Double blind trial of arnica in acute trauma patients. Homeopathy 41: 54-55 (1991)
11	Fisher P, Greenwood A, Huskisson EC, Turner P, Belon P: Effect of homeopathic treatment on fibrositis (primary fibromyalgia). Br Med J 299: 365-366 (1989)
12	Andrade L, Ferraz MB, Atra E, Castro A, Silva MSM: A randomized controlled trial to evaluate the effectiveness of homeopathy in rheumatoid arthritis. Scand J Rheumatol 20: 204-208 (1991)
13	Beer AM, Heiliger F: Randomized, double blind trial of Caulophyllum D4 for induction of labour after premature rupture of membranes at term. Geburtshilfe Frauenheilkd 59: 431-435 (1999)
14	Solanki M, Ghandi PM: Is homeopathy only a placebo therapy? Homeopathic Heritage 20: 707-713 (1995)
15	Brigo B, Serpelloni G: Homeopathic treatment of migranes: A randomized double blind study of sixty cases (homeopathic remedy versus placebo). The Berlin Journal on Research in Homeopathy* 1: 98-106 (1991)
16	Whitmarsh TE, Coleston-Shields DM, Steiner TJ: Double-blind randomized placebo-controlled study of homeopathic prophylaxis of migraine. Cephalgia 17: 600-604 (1997)
17	Van Erp VMA, Brands M: Homeopathic treatments of malaria in Ghana. Br Homeopath J 85: 66-70 (1996)
18	Jansen G, Veer A, Hagenaaers J, Kuy A: Lessons learnt from an unsuccessful clinical trial of homeopathy. Results of a small scale double blind trial in proctocolitis. Br Homeopath J 81: 132-138 (1992)
19	Lamont J: Homeopathic treatment of attentiondeficit hyperactivity disorder. Br Homeopath J 86: 196-200 (1997)

Fortsetzung S. 28

20	Jacobs J, Jimenez LM, Gloyds S, Gale JL, Crothers D: Treatment of acute childhood diarrhea with homeopathic medicine: A randomized clinical trial in Nicaragua. Pediatrics 93: 719-725 (1994)
21	Kuzeff RM: Homeopathy, sensation of well-being and CD4 levels: A placebo-controlled randomized trial. Complement Ther Med 6: 4-9 (1998)
22	Mössinger P: Zur Behandlung der Otitis media mit Pulsatilla. Kinderarzt 16: 581-82 (1985)
23	Pinsent RJFM, Baker GPI, Ives G, Davey RW, Jonas S: Does arnica reduce pain and bleeding after dental extraction? Midland Homeopathy Research Group Newsletter 11:71-72 (1984)
24	Rahlfs VW, Mössinger P: Asa foetida bei Colon irritabile- Doppelblindversuch. Dtsch Med Wochenschr 104: 140-143 (1979)
25	Wiesenauer M, Gaus W: Double-blind trial comparing the effectiveness of the homeopathic preparation Galphimia potensation D6, Galphimia dilution 10-6 and placebo on pollinosis. Arzneimittelforschung 35: 1745-1747 (1985)
26	Wiesenauer M, Gaus W, Häussler S: Behandlung der Pollinosis mit Galphimie glauca: Eine Doppelblindstudie unter Praxisbedingungen. Allergologie 13: 359-363 (1990)
27	Wiesenauer M, Häussler S, Gaus W: Pollinosis- Therapie mit Galphimia glauca. Fortschritt Medizin 101: 811-814 (1983)
28	Shipley M, Berry H, Booster G, Jenkins M, Clover A, Williams I: Controlled trial of homeopathic treatment of osteoarthritis. Lancet i: 97-98 (1983)
29	Ustianowski PA: A clinical trial of Staphysagria in postcoital cystitis. Br Homeopath J 63: 276-277 (1974)
30	Leaman AM, Gorman D: Cantharis in the early treatment of minor burns. Arch Emerg Med 6: 259-261 (1989)
31	Ritter H: Ein homöopathischer doppelter Blindversuch und seine Problematik. Hippokrates 12: 472-476 (1966)
32	Labrecque M, Audet D, Latulippe LG, Drouin J: Homeopathic treatment of plantar warts. Canadian Medical Assoc Journal 146: 1749-1753 (1992)
33	Savage RH, Roe PF: A double blind trial to assess the benefit of Arnica montana in acute stroke illness. Br Homeopath J 66: 207-220 (1977)
34	Hildebrandt G, Eltze C: Über die Wirksamkeit einer Behandlung des Muskelkaters mit Rhus toxicodendron D4. Zwei Beiträge zur Pharmakologie adaptiver Prozesse. Wissenschaftliches Archiv der Hufelandgesellschaft für Gesamtmedizin 1: 3-18 (1983)
35	Chapman EH, Angelica J, Spitalny G, Strauss M: Results of a study of the homeopathic treatment of PMS. Journal of the American Inst Hom 87: 14-21 (1994)
36	Gaucher C, Jeulin D, Peycru P, Amengual C: A double blind randomized placebo controlled study of cholera treatment with highly diluted and succussed solutions. Br Homeopath J 83: 132-134 (1994)
37	Gibson RG, Gibson S, MacNeill AD, Watson Buchanan W: Homeopathic therapy in rheumatoid arthritis: Evaluation by double-blind clinical therapeutical trial. Br J Clin Pharmacol 9: 453-459 (1980)
38	De Lange de Klerk E, Blommers J, Kuik DJ, Bezemer PD, Feenstra L: Effects of homeopathic medicines on daily burden of symptoms in children with recurrent upper respiratory tract infections. Br Med J 309: 1329-1332 (1994)
39	Wiesenauer M, Lüdtke R: The treatment of pollinosis with Galphimia glauca D4-a randomized placebo-controlled double-blind clinical trial. Phytomedicine 2: 3-6 (1995)
40	Hofmeyer GJ, Piccioni V, Blauhof P: Postpartum homeopathic Arnica montana: a potency-finding pilot-study. Br J Clin Pract 44: 619-621 (1990)

*Anmerkung zu Nr.9 und 15: The Berlin Journal on Research in Homeopathy: in "pubmed" nicht gelistet

3.2 Anzahl der teilnehmenden Institutionen im Validitätsvergleich

Verteilung der Anzahl der teilnehmenden Institutionen bei n=40 Studien

Mono-/bizentrisch (1-2 Institutionen):	30 von 40 Studien
Drei oder mehr Institutionen:	Acht von 40 Studien
Anzahl der Institutionen nicht beschrieben:	Zwei von 40 Studien

Mehr als 2/3 der Studien wurde also mit maximal zwei beteiligten Institutionen durchgeführt. Tabelle 4 zeigt die Verteilung der internen Validität in Bezug auf die Anzahl der beteiligten Institutionen pro Studie.

Tabelle 4: Interne Validität in Abhängigkeit von der Zahl der beteiligten Institutionen pro Studie

Studie	Interne Validität (in Noten von 1-6) der Studien mit		
	1-2 Institutionen	3 oder mehr Institutionen	nicht beschriebener Institutionsanzahl
Studie 1	6,0		
Studie 2	6,0		
Studie 3	4,0		
Studie 4	5,0		
Studie 5	4,0		
Studie 6	1,0		
Studie 7	5,0		
Studie 8	6,0		
Studie 9			4,0
Studie 10	6,0		
Studie 11	6,0		
Studie 12	3,0		
Studie 13	3,0		
Studie 14	6,0		
Studie 15	3,0		
Studie 16	4,0		
Studie 17	4,0		
Studie 18	5,0		
Studie 19	6,0		
Studie 20	2,0		
Studie 21	5,0		
Studie 22		6,0	
Studie 23		5,0	
Studie 24		4,0	
Studie 25		4,0	
Studie 26		4,0	
Studie 27		4,0	

Studie 28		5,0	
Studie 29	6,0		
Studie 30	5,0		
Studie 31	5,0		
Studie 32	4,0		
Studie 33	5,0		
Studie 34	6,0		
Studie 35	4,0		
Studie 36			5,0
Studie 37	4,0		
Studie 38	2,0		
Studie 39		4,0	
Studie 40	3,0		
Median	5,0	4,0	4,5
IQR	(4,0;6,0)	(4,0;5,0)	-

Aus Tabelle 4 ist ersichtlich, dass die interne Validität bei multizentrische Studien im Vergleich zur mono- bzw. bizenrischen Studien im Median um eine Schulnote besser war. Dabei war die Anzahl der multizentrischen Studien um das nahezu vierfache kleiner als die Anzahl der mono-/bizenrischen Studien.

Die zwei Studien ohne Angabe zur Anzahl der beteiligten Institutionen erhielten einmal die Note 4,0, einmal die Note 5,0 bezüglich ihrer Validität.

3.3 Herstellung der Arzneimittel

Bei der Hälfte der Studien erfolgte die Herstellung beider Arzneimittel extern, also nicht durch die Apotheke der rekrutierenden Institution(en). Ein Drittel der Studien n=13 (entsprechend 32,5%) machten keine Angaben zu der Herkunft der Arzneimittel. Bei fünf Studien erfolgte die Herstellung der Arzneimittelgruppen an jeweils unterschiedlichen Lokalisationen, also nicht einheitlich extern oder intern.

Tabelle 5: Übersicht über den Herstellungsort (intern vs. extern) der Studienpräparate bei n=40 Studien

Herstellung	Experimentelles Arzneimittel	Kontrollarzneimittel (meistens Placebo)	Beide Arzneimittelgruppen
Interne Herstellung	2	3	2
Externe Herstellung	25	20	20
Herstellung nicht beschrieben	13	17	13
Herstellung nicht einheitlich			5

3.4 Informed Consent

Der Informed Consent wurde eingeholt bei 16 von 40 Studien, davon war bei sechs Studien der Inhalt klar beschrieben, bei zehn Studien war der Inhalt nicht angegeben oder nicht klar nachvollziehbar. Bei 23 von 40 Studien, also mehr als der Hälfte, gab es keinerlei Angaben bezüglich eines Informed Consent. Bei einer Studie war aus dem Text eindeutig zu entnehmen, dass kein Informed Consent stattgefunden hat: „Die Kranken erfuhren überhaupt nicht, dass ein Versuch mit ihnen vorgenommen wurde, ebenso wenig, was sie empfangen hatten.“ (Ritter 1966).

3.5 Einschlusskriterien

Die Einschlusskriterien wurden für 32 der 40 Studien angegeben. Davon waren sie in 13 Fällen (32,5%) präzise formuliert, in 19 Fällen (47,5%) waren sie unzureichend beschrieben. In acht Studien (20%) wurden keine Einschlusskriterien angegeben.

Beispiele:

1) Undifferenzierte Einschlusskriterien:

- Studie: „Behandlung von Heuschnupfen mit Galphimia glauca D4“ (Wiesenauer und Lüdtke 1995):
Einschlusskriterien: „Patienten, welche seit mindestens zwei Jahren in Behandlung wegen Heuschnupfen sind“
- Studie: „Postpartale Gabe von Arnica montana“ (Hofmeyer et al. 1990):
Einschlusskriterien: „Frauen mit Episiotomie oder Dammriss nach Geburt“

2) Differenzierte Einschlusskriterien:

- Studie: „Homöopathische Prophylaxe von Migräne“ (Whitmarsh et al. 1997)
Einschlusskriterien: „Migräne mit oder ohne Aura, diagnostiziert durch IHS (International Headache Society) Kriterien, erkennbare Attacken seit mindestens zwei Jahren, Häufigkeit der Attacken

(medizinisch dokumentiert) zwischen zwei bis acht pro Monat in jedem der letzten drei Monate, Alter 18-60 Jahre, vorliegender „Informed Consent“.

3.6 Rekrutierung und Randomisierung von Patienten

Anzahl der rekrutierten Patienten

Bei neun von 40 Studien war die Anzahl der rekrutierten Patienten nicht angegeben. In den restlichen 31 Studien wurden insgesamt 3539 Patienten rekrutiert. Der Median betrug 69, der interquartile Range IQR (34;139). Das heißt, bei 50% der Studien bewegte sich die Anzahl der rekrutierten Patienten zwischen $n=34$ und $n=139$ bei einem Minimum von elf Patienten und einem Maximum von 853 Patienten.

Anzahl der randomisierten Patienten

Eine Randomisierung wurde in 31 der 40 Studien durchgeführt. In den 31 Studien wurden insgesamt 2363 Patienten randomisiert, davon war bei drei Studien die angegebene Anzahl der randomisierten Patienten nicht eindeutig nachvollziehbar (Savage u. Roe 1977 mit $n=40$ Patienten, Gaucher et al. 1994 mit $n=80$ Patienten und Gibson et al. 1991 mit $n=20$ Patienten). Zusätzlich muss bei Gibsen et al. die Durchführung der Randomisierung angezweifelt werden, da fünf Krankenschwestern vier mit A,B,C und D beschriftete Arzneimittelflaschen mit je zweimal Verum und Placebo in Rotation verwendet haben wollen. Die tatsächliche Einhaltung dieses Randomisierungsplanes erscheint wenig wahrscheinlich.

Der Median der randomisierten Patienten betrug 74, das 1.Quartil 36, das 2.Quartil 121. Damit bewegt sich bei 50% der Studien die Anzahl der randomisierten Patienten im Bereich des interquartilen Range IQR von $n=36$ bis $n=121$ Patienten. Das Minimum lag bei zehn, das Maximum bei 243 Patienten.

Bei vier Studien wurde eine Randomisierung beschrieben, die Anzahl der Patienten jedoch nicht angegeben (Kennedy 1971, Fisher et al. 1989, Wiesenauer u. Gaus 1985, Shipley et al. 1983).

Keine Randomisierung fand bei neun der 40 Studien statt.

Tabelle 6: Zahl der rekrutierten und randomisierten Patienten in den 40 ausgewerteten Studien

Die Studien wurden jeweils aufsteigend nach der Zahl der berücksichtigten Patienten geordnet.

Die Mediane und der interquartile Range sind **fett markiert**.

Abkürzungen:

n.a.: nicht angegeben *k.R.*: keine Randomisierung *: Angabe unklar

Studie	Anzahl rekrutierter Patienten	Studie	Anzahl randomisierter Patienten
Studie 7	n.a.	Studie 1	k.R.
Studie 8	n.a.	Studie 2	k.R.
Studie 12	n.a.	Studie 3	k.R.
Studie 13	n.a.	Studie 4	k.R.
Studie 21	n.a.	Studie 8	n.a.
Studie 27	n.a.	Studie 11	n.a.
Studie 30	n.a.	Studie 19	k.R.
Studie 39	n.a.	Studie 22	k.R.
Studie 40	n.a.	Studie 25	n.a.
Studie 1	11	Studie 28	n.a.
Studie 2	15	Studie 29	k.R.
Studie 10	20	Studie 34	k.R.
Studie 18	20	Studie 37	k.R.
Studie 35	21	Studie 35	10
Studie 11	30	Studie 10	20*
Studie 9	31	Studie 18	20
Studie 14	34	Studie 9	31
Studie 28	36	Studie 14	34
Studie 33	40	Studie 30	34
Studie 4	42	Studie 21	36
Studie 19	43	Studie 13	40
Studie 22	44	Studie 33	40*
Studie 37	46	Studie 12	44
Studie 15	60	Studie 15	60
Studie 16	69	Studie 16	63
Studie 5	70	Studie 5	67
Studie 34	76	Studie 17	74
Studie 36	80	Studie 36	80*
Studie 17	92	Studie 20	87
Studie 20	92	Studie 23	100
Studie 23	100	Studie 6	101
Studie 24	119	Studie 7	118
Studie 6	139	Studie 24	119
Studie 31	147	Studie 27	121
Studie 3	195	Studie 31	147
Studie 29	200	Studie 40	161
Studie 25	213	Studie 39	164
Studie 26	243	Studie 38	175
Studie 38	358	Studie 32	174
Studie 32	853	Studie 26	243

Fortsetzung S. 34

Patientenanzahl gesamt	3539		2363
Median	69		74
IQR	(34;139)		(36;121)

3.6.1 Korrelation zwischen der Größe der Patientenpopulation und der internen Validität

Die nachfolgende Tabelle 7 und Abbildung 2 zeigen die Gegenüberstellung der Größe der Probandenpopulation einer Studie und ihrer jeweiligen internen Validität. Die Zahl der rekrutierten und der randomisierten Patienten wurde dazu jeweils vier Größenkategorien (<31,31-100,101-300,>300) zugeteilt. Zur Berechnung des Medians und des interquartilen Range der Validität wurde jede Studie nur einmal berücksichtigt.

