

Metaanalysen zur Psychopharmakotherapie: Garbage in – Garbage out?

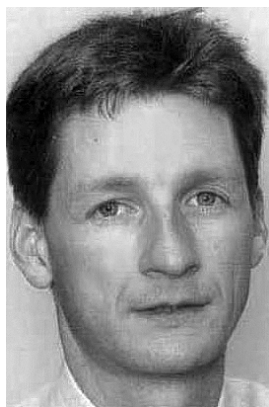
Meta-Analyses in Psychopharmacotherapy: Garbage In – Garbage Out?

Autor

Stefan Weinmann

Institut

Institut für Sozialmedizin, Epidemiologie und Gesundheitsökonomie, Charité – Universitätsmedizin Berlin



Dr. Dr. Stefan Weinmann

Wer Metaanalysen zu wichtigen Themen der medikamentösen Behandlung in der Psychiatrie macht, hat keine geringe Chance, diese auch gut publiziert zu bekommen. Große Zeitschriften nehmen sie seit den 90er-Jahren gerne auf, da davon ausgegangen wird, dass Metaanalysen im besten Fall klinische Expertise mit quantitativen Ergebnissen kombinieren und damit das Fachgebiet bereichern. Sie werden immer und immer wieder zitiert – oft als stärkster oder definitiver Nachweis der Wirksamkeit, Überlegenheit oder Unterlegenheit eines Medikamentes oder einer Medikamentengruppe. Mittlerweile erlernen immer mehr Forscher das Handwerk der Metaanalyse und führen diese auch für alle möglichen Fragestellungen durch, weil sie mit einer Publikation in einem *High-rank*-Journal rechnen. Es ist mittlerweile wenig strittig, dass Metaanalysen für einige Bereiche der Medizin überzeugende und vertrauenswürdige Aussagen liefern [1]. Sie gehen aber weiterhin pauschal als höchste Form der Evidenz in Leitlinien ein [2], ohne dass (im Unterschied zu den Primärstudien) die *Qualität* der Metaanalyse und die Frage nach deren Notwendigkeit und Aussagekraft ausreichend Berücksichtigung findet.

Ergebnisse von Metaanalysen befriedigen das Bedürfnis nach griffigen Zahlen und klaren Aussagen auf der Basis einer vorab definierten Methodik, die international akzeptiert ist und wenig Raum für Manipulation und Verfälschung lässt. Es wird oft argumentiert, dass dem Kliniker oft gar nichts anders übrig bleibt als methodisch hochwertige Informationssynthesen in Form von Metaanalysen von Studien für seine Therapieentscheidung heranzuziehen [3]. Hierbei werden Metaanalysen gelegentlich mit systematischen Reviews gleichgesetzt oder gar systematische Reviews fälschlicherweise als die methodisch am weitesten entwickelte Form narrativer Übersichtsarbeiten betrachtet [4]. Tatsächlich machen Metaanalysen nur auf der Basis einer systemati-

schen Literaturrecherche und a priori festgelegter Bewertungsstrategie auf der Basis eines Protokolls Sinn [5]. Denn besonders die quantitative und statistisch elaborierte Synthese der Ergebnisse von Einzelstudien zu einem einzelnen Effektmaß ist von der Auswahl der eingeschlossenen Studien und der Wahl der Ergebnisparameter abhängig.

Die Vorteile der Metaanalyse in der Psychopharmakotherapie sind aber meist zugleich ihre Nachteile – und mittlerweile haben sich zwei Lager pro und kontra Metaanalysen gebildet, welche die Diskussion zum Stellenwert medikamentöser Behandlung in der Psychiatrie dominieren, uns aber wichtige Aspekte aus dem Auge verlieren lassen. Einige seien hier genannt.