Beispiel: Wurden in einer Studie 58 Patienten rekrutiert und davon 32 Patienten randomisiert, so wurde die Note der internen Validität in der Kategorie 31-100 für die betreffende Studie nur einmal zur Berechnung der Werte herangezogen.

Die interne Validität verbesserte sich mit zunehmender Größe der Studienpopulation im Median um insgesamt 2,5 Schulnoten (Gesamtnotenbereich 1,0-6,0) bei einer Studienpopulation von zehn bis 853 Patienten. In der Kategorie der Patientenpopulation > 300 fanden sich allerdings nur zwei Studien.

Bei den neun Studien, welche keine Randomisierung angaben, betrug die interne Validität im Median 6,0 mit IQR (5,0;6,0).

Tabelle 7: Größe der Studienpopulation und zugehörige interne Validität (Angabe von Median und interquartilem Range)

*Anmerkung: * : Patientenanzahl unklar*

	Patientenanzahl						
Studien mit n	< 31	31-100	101-300	>300	Pat.-Anzahl nicht angegeben	Gesamt	Randomisation nicht durchgeführt
Anzahl der Studien mit rekrutierten Patienten	6	16	7	2	9	40	-
Anzahl der Studien mit randomisierten Patienten	3 (davon 1*)	14 (davon 2*)	10	0	4	31*	9
Interne Validität (Schulnoten 1,0-6,0)							
Median	5,5	5,0	4,0	3,0	4,5		6,0
IQR	(5,0;6,0)	(4,0;5,0)	(4,0;4,0)		(3,5;5,0)		(5,0;6,0)

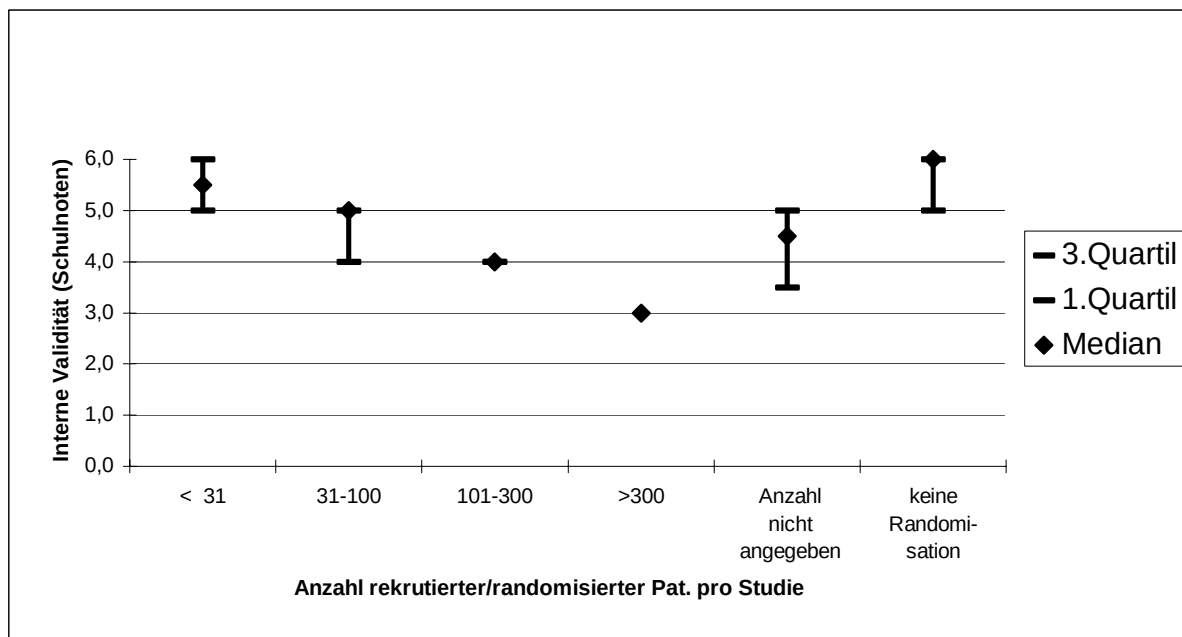


Abbildung 2: Interne Validität in Abhängigkeit von der Patientenanzahl pro Studie

3.6.2 Ergebnisse zur Differenz zwischen rekrutierten und randomisierten Patienten

In nur 19 von 40 Studien waren die Angaben zu beiden Patientenkollektiven vollständig.

- 1) Davon entsprach bei elf dieser 19 Studien die Anzahl der rekrutierten Patienten genau der Anzahl der randomisierten Patienten.
- 2) Bei acht Studien gab es eine Differenz im Verhältnis der rekrutierten zu randomisierten Patienten, einen Überblick zeigt Tabelle 8.1. Im Median betrug der Anteil der nicht- randomisierten Patienten am Kollektiv der rekrutierten Patienten 24%, IQR (54%;7%).

Tabelle 8: Verhältnis der rekrutierten zu randomisierten Patienten, n=8 Studien

	Rekrutierte Patienten	Randomisierte Patienten	Nicht-randomisierte Patienten	Anteil der nicht-randomisierten an rekrutierten Patienten in Prozent
Studie 32	853	174	679	80%
Studie 35	21	10	11	52%
Studie 38	363	175	188	52%
Studie 6	139	101	38	27%
Studie 17	92	74	18	20%
Studie 16	69	63	6	9%
Studie 20	92	87	5	5%
Studie 5	70	67	3	4%
Median				24%
IQR				(54%;7%)

Bei den restlichen 21 Studien fehlte zumindest einer der folgenden Datenbereiche:

- 1) Acht Studien (siehe auch 3.6.1): keine Angaben zur Anzahl der rekrutierten Patienten vorhanden, davon wurde in allen acht Fällen die Anzahl der randomisierten Patienten beschrieben.
- 2) Neun Studien: kein Hinweis auf eine Randomisierung feststellbar sowie in
- 3) Drei Studien: diesbezüglich fehlende Zahlenangabe.
- 4) Diese zwölf Studien enthielten alle eine Angabe zur Anzahl der rekrutierten Patienten.
- 5) Eine Studie: weder die Anzahl der rekrutierten, noch die Anzahl der randomisierten Patienten war bei angeblich stattgefundener Randomisierung angegeben (Kennedy 1971).

3.6.3 Begründung der Differenz zwischen rekrutierter und randomisierter Patientenpopulation

Insgesamt waren bei neun der 19 Studien Gründe für die Differenz zwischen der Anzahl der primär akquirierten und der Zahl der randomisierten Patienten beschrieben.

Valide Begründung

In sechs Fällen wurde die Differenz der Patientenzahlen mit der angegebenen Begründung plausibel erklärt und stellte keine Einschränkung der Validität dar.

Beispiel:

Whitmarsh et al. 1997: 69 rekrutierte Patienten nahmen für einen Monat an einem „Testlauf“ teil. Während dieser Zeit erhielten alle Patienten das Placebo und mussten ein Tagebuch über ihre Erkrankung führen. Gründe für den Ausschluss von sechs Patienten nach diesem Zeitraum waren:

- a) Häufigkeit und der Charakter der Migräne entsprach laut Tagebuch nicht den Einschlusskriterien
- b) Nicht-Ausfüllen des Tagebuches
- c) Nicht-Einnahme der Medikation (Placebo)
- d) Zusätzliche Erkrankung
- e) „withdrawal of Informed Consent“

Es verblieben 63 Patienten, welche in eine Verum und eine Placebo-Gruppe randomisiert wurden.

Validitätseinschränkende Begründung

In zwei Fällen waren die Gründe unzulänglich und wurden als validitätsmindernd gewertet:

- 1) Bei der Studie „Homöopathische Behandlung von Fußwarzen“ (Labrecque et al. 1992) ist die Differenz von 853 rekrutierten Patienten zu 174 randomisierten Patienten selbst mit den differenziert angegebenen Erläuterungen zu „Einschlusskriterien nicht erfüllt“ bei 596 Pat. sowie „verweigerte Teilnahme“ bei 83 Pat. schwer nachvollziehbar.
- 2) In einer weiteren Studie (Kuzeff 1998) wurde bei freier Auswahl der homöopathischen Präparate „bei 50 % der Patienten nach vier Sitzungen kein geeignetes Arzneimittel gefunden“. Es fehlte eine explizite numerische Angabe bezüglich der Zahl der rekrutierten Patienten.

Tabelle 9: Differenz zwischen der Anzahl rekrutierter und randomisierter Patienten sowie Vorliegen von Begründungen

Differenz in der Anzahl rekrutierter zu randomisierter Patienten	0	1-10	10-20	20-50	50-100	>100	Differenz nicht berechenbar	Gesamt
Anzahl der betreffenden Studien	11	3	2	1	-	2	21	40
Gründe angegeben	-	3	2	1	-	2	1	9
Gründe beeinflussen die Validität	-	-	1	-	-	1	1	3
Gründe beeinflussen die Validität nicht	-	3	1	1	-	1	-	6

3.7 Ergebnisse zur Durchführung der Randomisierung

3.7.1 Ausführung der Randomisierung

Zur Gewährleistung einer möglichst unabhängigen Randomisierung ist die Ausführung durch eine externe Organisation wünschenswert. Dies wurde in den 40 Studien überprüft mit folgendem Ergebnis:

Die Randomisierung erfolgte durch

- 1) eine externe Institution: in acht Studien.
- 2) die rekrutierende/die Studien durchführende Institution: in sechs Studien.

Keine Angaben waren vorhanden bei 16 Studien.

3.7.2 Geheimhaltung des Randomisierungsplans

Von den 31 der 40 Studien, welche eine Randomisierung durchführten, wurde bei zwölf Studien der Randomisierungsplan so ausführlich beschrieben, dass von einer Geheimhaltung ausgegangen werden kann oder die Wahrung der Geheimhaltung wurde ausdrücklich vermerkt.

In 17 Fällen gab es keine Beschreibung des Randomisierungsplans, welche die Beantwortung der Frage ermöglicht hätte.

In zwei Fällen wurde der Randomisierungscode gebrochen:

- 1) nach der Randomisierung wurden Patienten ausgeschlossen um „homogenere Gruppen zu erhalten“ (Savage 1977).
- 2) „on perhaps two occasions, the order of Randomisierung became known to the practitioner“ (Kuzeff 1998).

3.8 Auswertung der Studienpatienten

Anzahl der ausgewerteten Patienten

Bei 31 der 40 Studien war die Anzahl der ausgewerteten Patienten angegeben, die Gesamtzahl belief sich auf n=2236 Patienten.

Von den 31 Studien waren bei 26 die Werte eindeutig nachvollziehbar, bei fünf Studien waren die Zahlenangaben unklar. Für die oben errechnete Gesamtzahl und die weitere Bearbeitung wurde bei differierenden Angaben die höchstgenannte Patientenzahl verwendet.

Unklarheiten ergaben sich bei folgenden Autoren:

- 1) Campbell 1976 berichtet über 15 ausgewertete Patienten. Es wurden bei 15 rekrutierten Patienten laut Angabe zwei Patienten ausgeschlossen, in der Auswertung fanden sich einmal 13, einmal 15 Patienten.
- 2) Walach et al. 1997 berichten über 98 ausgewertete Patienten. Ein Zahlenwert lag im Ergebnisteil nicht vor. Man kann nur aufgrund der Aussage, dass sechs Patienten ausgeschlossen, jedoch laut „Intent to Treat“ evaluiert wurden folgern, dass alle 98 randomisierte Patienten ausgewertet wurden (32).
- 3) Jansen et al. 1992 berichten über 18 ausgewertete Patienten. Im Text sind zwei „drop outs at own request“ beschrieben, diese waren in der Auswertung nicht berücksichtigt. Ferner gab es sechs „drop outs for medical reasons“. Diese sechs Ausschlusskandidaten wurden in die tabellarische Präsentation der Ergebnisse einbezogen (n=18 Patienten), in die Textbeschreibung des Outcomes jedoch nicht (im Text Angabe von n= zwölf ausgewerteten Patienten).
- 4) Gaucher et al. 1994 berichten über 44 ausgewertete Patienten: Laut Studie verblieben von 80 randomisierten Patienten (diese Angabe wird durch keine weiteren Zahlen in der Studie gestützt) nur 44 „valid forms“ für die Evaluation. Es lag keine Erklärung vor, was unter einem „gültigen Formular“ verstanden wurde. In der Evaluation fehlte jegliche weitere Zahlenangabe bis auf einmalig n=42 Patienten für ein Teilergebnis.
- 5) Savage u. Roe 1977 berichten über 40 ausgewertete Patienten: Nach Randomisierung (n=40 Patienten) wurden zehn Patienten ausgeschlossen, um „homogenere Gruppen“ zu erhalten. Im Ergebnisteil wurden teilweise 40, teilweise 30 Patienten ausgewertet. Dieses Vorgehen bedeutet zusätzlich einen Verstoß gegen die Geheimhaltung des Randomisierungsplanes (siehe dazu 3.7.2.) und das „Intent to Treat“-Prinzip.

Bei neun Studien gab es keinen Hinweis darauf, wie viele Patienten ausgewertet wurden.

Prozentual belief sich der Anteil der ausgewerteten an randomisierten bzw. rekrutierten Patienten im Median auf 92%, IQR (77%-100%), das Minimum betrug 24%, das Maximum 100%.

Anzahl der ausgeschlossenen Patienten

Insgesamt wurden 524 Patienten ausgeschlossen bei einer Gesamtzahl von 2236 ausgewerteten und 3174 randomisierten beziehungsweise rekrutierten Patienten. *(Bei fehlender Zahlenangabe zur Randomisierung/nicht stattgefundener Randomisierung wurde die Anzahl der rekrutierten Patienten berücksichtigt!).*

Prozentual betrug der Anteil der ausgeschlossenen Patienten am Anteil der ausgewerteten Patienten im Median 16%, IQR (6%:34%) mit einem Minimum von 1% ausgeschlossenen an ausgewertete Patienten und einem Maximum von 324%.

In fünf Studien wurde kein Patient ausgeschlossen, das heißt, alle randomisierten bzw. aufgenommenen Probanden wurden ausgewertet.

Bei zehn der 40 Studien war die Anzahl der ausgeschlossenen Patienten nicht angegeben, bei 25 Studien lag eine Zahlenangabe vor.

Dabei wiesen drei der 25 Studien eine Diskrepanz zwischen der beschriebenen Anzahl der ausgeschlossenen Patienten und der Anzahl der ausgewerteten Patienten auf. Die Angaben waren letztendlich nicht nachvollziehbar:

- 1) Studie 18 (Campbell 1976): Angeblich wurden zwei Patienten ausgeschlossen; in der Auswertung bei 15 aufgenommenen Patienten teilweise alle, teilweise 13 Patienten ausgewertet (siehe 3.8. Punkt 1).
- 2) Studie 23 (Pinsent et al. 1984): Bei 100 Patienten wurden 59 ausgewertet. Es wurden jedoch 24 Patienten ausgeschlossen. Wurde möglicherweise ein Teil der ausgeschlossenen Patienten doch ausgewertet?
- 3) Studie 33 (Savage und Roe 1977): Nachträglicher Ausschluss von zehn Patienten, diese waren teilweise in den Ergebnissen nicht beschrieben, teilweise jedoch berücksichtigt (siehe 3.8. Punkt 5).

3.8.1 Ausschlussgründe und ihre Validität

Die zehn Studien mit komplett fehlender Information zur Anzahl der ausgeschlossenen Patienten wurden hinsichtlich der Ausschlussgründe nicht evaluiert. Die Ausschlussgründe sind irrelevant, wenn ohnehin nicht nachvollziehbar ist, wie viele Patientendaten verworfen wurden. Daher wurde dieser Verstoß in Bezug auf die Punktevergabe der internen Validität gleichbedeutend gewertet wie eine vorhandene Zahlenangabe mit nicht zulässigem Ausschlussgrund (Vgl. dazu Methodik 2.6.25. Punkt 9).

Von den 25 Studien mit vorliegenden Zahlenangaben zum Ausschluss von Studienpatienten lag in fünf Studien keine Begründung vor. In 20 Fällen waren Begründungen vorhanden, nur sechs davon konnten jedoch als valide eingestuft werden. Bei der Beurteilung der Validität wurde auch in Betracht gezogen, ob die Größenordnung der ausgeschlossenen Patienten damit einsichtig ist.

1) Beispiel für validen Ausschlussgrund:

Walach et al. 1997: Drei von 98 Pat. wurden ausgeschlossen wegen folgenden Verstößen gegen die Ein- bzw. Ausschlusskriterien: Einnahme von Substanzen, welche als mögliche antidotische Substanzen zum homöopathischen Arzneimittel klassifiziert waren, Schwangerschaft (32).

2) Beispiel für nicht- validen Ausschlussgrund:

Andrade et al. 1991: Elf von 44 Patienten wurden ausgeschlossen: Zehn Patienten (jeweils fünf in der Placebo bzw. Verum-Gruppe) beendeten den Versuch vor Ablauf von zwei Monaten wegen „lack of efficacy“, also fehlender Wirksamkeit. Diese zehn Patienten fanden in der Auswertung keinerlei Berücksichtigung. Ein Patient erlitt eine psychotische Krise und hatte somit einen validen Ausschlussgrund (Studienthematik war die Behandlung rheumatoider Arthritis)

Die Anzahl ausgewerteter und ausgeschlossener Patienten in Relation zur Gesamtzahl der Studienteilnehmer sowie eine Übersicht zur Angabe und Validität der Ausschlussgründe zeigt Tabelle 10.

Tabelle 10: Summation der Daten zur Auswertung der Patienten

Abkürzungen:

n.v.: nicht valide

n.a.: nicht angegeben

**:* Zahlenangabe unklar

V: valide

n.e.: Gründe nicht evaluiert

rand./rekrut.: randomisiert/rekrutiert

Anteil der ausgewerteten an rand./rekrut. Patienten in %	Anzahl der randomisierten/ rekrutierten Patienten	Anzahl der ausgewerteten Patienten	Anzahl der ausgeschlossenen Patienten	Anteil der ausgewerteten an rand./rekrut. Patienten in %	Anteil der ausgeschlossenen an ausgewerteten Patienten in %	Ausschlussgrund
Studie 1	11	10	1	91%	10%	n.v.
Studie 2	15	15*	2*	100%	13%	n.v.

Fortsetzung S. 42

Studie 3	195	46	149	24%	324%	n.v.
Studie 4	42	42	0	100%	-	-
Studie 5	67	60	7	90%	12%	n.v.
Studie 6	101	98*	3	97%	3%	v
Studie 7	118	n.a.	n.a.	-	-	n.e.
Studie 8	0	n.a.	n.a.	-	-	n.e.
Studie 9	31	31	0	100%	-	-
Studie 10	20*	n.a.	n.a.	-	-	n.e.
Studie 11	30	n.a.	n.a.	-	-	n.e.
Studie 12	44	33	11	75%	33%	n.v.
Studie 13	40	40	0	100%	-	-
Studie 14	34	34	0	100%	-	-
Studie 15	60	60	0	100%	-	-
Studie 16	63	60	3	95%	5%	n.v.
Studie 17	74	55	19	74%	35%	v
Studie 18	20	18*	8	90%	44%	n.a.
Studie 19	43	n.a.	n.a.	-	-	n.e.
Studie 20	87	81	6	93%	7%	v
Studie 21	36	n.a.	n.a.	-	-	n.e.
Studie 22	44	38	6	86%	16%	n.a.
Studie 23	100	59	24*	59%	41%	n.v.
Studie 24	119	85	34	71%	40%	n.v.
Studie 25	213	164	49	77%	30%	n.v.
Studie 26	243	201	42	83%	21%	n.v.
Studie 27	121	86	35	71%	41%	n.v.
Studie 28	36	33	3	92%	9%	n.v.
Studie 29	200	n.a.	n.a.	-	-	n.e.
Studie 30	34	n.a.	n.a.	-	-	n.e.
Studie 31	147	126	21	86%	2%	v
Studie 32	174	162	12	93%	1%	n.a.
Studie 33	40*	40*	10*	100%	33%	n.v.
Studie 34	76	n.a.	n.a.	-	-	n.e.
Studie 35	10	10	2	100%	20%	n.a.
Studie 36	80*	44*	n.a.	55%	-	n.e.
Studie 37	46	43	3	94%	7%	n.a.
Studie 38	175	170	5	97%	3%	v
Studie 39	164	132	32	81%	24%	n.v.
Studie 40	161	160	1	99%	1%	v
Summe Patienten	3148	2236	488	-----	-----	-----
	Anzahl der Studien					
* = Zahlenangabe unklar	3	5	3			-
v = valide						6
n.v. = nicht valide						14
n.a. = nicht angegeben		9	10			5
n.e. = Gründe nicht evaluiert						10
Kein Ausschluss			5			-

3.9 Intent to Treat

Die Einhaltung des „Intent to Treat“- Prinzips wurde in nur drei der 40 Studien nachvollziehbar durchgeführt.

In 19 Studien musste die Frage nach dem „Intent to Treat“- Prinzip als „nicht beschrieben“ bewertet werden. Definitionsgemäß inkludierte dies alle Studien, bei welchen ein nicht begründetes Missverhältnis in der Anzahl der randomisierten/rekrutierten Patienten zu ausgewerteten Patienten bestand. Es ist anzunehmen, dass bei einer gewissen Anzahl dieser Studien der Intent to Treat verletzt wurde, ein Nachweis jedoch aufgrund mangelnder Information nicht erbracht werden konnte.

Eine Missachtung des „Intent to Treat“- Prinzips konnte nur für diejenigen Studien postuliert werden, welche dies selbst bestätigten oder bei denen das Ausschlussverfahren darauf hinwies. Davon waren 18 Studien betroffen. Von diesen 18 Studien wurde bei vier Studien eine „best-case-worst-case“- Analyse durchgeführt. Ein „best-case-worst-case“- Analyse beschreibt für die Ausschlusspatienten jeder Studiengruppe jeweils das beste beziehungsweise das schlechteste Outcome. Die Analyse war jedoch in zwei Fällen nicht verwertbar:

Bei der ersten Studie (van Erp und Brands 1996) betrug die „drop-out“- Rate in den zwei Studiengruppen jeweils über 25% (elf von 41 bzw. acht von 31 Patienten), so dass eine „best-case-worst-case“- Analyse nicht mehr sinnvoll ist.

Bei der zweiten Studie (Rahlf's und Mössinger 1979) wurde für vier Patienten eine „best-case-worst-case“- Analyse bei insgesamt 34 ausgeschlossenen Patienten durchgeführt.

3.10 Verblindung der Studienteilnehmer

Eine Verblindung fand bei 38 der 40 Studien statt, davon wurde in 34 Studien doppelt verblindet. Wenn ein Auswerter vorhanden und dieser zusätzlich zu Behandler und Studienpatient verblindet war, wurde die Konstellation einer „doppelten Verblindung“ gleichgestellt. Dieser Fall traf für drei Studien zu. In zwei Fällen wurde die Verblindung des Auswerter's und des Patienten, nicht jedoch des Behandler's beschrieben. Diese Situation entspricht einer Einfach-Verblindung und wurde auch als solche ausgewertet. Die weiteren zwei genannten Einfach-Verblindungen beschrieben eine Verblindung von „nur Arzt“ oder „nur Patient“.

Grundsätzlich ist bei jeder Studie wünschenswert, dass auf die Methodik der Verblindung eingegangen und diese dem Leser nachvollziehbar dargestellt wird. Dies war in 19 der 38 verblindeten Studien der Fall, davon entfielen zwei auf eine einfache Verblindung, 17 auf eine doppelte Verblindung. Der Verblindungsvorgang war in acht Arbeiten unzulänglich beschrieben, 13mal war keine Beschreibung vorhanden.

Einen Überblick zur Verblindung gibt Tabelle 11.

Tabelle 11: Verblindung der Studienteilnehmer

Verblindung der Studienteilnehmer	Einfach verblindet	Doppelt verblindet	Arzt	Patient	Auswerter	Niemand	Summe
	4	34	34	38	5	2	
1. Beschreibung des Verblindungsprozesses mit							
a) nachvollziehbarer Methodik	2	17					19
b) unklarer Methodik	2	6					8
2. Keine Beschreibung des Verblindungsprozesses		11				2	13

Anzumerken ist, dass bei neun der Studien, welche eine Verblindung beschrieben, keine Randomisierung der Studien statt gefunden hat! Da dieser Punkt in unserem Fragebogen qualitativ nicht ausgewertet wurde, jedoch eine eindeutige Beeinträchtigung der Validität darstellen kann, wird im Rahmen der Diskussion unter 4.1 dazu Stellung genommen.

3.11 Vergleich der Ausgangsrisiken mittels „Tabelle 1“

In 21 von 40 Studien war eine „Tabelle 1“ vorhanden. In den restlichen 19 Studien fehlte diese vollständig.

Inhaltlich bestand bei nur fünf der 21 Studien eine mangelhafte Ausführung der „Tabelle 1“. Dabei ist jedoch zu berücksichtigen, dass die Anforderungen bezüglich der inhaltlichen Angaben („mindestens zwei risiko-relevante Items vorhanden“) stark minimiert wurden (vgl. dazu Methodik 2.2.4 Punkt 16.). Es wurde ferner bei zwei Studien akzeptiert, dass rein prozentuelle Angaben zur Verteilung der Studienteilnehmer auf die risiko-relevanten Items gemacht wurden.

Unter den 16 Studien, bei denen eine risiko-relevante Beschreibung der Studienpopulation dargestellt war, lag eine Gleichverteilung der Items in zwölf Fällen vor.

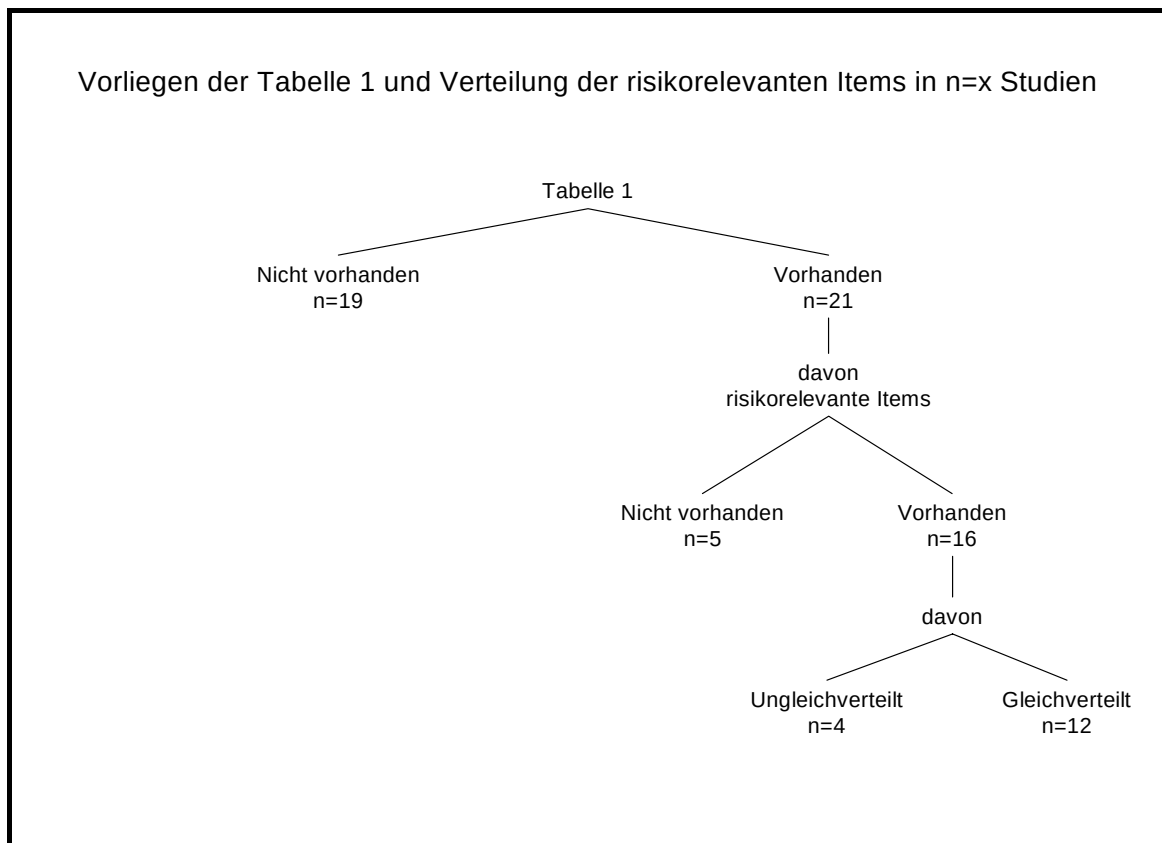


Abbildung 3: Vorliegen der Tabelle 1 und Verteilung der risiko-relevanten Items in n=x Studien

Die fünf Studien mit unzureichender Tabelle 1 wurden in Bezug auf die Punktevergabe der internen Validität ebenso mit „null“ Punkten bewertet wie die 19 Studien ohne Tabelle 1.

3.12 Gleichbehandlung der Studienteilnehmer

Die Auswertbarkeit dieser Fragestellung gestaltete sich schwierig, da in der Mehrzahl der Studien die notwendigen Informationen nicht greifbar waren. In der Regel war das Studienprotokoll, insbesondere der Umgang mit den Studienpatienten und entsprechende Verhaltensanweisungen nicht ausreichend erläutert. Aus diesem Grund konnte diese Fragestellung auch nicht als Kriterium zur Bewertung der internen Validität herangezogen werden.

Insgesamt war bei 22 der 40 Studien die Fragestellung nicht zu beantworten. Soweit eruierbar, konnte man bei zumindest sieben Studien eine Gleichbehandlung annehmen, bei elf Patienten musste man dies ausschließen.

3.13 Gibt es Ansätze, den speziellen Anforderungen der Homöopathie gerecht zu werden?

Genau bei der Hälfte der Studien (20 von 40 Studien) waren Bemühungen ersichtlich, einen Interventionsplan zu implementieren, der zumindest ansatzweise homöopathische Behandlungsgrundlagen würdigt.

Die nachfolgenden Behandlungsvarianten sowie Kombinationen daraus wurden implementiert:

- 1) Mehrere homöopathische Arzneimittel wurden präselektiert und aus diesem Repertoire das am Meisten zutreffende Präparat für den jeweiligen Studienpatienten ausgewählt. In vier Studien wurde diese Variante gewählt, zweimal mit einer Vorauswahl von acht Präparaten, einmal mit elf und einmal mit 200 Präparaten. Bei einer der Studien mit acht Präparaten durfte auch eine Kombination aus zwei Arzneimitteln verabreicht werden.
- 2) Die umgekehrte Version wurde in zwei Studien appliziert: Es fand eine Vorauswahl an homöopathischen Arzneimitteln statt, anschließend wurden nur Patienten in die Studie aufgenommen, deren Symptombild mit den betreffenden Mitteln behandelbar war.
- 3) Drei Studien verwendeten für alle Studienteilnehmer das selbe Arzneimittel, es fand jedoch eine individuelle Dosisanpassung je nach „Reizlage“ und Ausprägung der Beschwerden statt.
- 4) Bei neun Studien wurden die homöopathischen Substanzen individuell für jeden Studienpatienten selektiert. Dabei wurde teilweise ein standardisiertes Computerprogramm („RADAR“) zu Hilfe genommen, teilweise bestimmte ein Gremium aus mehreren homöopathisch geschulten Behandlern gemeinschaftlich das passende Therapeutikum.

Anmerkung: RADAR ist ein spezielles Softwareprogramm welches ein homöopathisches Repertorium zur Verfügung stellt. Es bietet unter Anderem eine Hilfestellung zur Auswahl homöopathischer Arzneimittel an, zum Beispiel nach symptomorientierten Krankheitsbildern.
- 5) Drei Studien orientierten sich an dem dualen Grundkonzept einer Basis- und einer Akuttherapie. Das heißt, ein Arzneimittel wird gewählt, welches die Grundkonstitution des Patienten berücksichtigt und ein Arzneimittel, welches das aktuelle Symptombild berücksichtigt. Zwei dieser Studien

entschieden dabei individuell die Auswahl der passenden Präparate wie bereits in 3) beschrieben.

- 6) Bei einer Studie wurde folgendes Konzept gewählt: Nach individueller Arzneimittelselektion fand eine Zwischenevaluation statt, bei Besserung der Beschwerden bekamen die Patienten im weiteren Studienverlauf ein Placebo, bei Verschlechterung ein Verum. Begründet wurde diese Vorgehensweise damit, dass es in der Homöopathie lege artis sei, nur bei Symptomverschlechterung die aktive Behandlung fortzusetzen, und zwar gegebenenfalls mit einer Umstellung der Arzneimittel.

Aus den oben aufgeführten Ansätzen ergeben sich diverse Problematiken. In der Diskussion unter Punkt 4.4 wird die Thematik nochmals aufgegriffen und die Möglichkeiten einer validen Studiendurchführung unter Berücksichtigung homöopathischer Behandlungsgrundsätze erörtert.

Wie im Methodik-Teil unter Punkt 25. vermerkt, wurde diese Fragestellung nicht als Validitätskriterium verwendet, eine qualitative Analyse der hier beschriebenen Vorgehensweisen fand somit nicht statt.

Es sollte jedoch überprüft werden, ob ein Unterschied in der Validitätseinstufung zwischen Studien mit homöopathischen Therapiekonzepten im Vergleich zu konventionellen Therapiekonzepten ermittelt werden konnte. Um feststellen, ob eine Differenz zwischen dem subjektiven Eindruck und der objektivierten, internen Validität besteht, wurden die jeweiligen berechneten Mediane einander gegenüber gestellt.

Das Ergebnis stellt sich wie folgt in Tabelle 12 dar:

Tabelle 12: Subjektive und interne Validität der Studien mit homöopathischem versus konventionellem Therapiekonzept

	Homöopathisches Therapiekonzept n=20 Studien	Konventionelles Therapiekonzept n=20 Studien
	Interne Validität (Schulnoten 1,0-6,0)	
Median	3,5	5,0
IQR	(4,0-5,0)	(4,0-6,0)
	Subjektive Validität (Schulnoten 1,0-6,0)	
Median	4,3	5,5
IQR	(3,3-6,0)	(5,0-6,0)

Die Gegenüberstellung zeigt bei der subjektiven wie bei der internen Validität eine Verbesserung von 1,2, bzw. von 1,5 Schulnoten im Median bei Studien mit homöopathischen Therapieansätzen im Vergleich zu konventionellen Therapieansätzen.

Die subjektive Einschätzung fiel insgesamt kritischer aus: beim Vergleich der Mediane beläuft sich die Differenz immerhin auf eine Verschlechterung um 0,8 (homöopathisches Therapiekonzept) bzw. 0,5 Noten (konventionelles Therapiekonzept) gegenüber der internen Validitätsbeurteilung.

3.14 Formulierung des Zieles einer Studie

29 der 40 Studien (72,5%) gaben die Hauptfragestellung klar verständlich an. Bei elf Studien (27,5%) wurde die Zielsetzung der Studie nicht eindeutig verfasst, teilweise fehlte eine Aussage dazu komplett. Die häufigsten Problematiken waren:

- 1) Eine völlig unklare bis hin zur nicht vorhandenen Angabe der Fragestellung. Der Leser muss die Zielsetzung aus dem beschriebenen Studienablauf rückfolgern.
- 2) Eine undifferenziert formulierte Fragestellung. Beispiele: „Der Vorteil („benefit“) von Arnika montana beim akuten Hirninfarkt“, „Homöopathische Behandlung bei Osteoarthritis“. In Anbetracht der Komplexität solcher Krankheitsbilder wurde die Zielsetzung nicht ausreichend eingegrenzt. Als Beispiel zum Vergleich: „Die analgetische Wirksamkeit von Cantharis bei geringgradigen Verbrennungen“.

3.15 Evaluation der Ergebnisangaben

Die Fragestellung wurde in drei wesentliche Punkte unterteilt:

3.15.1 Sind grundsätzlich Zielgrößen angegeben, mit denen die Ergebnisse gemessen wurden?

Die Frage konnte in 36 der 40 Studien mit „ja“ beantwortet werden, in drei der Studien mit „nein“.

Bei drei der mit „nein“ beantworteten Studien sind die Auswertungen der Studienergebnisse unklar:

- 1) Savage und Roe 1977: "nine (47%) of these 19 women responded to placebo.....giving a total placebo response rate of 57%." Es wird im gesamten Ergebnisteil nicht erklärt, welche Größen wie gemessen wurden oder auf was sich derartige Zahlenangaben beziehen.
- 2) Ustianowski 1974: Im Ergebnisteil wird in reinen Prozentangaben über eine Besserung/Verschlechterung oder gleichbleibende Symptomatik mit jeweils Verum oder Placebo berichtet.
- 3) Solanki und Ghandi 1995: „Of 34 cases, 11 (57,89%) cases out of 19 cases responded extremely well in category (a) while 2 (13,13%) cases also responded extremely well out of 15 cases in the controlled group" Anhand der Studienbeschreibung kann die Auswertung der Daten nicht nachvollzogen werden, zumal unklar ist, welche Zahlenwerte zur Verum- und welche zur Placebo-Gruppe gehören.

3.15.2 Waren die Skalierungen der Zielgrößen diskriminativ?

In 24 von 40 Fällen, also über 50% der Studien, konnte diese Frage nicht positiv beantwortet werden. In 13 der 40 Studien waren die geforderten Kriterien für die Zielgrößen und Messskalen erfüllt (siehe Methodikteil 2.6 Punkt 21. 2).

Als unzureichend wurden folgende Maßangaben gewertet:

- 1) Besserung der Symptomatik: viel/wenig
oder
Grad der Dehydrierung: wenig, mittel, viel.
- 2) Skalierung nur in Form einer Drei-Punkte-Skala statt einer Zehn-Punkte-Skala wie z.B. einer Visual-Analog-Scale (häufig anzutreffendes Problem)
- 3) Vage und relative Angaben ohne Bezug und Maßzahl: diese sind nicht verwertbar
- 4) Rein subjektive Angaben/Eigenevaluationen von Patienten ohne objektifizierbare Kriterien
- 5) Zeitabstände, welche nicht engmaschig genug gewählt wurden (Beispiel: Beurteilung des Verlaufs allergischer Reaktionen zweimal in 14 Tagen)
- 6) Unspezifische Messkriterien wie die Erhebung der Kreatininkinase zur Beurteilung der Stärke eines Muskelkaters. Dieser Laborwert kann bei diversen Ereignissen eine Erhöhung zeigen und ist daher kein guter Indikator.

3.15.3 Sind die Evaluationskriterien dem Krankheitsbild angemessen?

Bei nur 13 der Studien sind die Evaluationskriterien ausreichend differenziert um dem untersuchten Krankheitsbild gerecht zu werden. Dies lag teilweise an zu großzügig formulierten Einschlusskriterien, welche eine zu breit gefächerte, inhomogene Studienpopulation bezüglich des Beschwerdespektrums ergaben.

3.16 Durchführung einer Eigenevaluation

Das Ausmaß und der qualitative Wert einer kritischen Auseinandersetzung war nicht Gegenstand dieser Auswertung. Ebenso wenig sollte damit überprüft werden, ob diese Eigenevaluation im Vergleich mit den hier erhobenen Validitätskriterien übereinstimmte. Ziel war ausschließlich eine Erfassung der Autoren, welche zu ihrer eigenen Arbeit Stellung bezogen. In 18 von 40 Studien, also bei nahezu der Hälfte aller Autoren, wurde eine kritische Anmerkung bezüglich der eigenen Arbeit vorgenommen. Bei 22 Autoren war kein Kommentar vorliegend.

3.17 Vergleich der subjektiven Validität mit der internen Validität

Dieser Vergleich diente zur Kontrolle der Qualität der subjektiven Einschätzung der Studien anhand der bewerteten internen Validitätsmerkmale.

Die subjektive Validität betrug im Median 5,0, IQR (3,9-6,0), die interne Validität wich davon um genau eine halbe Schulnote nach oben ab und betrug im Median 4,5, IQR (4,0-6,0).

Wie aus Abbildung 4 zu entnehmen ist, belief sich die maximale Differenz zwischen den beiden Validitätsgruppen auf 2,0 Schulnoten. Die Verteilung der Abweichungen war wie folgt:

Abweichung um

- 1) Zwei Schulnoten: Dreimal
- 2) Eine Schulnote: Achtmal
- 3) 0,5 Schulnoten: Neunmal

Eine völlige Übereinstimmung in der Notenvergabe wurde bei der Hälfte der Fälle, also in 20 Studien erzielt.

Einen Überblick über die Notenverteilung gibt Abbildung 4:

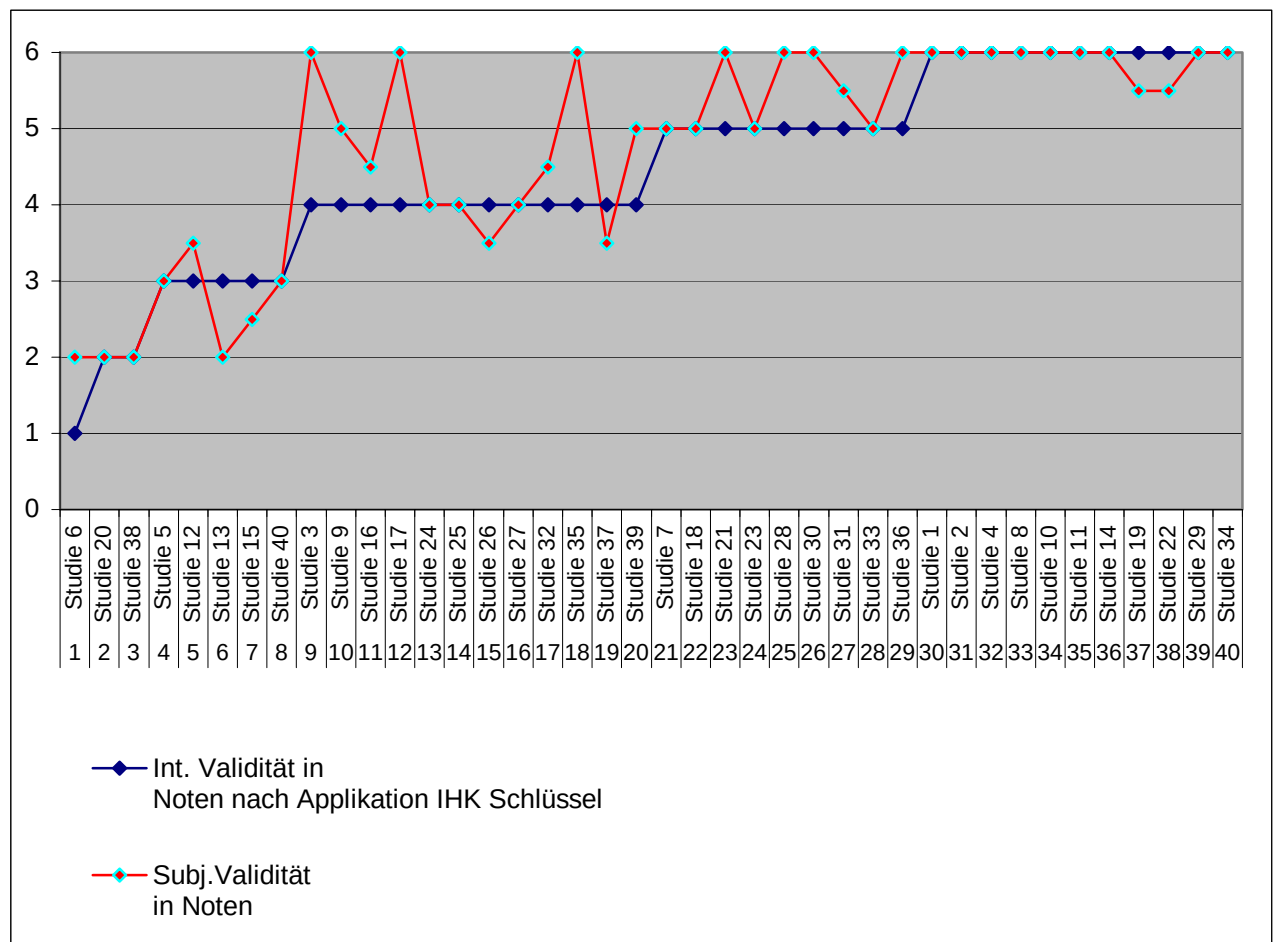


Abbildung 4: Vergleich der subjektiven und der internen Validität

Abkürzungen:

Int. Validität: Interne Validität

Subj. Validität: Subjektive Validität

3.18 Häufigkeitsverteilung der Validitätskriterien

Um einen Überblick über die Häufigkeitsverteilung von beachteten und missachteten Validitätskriterien zu erlangen, wurde die Punkteverteilung der internen Validität betrachtet. Wie in Methodik 2.6.,25. berichtet, konnte jede Studie maximal zwei Punkte für jedes der 18 Validitätskriterien erzielen. Insgesamt betrug die maximal erreichbare Punktzahl 36 Punkte pro Studie.

Es zeigte sich, dass alle Studien mindestens vier Punkte erreicht haben, nur eine Studie erreichte 33 Punkte. Die Punkteverteilung lag im Median bei 17,5 Punkten, IQR (10,0;23,0) bei einem Minimum von 4,0 und einem Maximum von 33,0 Punkten.

Abbildung 5 zeigt die kumulative Verteilung der Studien je nach Anzahl der erreichten internen Validitätspunkte:

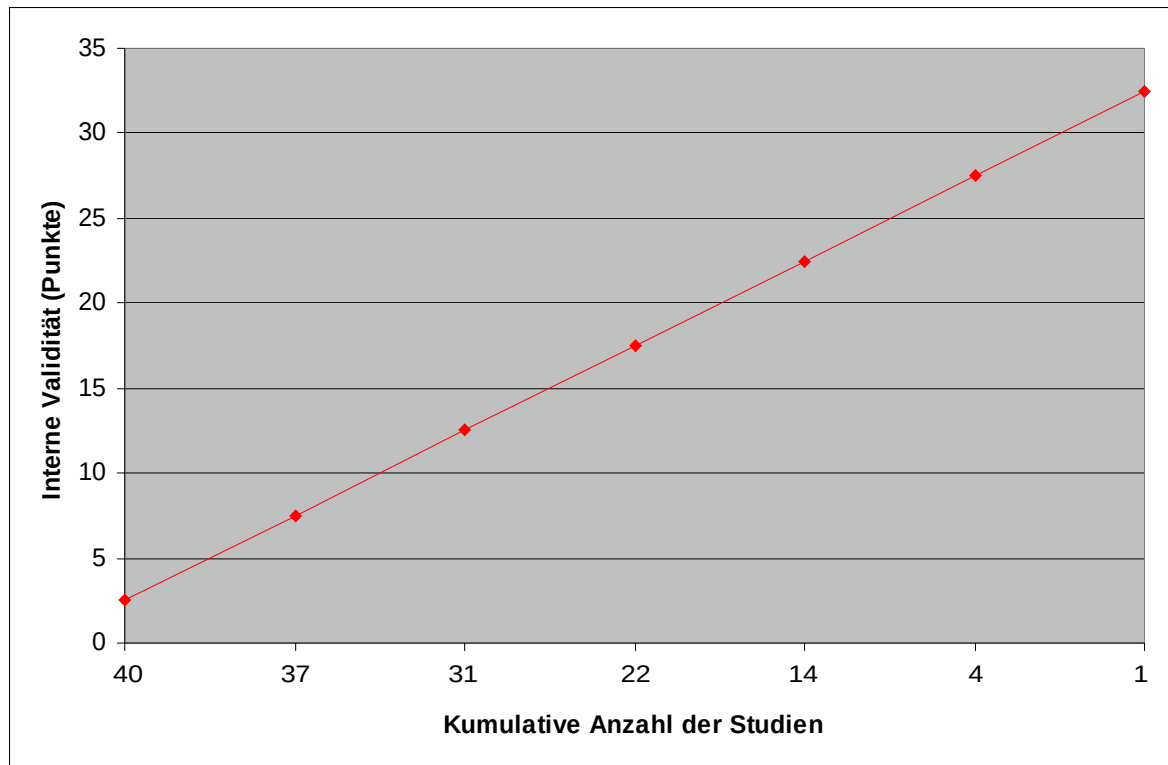


Abbildung 5: Kumulative Darstellung der Studienanzahl nach Validitätspunkten

Betrachtet man die Gesamtpunktzahl, die sich für jedes Validitätskriterium im Verlauf der Bearbeitung der 40 Studien subsummieren lässt, so lassen sich folgende Feststellungen machen:

- Die maximale Punktzahl, die für jedes der Kriterien bei optimaler Beantwortung im Sinne von „kein Validitätsverstoß“ erzielt werden konnte, war 80 Punkte (je 2 Punkte bei 40 Studien). Im Median wurden 36,5 Punkte, IQR (23;52) erreicht, das Minimum lag bei sechs Punkten, das Maximum bei 74 Punkten.
- Die Häufigkeit, mit der gegen ein bestimmtes Kriterium verstoßen wurde (Vergabe der Punktzahl „null“) betrug im Median 21. Das heißt, dass jedes Validitätskriterium im Median 21mal missachtet wurde, IQR (11;27) bei einem Minimum von zwei und einem Maximum von 38 Verstößen.

Die häufigsten verletzten Validitätskriterien (Vergabe von null Punkten):

- 1) Kriterium zehn: „Intent to Treat“ (siehe 3.9)
- 2) Kriterium fünf: nicht nachvollziehbare Differenzen und invalide Begründungen beim Vergleich der randomisierten zu rekrutierten Patientenzahlen (siehe 3.6.2)
- 3) Kriterium neun: Angaben zur Anzahl ausgeschlossener Patienten (siehe 3.8.)

Die am meisten beachteten Validitätskriterien (Vergabe von zwei oder einem Punkt):

- 1) Kriterium 16: Angabe der Zielgrößen, wobei nur dieser Teilbereich der Auswertung zu den Zielgrößen den Validitätsanforderungen in der Mehrzahl der Studien entsprach. Die weiteren Bewertungskriterien, welche sich mit der Skalierung der Zielgrößen und der damit beabsichtigten Erfassung des Krankheitsbildes beschäftigte, wiesen ein erheblich schlechteres Ergebnis auf (siehe 3.15)
- 2) Kriterium elf: Art der Verblindung (siehe 3.10)
- 3) Kriterium 15: Formulierung der Hauptfragestellung (siehe 3.14)

Die exakte Punkteverteilung der internen Validitätskriterien mit maximal erreichter Punktzahl pro Kriterium sowie Häufigkeit der „null-Punkte-Vergabe“ und somit einen Überblick über die Einhaltung bzw. Missachtung der Kriterien zeigt Tabelle 14.

Tabelle 14: Punkteverteilung der internen Validitätskriterien

Studien- nummer	Interne Validitätskriterien: Punkteverteilung (0-2 Punkte pro Studie)																		Gesamt- punkt- zahl (max. 36) der int. Validität pro Studie
	1) Herstellung d. Arzneimittel	2) Informed consent	3) Einschlusskriterien	4) Randomisierung	5) Begründung zur Anzahl rekrutierten u. randomisierten Pat.	6) Durchführung d. Randomisation	7) Geheimhaltung d. Randomisierungsplanes	8) Angaben zur Anzahl ausgewerteter Pat.	9) Angaben zur Anzahl ausgeschlossener Pat.	10) Intent to treat	11) Art der Verblindung	12) Beschreibung des Verblindungsprozesses	13) Vorhandensein der Tabelle 1	14) Verteilung der risiko-relevanten items	15) Formulierung der Hauptfragestellung	16) Angabe der Zielgrößen	17) Skalierung der Zielgrößen	18) Erfassung d. Krankheitsbildes	
1	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	2	0	0	6
2	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	2	0	0	4
3	2	2	1	0	0	0	0	2	0	0	1	1	2	2	0	2	2	2	19
4	0	0	1	0	0	0	0	2	2	0	2	0	0	0	2	2	0	0	11
5	2	1	2	2	2	0	1	2	0	0	2	0	0	0	2	2	2	2	22
6	2	1	2	2	2	2	2	0	2	2	2	2	2	2	2	2	2	2	33
7	2	0	0	2	0	2	2	0	0	0	2	1	0	0	2	2	0	0	15
8	2	2	0	1	0	2	1	0	0	0	0	0	0	0	0	2	0	0	10
9	2	0	1	2	0	0	1	2	2	0	2	0	0	0	2	2	0	2	18
10	2	2	1	1	0	1	2	0	0	0	2	1	0	0	0	2	0	0	14
11	0	0	1	1	0	0	1	0	0	0	2	2	0	0	0	2	0	0	9
12	2	1	2	2	0	0	2	2	0	0	1	2	2	2	2	2	2	2	26
13	0	0	2	2	0	0	1	2	2	2	2	0	2	1	2	2	2	2	24
14	0	0	1	2	0	0	1	2	2	0	0	0	0	0	2	0	0	0	10
15	0	0	1	2	0	0	1	2	2	0	2	2	2	2	2	2	2	2	24
16	2	1	2	2	2	0	1	2	0	0	2	0	2	1	2	2	0	0	21
17	0	1	1	2	2	0	1	2	2	0	2	1	0	0	2	2	0	0	18
18	0	0	2	2	0	1	1	0	0	0	2	0	2	1	2	2	0	0	15
19	0	1	1	0	0	0	0	0	0	0	2	1	0	0	2	2	0	0	9
20	2	2	2	2	2	0	2	2	2	0	2	2	2	2	2	2	2	0	30
21	2	0	1	2	1	1	0	0	0	0	2	2	0	0	2	2	0	0	15
22	0	0	1	0	0	0	0	2	0	0	2	1	0	0	2	2	0	0	10
23	0	0	0	2	0	2	1	2	0	0	2	2	0	0	2	2	2	0	17

Fortsetzung S. 55

24	2	0	1	2	0	0	1	2	0	0	2	2	2	2	2	2	0	0	20
25	2	0	2	1	0	2	2	2	0	0	2	2	0	0	2	2	0	0	19
26	2	0	2	2	0	2	1	2	0	0	2	2	2	2	2	2	0	0	23
27	2	0	2	2	0	2	1	2	0	0	2	2	2	2	2	2	0	0	23
28	2	0	1	1	0	0	1	2	0	0	2	0	0	0	0	2	2	2	15
29	2	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	5
30	0	1	1	2	0	0	1	0	0	0	2	0	0	0	2	2	2	2	15
31	1	0	1	2	0	1	2	2	2	0	2	2	0	0	0	2	0	0	17
32	1	2	2	2	1	1	2	2	0	0	2	2	0	0	2	2	0	0	21
33	0	0	0	1	0	1	0	0	0	0	2	2	2	1	0	2	2	0	13
34	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	4
35	2	1	2	2	1	2	2	2	0	0	2	2	0	0	2	0	0	0	20
36	0	0	1	1	0	0	1	0	0	0	2	2	0	0	2	2	0	0	11
37	2	2	1	0	0	0	0	2	0	0	2	1	2	2	2	2	2	2	22
38	0	1	2	2	2	0	2	2	2	2	2	0	2	2	2	2	2	2	29
39	2	0	1	2	0	0	2	2	0	0	2	2	2	2	2	2	0	0	21
40	0	1	1	2	0	1	2	2	2	0	2	2	2	2	2	2	0	2	25
Erreichte Punktzahl (max.80)/ Kriterium bei n=40 Studien	42	22	45	55	15	23	41	52	22	6	72	46	32	28	58	74	26	24	
Häufigkeit des Verstoßes (Kriterium mit "0" beantwortet)	18	24	8	9	31	25	11	14	29	38	2	13	24	24	12	3	27	28	

4 Diskussion

4.1 Einfluss von Studien mit geringer Validität auf die Qualität von Übersichtsarbeiten

Um die Fragestellung unserer Arbeit zu beantworten, wurden nur wissenschaftliche Arbeiten verwendet, die Eingang in ein systematisches Cochrane-Review gefunden haben. Diesem Kriterium liegt die Annahme zugrunde, dass die Studien bezüglich ihrer Qualität bereits präselektiert sind. Insbesondere bei Meta-Analysen und Systematischen Reviews geht der Leser von einem Gütesiegel aus und erwartet Transparenz und Validität - Kriterien, welche von der Cochrane Collaboration selbst angestrebt werden (12).

Der gemeinsame Konsensus der fünf Reviews und der beiden Meta-Analysen, aus welchen die nach Durchlauf unseres Auswahlverfahrens verbliebenen Studien stammten, war eindeutig und stimmt mit den Ergebnissen dieser Arbeit überein:

- 1) Die Evidenz ist nicht ausreichend, um eine Aussage über die Wirksamkeit der homöopathischen Behandlungen zu machen (Aussage aller fünf Reviews und beider Meta-Analysen)
- 2) Bei nahezu allen Arbeiten wurde die mangelhafte Qualität aufgrund methodischer Fehler der Studien konstatiert und die ungenügende Evidenz damit in Verbindung gebracht. Teilweise wurden diese Mängel als gravierend betrachtet. Ernst (1999) stellte in seinem Review über homöopathische und konventionelle Behandlungsmethoden mit insgesamt sechs Studien fest: „None is free from serious methodological flaws“ (7).

In den Meta-Analysen wurde bemerkt, dass auch ein anerkanntes Scoring-System für die Qualitätseinstufung wie beispielsweise das „Jadad-Score-Assessment“ (3, 14) die Validität nicht adäquat bewerten konnte (20) und systematische Fehler (Bias) und Validitätsmängel der zugrunde liegenden Studien damit nicht unbedingt aufgedeckt wurden (5). Diese Einschätzung stützte sich hauptsächlich auf die Erfahrung der Autoren in der Einschätzung der Studien. Obwohl Mängel der primären Studien in beiden Meta-Analysen erkannt wurden, wurde aus den Ergebnissen abgeleitet, dass ein homöopathischer Effekt, der nicht Placebo-abhängig ist, nicht ausgeschlossen werden kann.

Damit wird ein Fehler begangen: Die mangelnde Aussagekraft und die damit assoziierte Gefahr einer falschen Schlussfolgerung aufgrund der ungenügenden Qualität der „Basis“ systematischer Reviews und Meta-Analysen wird zwar registriert, die notwendigen Konsequenzen werden aber nicht abgeleitet. Für den „Konsumenten“ von systematischen Reviews und Meta-Analysen sind letztendlich die Schlussfolgerungen relevant. Im Verlauf der Bewertung geäußerte Bedenken und Hinweise hinsichtlich qualitativer Einschränkungen werden möglicherweise nicht erkannt oder registriert, vor Allem, wenn sie nicht beweisbar sind. Die korrekte Konsequenz einer Meta-Analyse, die auf methodisch mangelhaften Studien beruht ist lediglich, dass auf der Basis der vorliegenden Information keine verlässliche Aussage getroffen werden kann. Dieses Problem wurde bereits von Porzsolt et al. (2005) beschrieben (25).

Unter Einbeziehung der Ergebnisse der vorliegenden Arbeit können folgende einschränkende Einflussfaktoren auf die Aussagekraft von Übersichtsarbeiten bei qualitativ mangelhaftem Studienmaterial identifiziert werden:

1) Ein- und Ausschlusskriterien

Ein- und Ausschlusskriterien für die Aufnahme von Studien sind nicht einheitlich und hinsichtlich grundsätzlicher Validitätsmerkmale wie Randomisierung und Verblindung nicht ausreichend stringent definiert.

- Alle Übersichtsarbeiten und beide Meta-Analysen akzeptierten Studien, in welchen „randomisiert oder verblindet“, „quasi randomisiert“ oder „nur randomisiert“ wurde. Das heißt, es wurden auch Studien aufgenommen, welche nicht randomisiert waren. Interessanterweise wurden jedoch in jeder der nicht randomisierten Studien die Therapien maskiert (verblindet). Eine Doppel-Verblindung ohne zufällige Zuteilung (Randomisierung) ist schwer realisierbar und eine Einfach-Verblindung ohne Randomisierung ist nicht sinnvoll (siehe Ergebnisteil 3.10).

Randomisierte, kontrollierte Studien und eine doppelte Verblindung sind Goldstandard und gerade zum Ausschluss des Placebo-Effekts in der Homöopathie erforderlich. Dass dies von Seiten der homöopathischen Gemeinschaft als teilweise nicht vereinbar mit ihren Behandlungskonzepten betrachtet wird (27,35) wird in Punkt 4.4 näher erörtert.

- Essentiell für eine klare Aussage in systematischen Reviews ist die möglichst einheitliche und klar eingegrenzte Studiena Auswahl hinsichtlich der untersuchten Fragestellung. Bei nahezu einem Drittel (elf von 40) der hier bewerteten Studien war die Fragestellung nicht exakt formuliert und konnte allenfalls aus der Studienbeschreibung rückgeschlossen werden. Bei unklar definierten Endpunkten und nicht beschriebenen Zielgrößen kann dieses Problem schwierig zu lösen sein. Es ist gut vorstellbar, dass Unklarheiten in der Zielsetzung der Primärstudien die Stellungnahme eines Reviews zu einer bestimmten Thematik erschwert. Der Leser systematischer Reviews, der klinische Entscheidungen treffen muss, erwartet jedoch eine Antwort auf eine dezidierte Fragestellung.
- In einem Vergleich systematischer Reviews zur Alternativmedizin von Linde und Willich (2003) zeigte sich, dass unterschiedliche Einschlusskriterien einen bestimmenden Faktor für differierende Stichprobengrößen und Populationen darstellen und damit zu unterschiedlichen Schlussfolgerungen führen (22). Auch Ziegler et al. (2004) wiesen darauf hin, dass „schon kleine Unterschiede im Studiendesign, z.B. bei den Ein- und Ausschlusskriterien, zu verschiedenen Ergebnissen in dem Grad des Wirksamkeitsunterschieds zwischen zwei Therapieverfahren führen können“ (37). In unserer Arbeit erwiesen sich in 19 von 40 Studien die Einschlusskriterien als zu unklar formuliert, um das Patientenkollektiv präzise zu beschreiben. In acht Studien fehlten die Einschlusskriterien vollständig. Damit wird die Aussagekraft von Übersichtsarbeiten hinsichtlich der Repräsentativität und des Wirksamkeitsnachweises für eine bestimmte Patientenpopulation stark eingeschränkt.
- Positiv zu betrachten ist die Tatsache, dass gerade die Cochrane-Collaboration bemüht ist, einen Publikationsbias durch Einschluss nicht-deutschsprachiger und englischsprachiger Publikationen zu vermeiden. Leider musste im Rahmen der vorliegenden Arbeit auf den Einschluss anderssprachiger Publikationen zu Gunsten der Durchführbarkeit verzichtet werden. Eine Verzerrung ist jedoch in unserem Fall auszuschließen, da bei der Vorauswahl viele Studien bereits aufgrund anderer

Ausschlusskriterien ausgesondert wurden und von den verbleibenden 129 Studien nur neun Studien aufgrund der Sprache und der Publikation zurück gewiesen werden mussten (siehe 3.1, Tabelle 2).

2) Assessment von Validitätskriterien

Die Verwendung verschiedener Assessment-Systeme zur qualitativen Beurteilung von Studien beeinflusst die Validitätseinstufung durch unterschiedliche Gewichtung von Kriterien (16). Der „Informed Consent“ und das „Intent to Treat“-Prinzip wurde in unserer Arbeit am häufigsten missachtet (in 40% der Studien erfolgte kein Informed Consent, ein Intent to Treat konnte nur in 5% der Studien nachgewiesen werden).

Auf Rang drei der am häufigsten missachteten Elemente erwiesen sich die fehlenden Angaben zum Ausschluss von Patienten. Die lückenhafte Berichterstattung der Ergebnisse ist hinsichtlich der Tatsache problematisch, dass bei nahezu einem Viertel der Studien (neun von 40) zusätzlich die Anzahl der ausgewerteten Patienten nicht vorlag. Diese Studien sollten die Mindestanforderungen für eine Evaluation in einem systematischen Review nicht erfüllen, da keine verlässlichen Aussagen zu den Ergebnissen getroffen werden können. Lückenhafte Berichterstattungen erschweren die Anwendung von Assessment-Systemen und führen möglicherweise zu inkorrekten Beurteilungen. Linde et al. (2001) schrieb angesichts dieser Problematik bei der Qualitätsevaluation von 207 randomisierten, kontrollierten Studien: „A few words (or the lack thereof) can determine whether a trial is considered „high“ or „low“ quality“ (21).

3) Interventionsangaben

Die untersuchten Interventionen müssen klar definiert sein. Das bedeutet, dass zwischen tatsächlich homöopathischen Therapien und Präparationen aus dem Bereich der Naturheilkunde und ähnlichen Bereichen differenziert werden sollte.

Die Mehrzahl der unter dem Suchbegriff „homeopath*“ erhobenen Suchergebnisse wies ein buntes Spektrum an untersuchten Präparaten aus dem gesamten Bereich der Naturheilkunde, Komplementärmedizin und nicht zuletzt tatsächlich der Homöopathie auf.

Auch bei Anwendung weniger rigoroser Ein- und Ausschlusskriterien als der hier implementierten entspricht ein großer Anteil der Studienpräparate nicht dem Grundbegriff eines homöopathischen Arzneimittels (13). Aus diesem Grund wurden 14 von 29 Reviews und von den verbliebenen 258 Einzelstudien nochmals etwa die Hälfte ausgeschlossen. In dem kleinen Kollektiv der verbleibenden fünf Reviews und der beiden Meta-Analysen beziehen sich alle Titel auf homöopathische Therapieoptionen. Dennoch sind in beiden Meta-Analysen sowie in einem der Reviews Studien enthalten, welche nicht klassische homöopathische Mittel verwendeten wie z.B. Traumeel® und Rheumaselect®. Gerade aus Sicht der Homöopathen, für die die korrekte Präparation der Arzneimittel wesentlich ist und in direktem Zusammenhang mit der Wirkungsweise steht, wird bei der Gegenüberstellung solcher Studien nicht Vergleichbares miteinander verglichen (4,36).

Katrak (2004) untersuchte 121 publizierte "Critical Appraisal Tools" und stellte eine beträchtliche Variabilität hinsichtlich der inhaltlichen Komponenten und der Zielsetzung fest (16). In dem Review fällt auf, dass die am geringsten vertretenen Items der für systematische Reviews designierten Appraisal Tools die Randomisierung, Verblindung und Intervention betrafen - ein Feststellung, welche unsere Ergebnisse bestätigten. Katrak beschrieb weiter, dass nur wenige Critical Appraisal Tools mit qualitativem Fokus identifiziert werden konnten - trotz der Hinweise vieler Guidelines auf die Anforderungen bezüglich methodischer Gesichtspunkte von Studienpublikationen.

Es gibt sicher keinen Gold-Standard für die Evaluation der Validität von Studien, zumal, wie in dem Review von Katrak festgestellt werden konnte, die Orientierung eines Critical Appraisals auch von der Art und Zielsetzung der Studie (experimentell, retrospektiv etc.), respektive einer Meta-Analyse oder eines systematischen Reviews abhängt.

Mangelhaftes Ausgangsmaterial darf jedoch nicht zu falschen Schlussfolgerungen führen. Daher ist die Notwendigkeit eines gewissen Konsensus hinsichtlich grundlegender Qualitäts-Standards für die Publikation von Einzelstudien wie auch deren Aufnahme in Übersichtsarbeiten durchaus gegeben.

4.2 Grundsätzliche validitätseinschränkende Faktoren

Wie bereits eingangs erwähnt, hat diese Analyse gezeigt, dass es eine Reihe validitätskompromittierender Faktoren gibt, welche bei jeder Studiendurchführung berücksichtigt werden sollten, unabhängig vom thematischen Hintergrund:

1) Detaillierte Studienprotokolle mit vollständigen Informationen

Grundvoraussetzung für eine valide Studie ist die Formulierung klarer Studienprotokolle. Dabei sollte nicht nur auf ein gut durchstrukturiertes Studiendesign, sondern auch auf eine akkurate und vor allem detaillierte Berichterstattung geachtet werden.

In den analysierten Studien mussten beschriebene Vorgehensweisen wiederholt hinterfragt werden, weil diese nicht präzise formuliert waren.

Beispiele:

- Eine Randomisierung wurde angegeben, wobei Informationen zum Randomisierungsplan fehlten.
- Eine Verblindung wurde ohne Angabe der Details beschrieben.
- Der Informed Consent wurde eingeholt ohne dessen Inhalt zu beschreiben.
- Endpunkte waren ohne Angabe der verwendeten Messgrößen und/oder deren Skalierung angegeben.

2) Unvollständige Patientendaten

Angaben zu Patientenzahlen müssen vollständig und nachvollziehbar sein. Dazu gehören folgende Informationen:

- Rekrutierte oder randomisierte Patientenanzahl sowie zahlenmäßige Verteilung der Patienten auf die verschiedenen Studiengruppen
- Anzahl der ausgeschlossenen Patienten mit Angabe, in welche Gruppe sie ursprünglich randomisiert waren, um den Intent to Treat zu verifizieren.

- Anzahl der ausgewerteten Patienten pro Studiengruppe bzw. Anzahl der ausgewerteten Patienten pro untersuchte Zielgröße in ihrer jeweiligen Studiengruppe.

Werden bestimmte Auswertungskriterien im Methodikteil einer Studie definiert, so erwartet der Leser diesbezüglich eine Darstellung über der Ergebnisse.

Studien, bei denen weder Evaluationskriterien definiert wurden noch eine Auswertung von Zielgrößen statt fand, sind unzulänglich.

Ein Beispiel hierfür ist die Arbeit von Solanki und Ghandi (1995) (32):

Placebo-kontrollierte Studie, Behandlung von Amöben- und Gardia- Infektionen mit verschiedenen homöopathischen Arzneimitteln.

Die Ergebnisse bestanden aus folgenden Angaben:

1. Auswertungstabelle: Jeweils Placebo- und Verum-Spalte mit „cured/not cured“
2. Balkengraphik, x-Achse unbeschriftet. Titel: Häufigkeit der verordneten Arzneimittel.

Messgrößen, Endpunkte und Art der Evaluation sind nicht nachvollziehbar!

Es kann nicht Aufgabe des Lesers sein, über fehlende Daten zu spekulieren, durch Berechnung aus den restlichen Angaben Rückschlüsse zu ziehen oder fehlende Textangaben aus Tabellen und Graphiken zu rekonstruieren.

Die Verwendung absoluter, numerischer Zahlenangaben erleichtert eine konsequente Verfolgung der Daten über den Verlauf der Studie. Relative Angaben wie Prozentangaben erschweren die Transparenz, insbesondere wenn der Bezug nicht eindeutig ersichtlich ist.

Statistischen Angaben sollten der Studienpopulation und Verteilung der Daten angepasst sein.

In unserer Arbeit fehlten bei neun von 40 Studien, also bei nahezu 25% die Angaben sowohl zur Anzahl der ausgeschlossenen wie zur Anzahl der ausgewerteten Patienten. Dabei sind die Studien, bei welchen unstimmgige oder fehlende Daten zur initialen Probandenzahl vorlagen nicht berücksichtigt (siehe Ergebnisse 3.6 und 3.8). Derartige Lücken in der Berichterstattung machen eine Akzeptanz der Auswertung und Aussage unmöglich.

3) Beschreibung der Ausgangspopulation und Risiko-Verteilung auf die Studiengruppen

In dieser Arbeit war in insgesamt 27 der 40 Studien (nahezu 70%) die Zusammensetzung der Studienpopulation nicht ausreichend nachvollziehbar (unzulänglich formulierte oder fehlende Einschlusskriterien, siehe Ergebnisse 3.5).

Die Beschreibung der Patientencharakteristika (üblich in Tabelle 1) fehlte ebenfalls in 19 der 40 Studien, in fünf Fällen waren die darin berichteten Angaben nur bedingt geeignet, um das Ausgangsrisiko der eingeschlossenen Population abzuschätzen. Die Risiko-Verteilung in den untersuchten Populationen ist somit in über der Hälfte der Studien nicht einzuordnen. Nur mit Kenntnis der Ausgangspopulation kann der praktisch tätige Mediziner entscheiden, ob die Resultate für sein Patientenkollektiv relevant sind. Das bedeutet, die Formulierung der Ein- und Ausschlusskriterien und die Darstellung der „Tabelle 1“ hat nicht nur Einfluss auf die Validität einer Studie, sondern direkte Auswirkungen auf die klinische Applikation von Studienergebnissen (8).

Leider konnten im Rahmen dieser Arbeit die Ursachen für die unpräzisen Formulierungen und lückenhaften Angaben zu den Studienkollektiven nicht festgestellt werden. Die Begründungen zu den Patienten-Ausschlüssen fehlten oder waren nur bedingt plausibel. Als Begründung fehlender Auswertungsdaten wurden wiederholt Verstöße gegen das Studienprotokoll (unerlaubte Medikamente eingenommen, sonstige, nicht näher beschriebene Protokollverstöße) sowie mangelnde (Nach-) Betreuung von Patienten (Patienten nicht mehr auffindbar, unvollständige oder falsche Datenerhebungen, Behandlung abgebrochen) angeführt.

Unklar ist, ob mangelnde Erfahrung der Studienleiter in der Durchführung und möglicherweise in der Publikation zu der Problematik beitrug. Letztendlich ist nicht eruierbar, in wie weit gewisse Validitätskriterien beachtet wurden, jedoch eine unvollständige Berichterstattung zu Missverständnissen führte. Bellavite et al. vermutete in seiner Evaluation der homöopathischen Studienlage durchaus einen der Hauptgründe für die geringe Zahl gut gestalteter Studien im Defizit entsprechend ausgebildeter Homöopathen mit wissenschaftlichen Qualifikationen und Interessen (2).

Aus analytischer Sicht wären viele der hier angeführten Mängel ohne größeren Aufwand zu beheben und hätten damit eine erhebliche Verbesserung der Stu-

dienqualität zu Folge. Bei Jonas et al. gelangte man nach einer sehr sorgfältigen Untersuchung von 59 homöopathischen Studien anhand von 33 Validitätskriterien zu ähnlichen Schlussfolgerungen: Eine Verbesserung der Berichterstattung, insbesondere hinsichtlich der Studiencharakteristiken und der Beschreibung der Studienpopulation würde eine deutliche Aufwertung der Qualität erbringen (15). Ein Zusammenhang zu der homöopathischen Thematik kann bei diesen Fehlerquellen nicht hergestellt werden. Somit ist zu erwarten, dass die beobachteten Mängel für Studien auf dem Gebiet der Homöopathie nicht spezifisch sind.

4.3 Kriterien mit unklarem Einfluss auf die Validität

1) Monozentrische versus multizentrische Studiendurchführung

Multizentrische Studien bieten die Möglichkeit größerer Objektivität. Die Voraussetzung dafür ist ein ausführliches und einheitliches Studienprotokoll, welches von den Studienleitern an den verschiedenen Lokalisationen eingehalten wird. Nur so lässt sich eine Vergleichbarkeit der Ergebnisse gewährleisten oder lassen sich Gründe für verschiedene Outcomes nachvollziehen.

In den acht multizentrischen Studien wurden im Median 120 (IQR 72-189) Patienten rekrutiert. In sechs der acht multizentrischen Studien (siehe Tabelle 7 im Ergebnisteil 3.6.1) wurden > 101 Patienten rekrutiert, während in mehr als 50% der hier evaluierten Studien die Zahl der rekrutierten Probanden geringer als 100 war. Tabelle 7 zeigt, dass in der Regel die Studien, welche eine größere Probandenzahl aufwiesen, auch eine bessere Einstufung der Validität erzielten.

2) Herstellung der Medikamente

Es ist als Mangel zu werten, dass in einem Drittel der Studien die Herkunft der Arzneimittel nicht bekannt gegeben wurde. Zudem stimmte in fünf Fällen die Herkunft des Placebos nicht mit der des Studienpräparates überein (siehe Ergebnisse 3.3). Deshalb bleibt ungeklärt, ob tatsächlich eine Übereinstimmung von Farbe, Form und Geschmack zwischen Verum und Placebo gewährleistet war.

Der Frage der Herkunft gewinnt möglicherweise an Bedeutung vor dem Hintergrund, dass die korrekte (und aufwendige) Herstellung des Arzneimittels unter Verwendung qualitativ hochwertiger Rohstoffe aus Sicht der Homöopathen essentiell für seine Wirkungsweise ist (13). Das heißt, dass die Transparenz eingeschränkt ist, wenn die Herstellung eines homöopathischen Arzneimittels aus homöopathischer Sicht als bedeutendes Kriterium zu betrachten ist.

3) Gleichbehandlung der Studiengruppen

Die Frage nach der Gleichbehandlung der Studiengruppen (außer der experimentellen Therapie) ist bei vielen Studien schwierig zu beantworten, da häufig nicht genug Informationen vorliegen, um eine eindeutige Aussage zu treffen. In unserer Analyse konnte bei 22 der 40 Studien diese Thematik aufgrund der spärlichen Beschreibung des Studienprotokolls ebenfalls nicht evaluiert werden.

Wie dem Leser die Bemühung um eine erhöhte Transparenz durch zusätzliche Informationen vermittelt werden kann, wird an folgenden herausgegriffenen Beispielen demonstriert:

- Alle Patienten erhielten zusätzlich eine Liste mit Verhaltensregeln und Diätvorschriften (34)
- Alle eingeschlossenen Patienten erhielten die selben Verbände (Patienten mit Verbrennungen) (19)
- Tabellarische Aufschlüsselung zusätzlich angewendeter geburtshilflicher Maßnahmen (Akupunktur, Vollbad, Analgetika-Gabe) mit Zuordnung zur Homöopathie oder Placebogruppe (1)

4) Eigener “Critical Appraisal”

In keinem Standard-Assessment-System ist die eigene kritische Auseinandersetzung als Bestandteil einer validen Studie gefordert oder als Bewertungskriterium etabliert. Es liegt jedoch nahe, dass abschließende, kritische Eigenanalysen den Qualitätsstandard von Studien fördern können. Durch die Tatsache, dass in der Regel misslungene Projekte selten publiziert werden, fehlt ein wesentliches Spektrum des wissenschaftlichen

Erfahrungsaustausches. Die Problematiken, die sich im Zuge der Durchführung einer Studie ergeben haben, sind dem Studienleiter respektive Autor am klarsten ersichtlich. Damit bietet sich die Chance, in Retrospektive Verbesserungsansätze zu evaluieren wie auch gelungene Konzepte und Vorgehensweisen hervorzuheben und künftigen Forschungsarbeiten mögliche Wiederholungsfehler zu ersparen.

4.4 Spezielle Problematik der Homöopathie

Die lebhafte Diskussion um den Wirksamkeitsnachweis der Homöopathie ist nach wie vor aktuell und wird wiederbelebt durch Veröffentlichungen wie die von Shang et al. (2005): "Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy". In der Meta-Analyse wurden 110 Placebo-kontrollierte homöopathische Studien mit Studien aus der konventionellen Medizin (aus der Cochrane Datenbank) hinsichtlich Erkrankung und Outcomes abgeglichen. Bei beiden Gruppen zeigte sich, dass qualitativ schlechtere Studien und solche mit niedrigeren Fallzahlen häufiger einen postulierten Unterschied nachweisen konnten als qualitativ bessere Studien mit höherer Fallzahl. Nach Berücksichtigung des Bias auf beiden Seiten war die Evidenz zugunsten der Effektivität konventioneller Interventionen größer. Dies veranlasste den Autor zu der Schlussfolgerung, die klinischen Effekte der Homöopathie Placebo-Effekten gleichzusetzen (30)

Milgrom (2005) diskutierte, ob konventionelle Studienkonzepte und insbesondere der konventionelle Gold-Standard der randomisierten, placebo-kontrollierten Studien geeignet seien, um die Effektivität der Homöopathie nachzuweisen. Aufgrund des holistischen Grundgedankens würde eine Separierung der Einheit Behandler-Arzneimittel-Patient durch die Struktur von randomisierten, placebo-kontrollierten Studien grundsätzlich mit dem homöopathischen Wirkprinzip in Konflikt stehen (23).

Bei der nachgewiesenen eingeschränkten Studienqualität drängt sich die Frage auf, in wie weit Schlussfolgerungen wie Shangs (29) legitim sind und ob auf dieser Basis die Problematik des Gold-Standards überhaupt diskutierbar und für die

unzureichende Aussagekraft homöopathischer Studien verantwortlich zu machen ist. Möglicherweise wäre es sinnvoller, primär die festgestellten Validitätsmängel zu korrigieren und dann die Eignung methodischer Konzepte zu evaluieren.

1) Integration homöopathischer Behandlungskonzepte

Im Ergebnisteil 3.13 wurden verschiedene Therapievarianten beschrieben, welche bei den bearbeiteten Studien verwendet wurden. Nachdem die Studien mit homöopathischen Behandlungsansätzen ein im Median um 1,5 Schulnoten (interne Validität) besseres Ergebnis erzielten, soll auf diese Konzepte eingegangen und ihre speziellen Problematiken erörtert werden.

- Präselektion homöopathischer Arzneimittel, Einschluss von Patienten nur, wenn das betreffende Symptombild damit behandelbar ist

Problematik:

- a) Unter Umständen langwierige und aufwendige Rekrutierungszeit
- b) Gefahr, dass auch unpassende Patienten aufgenommen werden um die notwendige Studiengröße zu erreichen

- Kombination aus Basis- und symptomorientierter Akuttherapie

Problematik:

- a) Die homöopathische Behandlung kann dabei nur als Gesamtes verglichen werden, die einzelne Arzneimittelwirkung kann nicht geprüft werden
- b) Alle Placebo-Patienten müssen ebenfalls eine duale Therapie erhalten

- Präselektion mehrerer homöopathischer Arzneimittel, dann Auswahl des geeigneten Präparates je nach Patient aus diesem Pool
- Individuelle Arzneimittelselektion entweder durch einen homöopathisch geschulten Behandler, ein „Expertengremium“ oder durch ein Computersystem
- Gleiches Arzneimittel für alle Studienpatienten jedoch individuelle Dosisanpassung (unterschiedliche Potenzen)

Problematik:

- a) Die Reihenfolge der Rekrutierung, Anamnese, Medikamentenzuordnung und Randomisierung ist streng zu beachten
- b) Aufwendiges Studienkonzept: Für jedes der verwendeten homöopathischen Arzneimittel wird ein korrespondierendes Placebo benötigt. Dies gilt auch für gleiche Mittel mit unterschiedlicher Potenz (also unterschiedlicher „Dosierung“)!
Fraglich ist, ob der erforderliche Aufwand, welcher je nach Anzahl der Präparate erheblich ist, erbracht werden kann: Bei 200 individuell selektierten Therapien (wie in der Studie von Gibson 1978 (9) beschrieben, siehe Ergebnisse Tab.3) ist die Bereitstellung eines jeweiligen Kontrollarzneimittels unter Umständen ein schwieriges Unterfangen und in diesem Umfang sicher nicht mehr glaubwürdig.

Erleichternd mag die Tatsache sein, dass homöopathische Arzneimittel in der Regel geschmacksneutral sind und die Träger-substanz (Lactose bzw. Ethanol) günstig herzustellen ist. Somit eignet sich das Mittel in seiner unpräparierten Grundform sehr gut als Placebo (zur Herstellung darf auf die Informationen des deutschen Markführers homöopathischer Arzneimittel, der Deutschen Homöopathie-Union DHU-Arzneimittel GmbH & Co. KG, verwiesen werden (6)).

- Konsekutive Reevaluation des Therapieerfolges unter der Prämisse, nur bei Symptomverschlechterung mit der homöopathischen Behandlung fortzufahren oder einen Wechsel des homöopathischen Arzneimittels vorzunehmen (18). Das bedeutet im Rahmen einer Studienführung: der verblindete Behandler dokumentiert nach einem im Vorhinein definierten Zeitraum die aktuellen Symptome und trifft die Entscheidung, ob eine Weiterführung der bestehenden Therapie oder eine Änderung erfolgen soll.

Problematik:

- a) Bei Aufrechterhaltung des Randomisierungs-codes und der Verblindung weiß der Behandler nicht, ob das Ausbleiben des

Therapieerfolges auf eine Placeboeinnahme zurückzuführen ist oder auf ein unpassend gewähltes Homöopathikum.

- b) Dieses Konzept bietet jedoch auch die Chance, Komponenten der Behandler-Patient-Interaktion und deren Einfluss auf die Therapie zu evaluieren (z.B. durch Eigenevaluation: wann glaubte der Behandler/Patient, das Verum verabreicht/erhalten zu haben?)

Essentiell bei allen Konzepten ist eine akkurate Beschreibung des Vorgehens. Im Rahmen dieser Arbeit war bei nur acht der 20 Studien mit speziellem, homöopathischem Behandlungsansatz das Studiendesign ausführlich genug beschrieben, um die Reihenfolge der Ereignisse wie Anamnese, Randomisierung und Verabreichung der Medikamente nachzuvollziehen. Eine explizite Aussage, dass für jedes der homöopathischen Medikamente ein Placebo bereit stand, wurde in keiner der Studien formuliert.

2) Der Placebo-Effekt

Einfluss des Placebo-Effektes:

Der Placebo-Effekt ist letztendlich ein multifaktorielles Geschehen, der von diversen Komponenten abhängig ist. Diese sind im Wesentlichen die Erwartungshaltung und Vorprägung des/der Behandler(s), die Erwartungshaltung und Vorgeschichte der Patienten und die in den Patienten investierte Zuwendung. Der psychologische Effekt ist, zumindest wenn man den Ergebnissen dieser Arbeit Rechnung trägt, aufgrund der methodischen Mängel nicht mehr durchschaubar und vom pharmakologischen Wirkmechanismus der Präparate nicht differenzierbar.

Am Beispiel von Gibsons Studie „Salicylates and homeopathy in rheumatoid arthritis: preliminary observations“ lässt sich die Problematik gut demonstrieren (9):

Die Studie wurde einfach verblindet durchgeführt. Die drei Patientengruppen wurden von unterschiedlichen Ärzten betreut: Die Homöopathie-Gruppe wurde von Homöopathen einer homöopathischen Klinik, die Placebo- und Salicylat-Gruppe von konventionellen Medizinerinnen einer Rheumaklinik betreut.

Zudem durften Patienten in der Homöopathie-Gruppe ihre konventionelle Behandlung beibehalten, Patienten in der Salicylat-Gruppe mussten diese abbrechen, über die Placebo-Gruppe liegt keine Information vor.

Damit ist nicht gewährleistet, dass die Zuwendung in beiden Gruppen vergleichbar war. Der größere Zeitaufwand für die homöopathische Anamnese und Weiterbetreuung ist bekannt und muss als „confounder“ berücksichtigt werden.

Zudem unterliegt die Studie einem Bias bezüglich der Erwartungshaltung auf Seite der Ärzte wie der Patienten. Allein die Überzeugung vieler Patienten, dass Schulmediziner weniger Gesprächsbereitschaft und Engagement bei der Behandlung zeigen und die emotionalen Bedürfnisse ihrer Patienten vernachlässigen (17) kann den Behandlungserfolg beeinflussen.

3) Wirksamkeitsnachweis der Homöopathie: Pharmakologischer Effekt, psychologischer Effekt oder beides?

Bei allen evaluierten 40 Studien bestand das Studienziel (soweit nachvollziehbar) im Wirksamkeitsnachweis der verwendeten homöopathischen Präparate. Letztendlich stellt sich für den Wissenschaftler die Frage, wonach er beim spezifischen Wirksamkeitsnachweis der Homöopathie suchen soll: dem psychologischen Effekt, dem pharmakologischen Effekt oder dem Zusammenwirken beider Komponenten? Vincent und Furnham beschrieben bereits 1999 die wissenschaftliche Untersuchung der Komponenten des therapeutischen Prozesses in der Komplementärmedizin als „greatly neglected area, but one with immense possibilities for research“ (33).

Der kontrollierte, randomisierte, doppelt verblindete Versuch stellt für diese Fragestellung nach wie vor das klassische Studienkonzept dar. Notwendig sind zusätzliche Evaluationskriterien, welche die Erwartungshaltung der Patienten und der Behandler sowie deren Vorprägung bezüglich alternativer Behandlungsmethoden dokumentieren. Shekelle et al. forderten dies bereits bei der Durchführung pragmatischer Studien in der Komplementärmedizin ebenso wie Informationen zum Ausbildungs- und Erfahrungsstand der Behandler (31). Für die Entwicklung von randomisierten, kontrollierten Studien zur Differenzierung

des psychologischen Effektes vom pharmakologischen Effekt sind diese Informationen essentiell- ebenso wie die detaillierte Berichterstattung.

4.5 Ist eine Bearbeitung von Studien in dargestellter Weise zur adäquaten Validitätsbewertung notwendig?

Ist es möglich, anhand eines standardisierten Verfahrens eine subjektive Einschätzung vorzunehmen oder ist eine formale Punktebewertung notwendig? Das heißt, kann man durch das subjektive Verfahren ähnliche Ergebnisse erzielen wie durch das aufwendige Verfahren?

Bei dieser Untersuchung handelt es sich um ein wissenschaftliches Projekt, welches mit entsprechendem Aufwand verbunden war. Es sollte nicht als Aufforderung verstanden werden, den selben Aufwand bei allen Studien zu betreiben, zumal die Anwendbarkeit in der Praxis definitiv nicht gegeben ist.

In einer sorgfältigen Analyse wurde gezeigt, dass Nichtbeachtung von Validitätskriterien zu falschen Interpretationen führen kann. Es bedarf keiner speziellen epidemiologischen Kenntnisse oder Erfahrungen, um diese Merkmale zu identifizieren. Ein geschärftes Bewusstsein für die zu beachtenden Kriterien reicht aus, um eine subjektive Einschätzung wissenschaftlicher Arbeiten zu vorzunehmen. Diese Arbeit zeigt, dass die subjektive Bewertung nicht erheblich vom Punkteverfahren abweicht. Diese leichte Erlernbarkeit eines Critical Appraisals wurde von verschiedenen Gruppen beschrieben (26).

4.6 Ist eine standardisierte Bewertung von Studienmaterial für systematische Reviews und Meta-Analysen erforderlich?

Die Notwendigkeit für jeden klinisch wie wissenschaftlich tätigen Mediziner, sich mittels der vorhandenen Literatur weiter zu bilden, ist unbestreitbar. Niemand wird in der Lage sein, zu allen Fragestellungen der Medizin die jeweiligen Originalarbeiten zu lesen. Deshalb sind alle Kollegen auf eine möglichst valide, objektive und wahrheitsgetreue Wiedergabe wissenschaftlicher Beobachtungen in

Übersichtsarbeiten angewiesen. Mit Recht geht der Leser von einer Art Qualitätssiegel aus und erwartet ein hohes Niveau an Transparenz und Validität (10).

Dieses Niveau ist aber nur zu erreichen, wenn die „Bausteine“, auf welchen die Reviews beruhen, den geforderten Ansprüchen genügen und dieses Kriterium von den Verfassern der Reviews tatsächlich beurteilt werden kann.

Die Abfassung qualifizierter Reviews ist für den Erfolg unserer Gesundheitsversorgung von derart zentraler Bedeutung, dass man durchaus ein standardisiertes Zulassungsverfahren, welches neben epidemiologischem Werkzeug (skills) auch fachliches Wissen (knowledge) und eine von der Gesellschaft akzeptierte Einstellung (attitude) zur Gewährung von Gesundheitsleistungen, fordern könnte. Da durchaus wahrscheinlich ist, dass diese Kriterien nicht mehr von einem Wissenschaftler alleine erbracht werden können, könnte an die Übernahme solcher Aufgaben durch Expertenteams gedacht werden.

Meine Untersuchung hat gezeigt, dass in der Mehrzahl der Studien, die in einem Cochrane Review – welches den Gold-Standard wissenschaftlichen Arbeitens erreichen sollte – eingeschlossen waren, gravierende Mängel aufwies. Dieser Befund zeigt, dass das derzeit angewandte System zur Herstellung von Wissen wahrscheinlich nicht genügend kontrolliert ist, um die beschriebenen Mängel selbst zu entdecken und zu vermeiden.

Als bedeutendstes Ergebnis meiner Arbeit konnte ich zeigen, dass bei der Mehrzahl der untersuchten Studien Kriterien unbeachtet blieben, die heute als Grundlage valider wissenschaftlicher Arbeit und Berichterstattung erfüllt sein sollten. Viele dieser Kriterien sind durchaus unabhängig von der Thematik der Homöopathie zu sehen; es ist deshalb nicht auszuschließen, dass ähnlich grundlegende Mängel auch bei nicht-homöopathischen Studien zu finden sind.

Aus diesem Grund sind die Ergebnisse dieser Studie keineswegs nur für die Bewertung homöopathischer Studien geeignet. Sie eignen sich generell als Beispiel für die kritische Analyse klinischer Studien. Durch die Besonderheiten der Homöopathie entstehen jedoch bei einigen Kriterien, insbesondere bei der Randomisierung und der Verabreichung von Arzneimitteln zusätzliche Konflikte.

Aufgrund der generell eingeschränkten Validität aller untersuchten Studien ist eine Aussage zur Wirksamkeit der homöopathischen Therapie nicht möglich. Allerdings muss angesichts der Ergebnisse festgestellt werden, dass die Durchführung homöopathischer Studien und/oder ihre Darstellung nicht dem aktuellen Standard wissenschaftlichen Arbeitens entspricht. Ohne valide Studien wird es nicht möglich sein, die offene Frage zur Wirksamkeit der Homöopathie zu klären.

5 Zusammenfassung

Die Validität von Studien ist eine notwendige Voraussetzung, um ihre Ergebnisse bewerten zu können und die Studien in Übersichtsarbeiten zusammen zu fassen. Ist diese Voraussetzung nicht gegeben, kann weder eine fundierte Bewertung der Aussage einer Übersichtsarbeit statt finden noch über ihre Umsetzung in die klinische Versorgung entschieden werden.

Im Rahmen der vorliegenden Arbeit wurden 40 Studien zur Homöopathie, die in Meta-Analysen und Reviews der Cochrane Datenbank eingeschlossen waren, auf ihre Validität hin überprüft. Bei der Auswahl wurden nur Studien berücksichtigt, welche Arzneimittelzubereitungen entsprechend der klassischen Homöopathie nach Hahnemann verwendeten. Zur Bewertung der Validität wurde ein differenzierter Kriterienkatalog erarbeitet, welcher sich an international anerkannten Leitlinien der Evidenz-basierten Medizin zur qualitativen Bewertung von Studien orientiert.

Die Ergebnisse zeigten, dass in den Übersichtsarbeiten Studien eingeschlossen waren, in welchen grundlegende Validitätskriterien nicht eingehalten wurden. Dabei fiel insbesondere eine Missachtung der Randomisierung (fehlende Randomisierung in 9 von 40 Studien), der Angabe der Risikostratifizierung von Studiengruppen (keine „Tabelle 1“ in 19 von 40 Studien) und des „Intent to Treat“ (in 3 von 40 Studien eingehalten) auf. Nachgewiesen wurde ferner die lückenhafte Darstellung von Daten, insbesondere zur Angabe der ausgeschlossenen und ausgewerteten Patienten (in 9 von 40 Studien lag die Anzahl der ausgewerteten Patienten nicht vor) sowie Missverhältnisse in den numerischen Angaben der rekrutierten zu den randomisierten Patienten.

Neben der Identifizierung von Faktoren mit unklarem Einfluss auf die Validität der eingeschlossenen Studien (z.B. multizentrische versus monozentrische Studienführung, Herstellungsnachweis von Studienmedikamenten) stellte die spezielle Thematik der Homöopathie eine besondere Herausforderung dar.

Um die Notwendigkeit eines komplizierten Bewertungssystems zur Validitätsprüfung zu evaluieren, wurde ein Punktesystem entwickelt, welches den Vergleich zu einer anfänglich durchgeführten, subjektiven Einschätzung der Studien ermöglich-

te. Die Diskrepanz zwischen beiden Verfahrensweisen betrug im Median nur eine halbe Schulnote auf einer Skala von 1,0-6,0.

In der Mehrzahl der Studien, welche Eingang in die Reviews und Meta-Analysen fanden, lagen derart gravierende Mängel vor, dass aus den Daten keine Folgerungen abgeleitet werden konnten. Das bedeutet respektive, dass die Aussagekraft der Übersichtsarbeiten in Frage zu stellen ist.

Hinsichtlich der Aufnahme von Studien in Übersichtsarbeiten konnten als besonders kritische Faktoren identifiziert werden:

- 1) Die Applikation uneinheitlicher Auswahlkriterien und ein geringer Anspruch an gängige Validitätsmerkmale wie Randomisierung, Verblindung, Formulierung präziser Fragestellungen und differenzierter Ein- und Ausschlusskriterien bezüglich des Patientenkollektivs.
- 2) Die erschwerte Anwendung eines Assessment- Systems bei hohem Datenverlust durch fehlende Berichterstattung.
- 3) Eine hohe Interventionsbreite/inhomogene Interventionsauswahl durch Verwendung verschiedenster Präparate, welche nicht der homöopathischen Arzneimittelzubereitung entsprachen.

Als grundlegende Voraussetzungen für die Validität von Studien wurden festgestellt:

- 1) Die möglichst differenzierte Erarbeitung und Einhaltung von Studienprotokollen, inklusive der entsprechend detaillierten Berichterstattung.
- 2) Die zwingende Notwendigkeit einer kompletten Datenerhebung.
- 3) Die Risiko-Stratifizierung der Patientenpopulation.

Zusammenfassend lässt sich als wesentliches Ergebnis dieser Arbeit feststellen, dass in Übersichtsarbeiten aufgrund der deutlich eingeschränkten Validität der eingeschlossenen Studien weder Aussagen zur Wirksamkeit noch zur Unwirksamkeit gemacht werden können. Die Arbeit zeigt, dass selbst in Übersichtsarbeiten wissenschaftlich renommierter Datenbanken - wie der Cochrane Collaboration - unzulängliche Studien aufgenommen und Schlussfolgerungen gezogen werden, welche aufgrund der Validitätsmängel nicht zulässig sind.

6 Literaturverzeichnis

- (1) Beer AM, Heiliger F: Randomized, double blind trial of Caulophyllum D4 for induction of labour after premature rupture of membranes at term. Geburtshilfe Frauenheilkd 59: 431-435 (1999)
- (2) Bellavite P, Ortolani R, Pontarollo F, Piasere V, Benato G, Conforti A: Immunology and Homeopathy. 4. Clinical Studies-Part 2. Evid Based Complement Alternat Med 3: 397-409 (2006)
- (3) Clark O, Castro AA, Filho JV, Djubelgovic B: Interrater agreement of Jadad's scale. Cochrane Colloq 9th 2001, 1: op031
- (4) Clausius N: Kontrollierte klinische Studien zur Homöopathie. Eine systematische Übersichtsarbeit mit Metaanalyse. Med Dissertation, Ludwig-Maximilians-Universität München. KVC (Karl und Veronica Carstens Stiftung), Essen, S. 55-57 (1998)
- (5) Cucherat M, Haugh MC, Gooch M, Boissel J P: Evidence of clinical efficacy of homeopathy: a meta-analysis of clinical trials. Eur J Clin Pharmacol 56: 27-33 (2000)
- (6) Deutsche Homöopathie-Union DHU-Arzneimittel GmbH & Co. KG Ottostr. 24, 76227 Karlsruhe, <https://www.dhu.de>
- (7) Ernst E: Classical homoeopathy versus conventional treatments: a systematic review. Perfusion 12: 13-15 (1999)
- (8) Fäßler M: Analyse der Information der Table #1 in randomisierten klinischen Studien. Med Dissertation, Universität Ulm (2004)

- (9) Gibson RG, Gibson S, MacNeill AD, Gray GH, Dick WC, Watson Buchanan W: Salicylates and homeopathy in rheumatoid arthritis: preliminary observations. Br J Clin Pharmacol 6: 391- 395 (1978)
- (10) Guyatt GH, Sinclair J, Cook DJ, Glasziou P: User's Guides to the Medical Literature. J Am Med Assoc 281: 1836-1843 (1999)
- (11) Hahnemann S: Organon der Heilkunst. 6.Auflage, Burgdorf, Göttingen, S. 70-73 (§22-27), S. 132 (§124), S. 191-197 (§269-270), S. 198-199 (§273) (2001)
- (12) Higgins JPT, Green S, editors: Systematic reviews and the Cochrane Handbook. Cochrane Handbook for Systematic Reviews of Interventions 1.1 (updated May 2005); Sektion 1
<http://www.cochrane.org/resources/handbook/hbook.htm> (assessed 31 May 2005)
- (13) Homöopathisches Arzneimittelbuch 2007 (erste amtliche Fassung 1934), Deutscher Apotheker Verlag Stuttgart, Govi- Verlag- Pharmazeutischer Verlag GmbH Eschborn, S. 1/4- 4/4 (H 5.1-H5.2), S. 1/3-3/3 (H 5.4)
- (14) Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuary HJ: Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 17: 1-12 (1996)
- (15) Jonas WB, Anderson RL, Crawford CC, Lyons JS: A systematic review of the quality of homeopathic clinical trials. BMC Complement Altern Med 1: 12 (2001)
- (16) Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA: A systematic review of the content of Critical Appraisal Tools. BMC Med Res Methodol 4: 1-11 (2004)

- (17) Kirkcaldy B, Furnham A, Trimpop R, Knobloch J: Akzeptanz und Skepsis: Einstellungen und Verhalten von Patienten der Alternativ-und Schulmedizin. Zeitschrift für Gesundheitspsychologie 9: 57-66 (2001)
- (18) Kron M, English JM, Gaus W: Guidelines on methodology of clinical research in homeopathy. In: Ernst E, Hahn EG (Editors): Homeopathy. A Critical Appraisal. Butterworth-Heinemann, Oxford, S.9-47 (1998)
- (19) Leaman AM, Gorman D: Cantharis in the early treatment of minor burns. Arch Emerg Med 6: 259-261 (1989)
- (20) Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV, Jonas WB: Are the clinical effects of homeopathy placebo effects: a meta-analysis of placebo-controlled trials. Lancet 350: 834-843 (1997)
- (21) Linde K, Wayne BJ, Melchart D, Willich S: The methodological quality of randomized controlled trials of homeopathy, herbal medicines and acupuncture. Int J Epidemiol 30: 526-531 (2001)
- (22) Linde K, Willich SN: How objective are systematic reviews? Differences between reviews on complementary medicine. J R Soc Med 96: 17-22 (2003)
- (23) Milgrom LR: Are randomized controlled trials (RCTs) redundant for testing the efficacy of homeopathy? A critique of RCT methodology based on entanglement theory. J Altern Complement Med 11: 831-838 (2005)
- (24) Moher D, Schulz KF, Altman DG: The CONSORT Statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 357: 1191-1194 (2001)

- (25) Porzsolt F, Kajnar H, Awa A, Fässler M, Herzberger B: Validity of original studies in health technology assessment reports: Significance of standardized assessment and reporting. *Int J Technol Assess Health Care* 21: 1-4 (2005)
- (26) Remington TL, Foulk MA, Williams BC: Evaluation of evidence for inter-professional education. *Am J Pharm Educ* 70: 66 (2006)
- (27) Righetti M: Forschung in der Homöopathie. Burgdorf, Göttingen, S. 34-48 (1988)
- (28) Sackett DL, Richardson WS, Rosenberg W, Haynes RB: Ist die Evidenz zu einer Behandlung valide? In: Kunz R, Fritsche L, Evidenzbasierte Medizin, Zuckschwerdt, München, S.72-76 (1999)
- (29) Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB: Evidence-based Medicine: How to practice and teach EBM. Second Edition, Churchill Livingstone, Edinburgh (2000)
- (30) Shang A, Huwiler-Müntener K, Nartey L, Jüni P, Dörig S, Sterne JA, Pewsner D, Egger M: Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. *Lancet* 366: 726-732 (2005)
- (31) Shekelle PG., Morton S, Suttrop MJ, Buscemi N, Friesen C: Challenges in Systematic Reviews of Complementary and Alternative Medicine Topics. *Ann Intern Med* 142: 1042-1047 (2005)
- (32) Solanki M, Ghandi PM: Is homeopathy only a placebo therapy? *Homeopathic Heritage* 20: 707-713 (1995)
- (33) Vincent C, Furnham A: Complementary medicine: state of the evidence. *J R Soc Med* 92: 170-177 (1999)

- (34) Walach H, Haeusler W, Lowes T, Mussbach D, Schamell U, Springer W, Stritzl G, Gaus W, Haag G: Classical homeopathic treatment of chronic headache. *Cephalalgia* 17: 119-126 (1997)

- (35) Walach H: Wissenschaftliche Untersuchungen zur Homöopathie. KVC (Karl und Veronica Carstens Stiftung), Essen, S. 268-270 (2000)

- (36) Wein C: Qualitätsaspekte klinischer Studien zur Homöopathie. KVC (Karl und Veronica Carstens Stiftung), Essen, S. 41-42 (2002)

- (37) Ziegler A, Lange S, Bender R: Systematische Übersichten und Meta-Analysen. *Dtsch Med Wochenschr* 129: T11-T15 (2004)

7 Anhang

Literaturangaben zu den initial ausgewählten 29 Reviews und 3 Meta-Analysen sowie den 4 nicht ausgewertete Einzelstudien

1) Einzelstudien

Jacobs J, Jimenez LM, Gloyds S, Casares FE, Gaitan MP, Crothers D: Homeopathic treatment of acute childhood diarrhea. A randomized trial in Nicaragua. Br Homeopath J 82: 83-86 (1993)

Mössinger P: Untersuchung über die Behandlung der akuten Pharyngitis mit Phytolacca D2. Allg Homöopath Ztg 221: 177-83 (1976)

Mössinger P: Untersuchung zur Behandlung des akuten Fließschnupfens mit Euphorbium D3. Allg Homöopath Ztg 227: 89-95 (1982)

Mössinger P: Zur therapeutischen Wirksamkeit von Hepar sulfuris calcareum D4 bei Pyodermien und Furunkeln. Allg Homöopath Ztg 225: 22-27 (1980)

2) Reviews

Aikins Murphy P: Alternative therapies for nausea and vomiting of pregnancy. Obstet Gynecol. 91: 149-55 (1998)

Campbell K, Waters E, O'Meara S, Kelly S: Interventions for preventing obesity in children. Cochrane Database Syst Rev 3 (2001). Updates: Cochrane Database Syst Rev 2 (2002) and 3 (2005)

Criswell LA, Saag KG, Sems KM, Welch V, Shea B, Wells G, Suarez-Almazor ME: Moderate-term, low-dose corticosteroids for rheumatoid arthritis. Cochrane Database Syst Rev 2 (2002)

Ernst E, Barnes J: Are homeopathic remedies effective for delayed-onset muscle soreness: a systematic review of placebo-controlled trials. Perfusion 11: 4-8 (1998)

Ernst E, Pittler M H: Efficacy of homeopathic arnica: a systematic review of placebo-controlled clinical trials. *Arch Surg* 133:1187-1190 (1998)

Ernst E, Stevinson C: Ginkgo biloba for tinnitus: a review. *Clinical Otolaryngology and Allied Sciences* 24: 164-167 (1999)

Ernst E: Classical homoeopathy versus conventional treatments: a systematic review. *Perfusion* 12: 13-15 (1999)

Ernst E: Homeopathic prophylaxis of headaches and migraine: a systematic review. *Journal of Pain and Symptom Management* 18: 353-357 (1999)

Furlan AD, Brosseau L, Imamura M, Irvin E: Massage for low back pain. *Cochrane Database Syst Rev* 4 (2000). Update: *Cochrane Database Syst Rev* 2 (2002)

Garner SE, Eady EA, Popescu C, Newton J, Li Wan Po A: Minocycline for acne vulgaris: efficacy and safety. *Cochrane Database Syst Rev* 2 (2002). Update: *Cochrane Database Syst Rev* 1 (2003)

Gibbs S, Harvey I, Sterling JC, Stark R: Local treatment for cutaneous warts: systematic review. *Cochrane Database Syst Rev* 2 (2001). Update: *Cochrane Database Syst Rev* 3 (2003)

Glazener CM, Evans JH: Drugs for nocturnal enuresis children (other than desmopressin and tricyclics). *Cochrane Database Syst Rev* 3 (2000). Updates: *Cochrane Database Syst Rev* 4 (2003)

Glazener CM, Evans JH: Simple behavioural and physical interventions for nocturnal enuresis in children. *Cochrane Database Syst Rev* 2 (2002). Update: *Cochrane Database Syst Rev* 2 (2004)

Hondras MA, Linde K: Manual therapy for asthma. *Cochrane Database Syst Rev* 4 (2002). Update: *Cochrane Database Syst Rev* 2 (2005)

Hutton E, Mozurkewich E: Extra-amniotic prostaglandin for induction of labour. Cochrane Database Syst Rev 2 (2001)

Kavanagh J, Kelly AJ, Thomas J: Breast stimulation for cervical ripening and induction of labour. Cochrane Database Syst Rev 4 (2001). Updates: Cochrane Database Syst Rev 3 (2005)

Linde K, Jobst K A: Homeopathy for chronic asthma. Cochrane Database Syst Rev 2 (2000). Update: Cochrane Database Syst Rev 1 (2004)

Linde K, Melchart D: Randomized controlled trials of individualized homeopathy: a state-of-the-art review. J Altern Complement Med 4: 371-388 (1998)

Linde K, Mulrow CD: St John's wort for depression. Cochrane Database Syst Rev 2 (2000). Update: Cochrane Database Syst Rev 2 (2005)

Little C, Parsons T: Herbal therapy for treating rheumatoid arthritis. Cochrane Database Syst Rev 1 (2001)

Little CV, Parsons T, Logan S: Herbal therapy for treating osteoarthritis. Cochrane Database Syst Rev 1 (2001)

Lovell HG: Angiotensin converting enzyme inhibitors in normotensive diabetic patients with microalbuminuria. Cochrane Database Syst Rev 2 (2000). Updates: Cochrane Database Syst Rev 1 (2001)

Luckas M, Bricker L: Intravenous prostaglandin for induction of labour. Cochrane Database Syst Rev 4 (2000)

Melchart D, Linde K, Fischer P, Kaesmayr J: Echinacea for preventing and treating the common cold. Cochrane Database Syst Rev 2 (2000). Update: Cochrane Database Syst Rev 1 (2006)

Proctor ML, Murphy PA: Herbal and dietary therapies for primary and secondary dysmenorrhoea. Cochrane Database Syst Rev 3 (2004)

Smith CA: Homoeopathy for induction of labour. Cochrane Database Syst Rev 4 (2001). Update: Cochrane Database Syst Rev 4 (2003)

Stevinson C, Ernst E: Complementary/alternative therapies for premenstrual syndrome: a systematic review of randomized controlled trials. Am J Obstet Gynecol 185: 227-35 (2001)

Vernon H, McDermaid CS, Hagino C: Systematic review of randomized clinical trials of complementary/alternative therapies in the treatment of tension-type and cervicogenic headache. Complement Ther Med 7: 142-55 (1999)

Vickers AJ, Smith C: Homoeopathic Oscilloccinum for preventing and treating influenza and influenza-like syndrome. Cochrane Database Syst Rev 2 (2000). Update: Cochrane Database Syst Rev 1 (2004)

3) Meta-Analysen

Barnes J, Resch K L, Ernst E. Homeopathy for postoperative ileus: a meta-analysis. Clin Gastroenterol 25: 628-633 (1997)

Cucherat M, Haugh M C, Gooch M, Boissel J P. Evidence of clinical efficacy of homeopathy: a meta-analysis of clinical trials. Eur J Clin Pharmacol 56: 27-33 (2000)

Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV, Jonas WB: Are the clinical effects of homoeopathy placebo effects: a meta-analysis of placebo-controlled trials. Lancet 350: 834-843 (1997)

8 Tabellen- und Abbildungsverzeichnis

Tabelle 1:	Umrechnungstabelle für die Transformierung der Punktzahlen der internen Validität in Schulnoten durch Applikation der IHK-Punkteskala.....	22
Tabelle 2:	Summation des Ausschlussverfahrens.....	26
Tabelle 3:	Bearbeitete Studien.....	27
Tabelle 4:	Interne Validität in Abhängigkeit von der Zahl der beteiligten Institutionen pro Studie.....	29
Tabelle 5:	Übersicht über den Herstellungsort (intern vs. extern) der Studienpräparate bei n=40 Studien.....	30
Tabelle 6:	Zahl der rekrutierten und randomisierten Patienten in den 40 ausgewerteten Studien.....	33
Tabelle 7:	Größe der Studienpopulation und zugehörige interne Validität (Angabe von Median und interquartilem Range).....	34
Tabelle 8:	Verhältnis der randomisierten zu rekrutierten Patienten, n=8 Studien.....	36
Tabelle 9:	Differenz zwischen der Anzahl rekrutierter und randomisierter Patienten sowie Vorliegen von Begründungen.....	37
Tabelle10:	Summation der Daten zur Auswertung der Patienten.....	41
Tabelle 11:	Verblindung der Studienteilnehmer.....	44
Tabelle 12:	Subjektive und interne Validität der Studien mit homöopathischem versus konventionellem Therapiekonzept.....	47
Tabelle14:	Punkteverteilung der internen Validitätskriterien.....	54
Abbildung 1:	Flussdiagramm zur Studie von Campbell A (1976): Two pilot controlled trials of Arnica montana.....	22
Abbildung 2:	Interne Validität in Abhängigkeit von der Patientenanzahl pro Studie.....	33
Abbildung 3:	Vorliegen der Tabelle 1 und Verteilung der risiko-relevanten Items in n=x Studien.....	45
Abbildung 4:	Vergleich der subjektiven und der internen Validität.....	51
Abbildung 5:	Kumulative Darstellung der Studienanzahl nach Validitätspunkten.....	52

9 Danksagung

Allen, die zum Gelingen dieser Arbeit beigetragen haben, möchte ich von ganzem Herzen danken.

Dabei gilt ganz besonderer Dank meinem Doktorvater Prof. Dr. Franz Porzsolt. Er ermöglichte nicht nur diese Arbeit, sondern weckte mit seiner fast unerschöpflichen Energie und seinem unglaublich ansteckenden Enthusiasmus immer wieder die Motivation aufs Neue.

Für seine Geduld mit mir und dieser Arbeit bin ich zutiefst dankbar, ebenso wie für alle fachlichen wie auch privaten Impulse, die ich in den vielen Gesprächen in dieser Zeit vermittelt bekam!

Dank gilt ferner allen Freunden und Kollegen, die mir insbesondere in der strukturellen Ausarbeitung dieser Arbeit behilflich waren und mich davor bewahrten, in völlige Verzweiflung über die Tücken von „Word“ und „Excel“ zu geraten.

Für die Mühe der Korrektur- Lesung und für die kontinuierliche Motivation im „Endspurt“ sei ein spezielles Dankeschön an Dr. med. P. Fißl gerichtet.

Das innigste Dankeschön gebührt meiner Familie, meiner Mutter Gabriele Schlotz und meinem Bruder Alexander mit Kerstin, Robin und Julian sowie meinen Freunden, Steffi und Nikki mit Sophia-Marie, Sylvia und Lutz und den „zwei Jakobstälern“ Trudi und Manne, die stets mit großem Verständnis meiner kurzfristigen Zeitplanung begegnen und deren Türen mir immer offen stehen.