Die Technik der Metaanalysen wurde ursprünglich von Psychologen entwickelt, um eine Synthese von verschiedenen, aber ähnlichen Experimenten zu ermöglichen und einfache Empfehlungen auf der Basis einer heterogenen Sammlung von Einzelergebnissen abzuleiten. Das primäre Ziel der Metaanalyse ist es, im Falle des Vorliegens mehrerer Studien mit kleinerer Patientenzahl deren Einzelergebnisse zu poolen, um eine größere statistische Power zu bekommen. Die Metaanalyse ist daher ein Behelfsinstrument, wenn große methodisch hochwertige Studien mit eindeutigen Ergebnissen fehlen und ist letzteren nur in solchen Fällen vorzuziehen. Die Ergebnisse kleinerer Studien werden so zusammengefasst, als ob sie im Rahmen einer einzigen großen Studie aufgetreten wären. Denn kleinere Studien sind besonders dem Spiel des Zufalls ausgesetzt. Es wird allgemein angenommen, dass diese insbesondere anfällig sind für falsch negative Ergebnisse [6]. Je kleiner der Unterschied zwischen zwei Therapieoptionen, desto größer müsste die Studie sein, um einen solchen nachzuweisen. Extrem große Studien sind daher notwendig zur Entdeckung eines Wirksamkeitsunterschiedes zwischen zwei Antidepressiva [7].

Bibliografie

DOI 10.1055/s-0029-1220425
 Psychiat Prax 2009; 36:
 255–257
 © Georg Thieme Verlag KG
 Stuttgart · New York ·
 ISSN 0303-4259

Korrespondenzadresse

**Dr. med. Dr. P. H.
 Stefan Weinmann**
 Institut für Sozialmedizin,
 Epidemiologie und Gesund-
 heitsökonomie, Charité –
 Universitätsmedizin Berlin
 Luisenstraße 57
 10117 Berlin
 stefan.Weinmann@
 charite.de

Es wird daher angenommen, dass Metaanalysen mehrerer kleinerer Studien die Wirksamkeit von Interventionen aufdecken können, die durch die falsch negativen Ergebnisse der kleinen Studien „verhüllt“ waren. Es wird weiter angenommen, dass falsch positive Ergebnisse wahrscheinlicher sind, wenn eine große Zahl von kleineren Studien ein wirkungsloses Medikament testet (und womöglich nur die mit Positivergebnissen publiziert werden). Wenn alles zufallsbedingt wäre, ein Medikament eine definierte Effektstärke hätte und alle durchgeführten Studien verfügbar wären, könnten Metaanalysen diesen wahren Effekt darstellen. Sie können Zufallseffekte hervorragend ausgleichen – aber nicht systematische Fehler.

Und es ist nicht alles zufallsmäßig. Viele Biostatistiker hofften, dass Metaanalysen eindeutige Methoden objektiver Interpretation von Daten liefern könnten. Diese Hoffnung hat sich bisher in der Psychopharmakotherapie nicht erfüllt. Vielmehr gibt es immer mehr Metaanalysen, die divergente Ergebnisse erbringen, weil sie auf unterschiedlichen Studien(einschlüssen) beruhen, andere statistische Methoden verwenden, Studien, die ganz unterschiedlich sind, gleich oder nicht gleich behandeln oder gar ausschließen, und weil schlussendlich die Statistik aus inadäquaten Daten keine adäquate Synthese machen kann. So können Metaanalysen nicht besser sein als die Güte der qualitativ besten Primärstudien. Und bezüglich der Studienqualität bei Medikamentenstudien in der Psychiatrie sind in den letzten Jahren doch erhebliche Zweifel aufgetreten, die sicherlich ihre Berechtigung haben [8,9]. Genau dieses Problem wird mit dem Begriff „garbage in – garbage out“ bezeichnet: Metaanalysen können zwar manchmal auf die unzureichende Datenqualität hinweisen, diesen Daten aber auch einen objektiven Touch geben – und Kliniker und Entscheidungsträger zu einfachen Schlussfolgerungen verleiten.

Einerseits ist in der Psychopharmakotherapie aufgrund der kleinen Stichprobengrößen der Bedarf an guten Metaanalysen groß (eine Untersuchung von 1941 Studien zur Schizophrenie zeigte beispielsweise, dass nur 3 % groß genug waren, um einen bedeutenden Unterschied zwischen den Interventionen zu zeigen [10]); andererseits ist das Risiko für Verfälschungen hier besonders groß.

Wann machen Metaanalysen Sinn?

Metaanalysen machen Sinn, wenn es eine Reihe von ähnlichen Studien mit kleinen Patientenzahlen gibt, die vergleichsweise konsistente Ergebnisse bringen, aber aufgrund der hohen Einzelvarianzen in den Studien klare Aussagen verhindern. Für diesen „mittleren Bereich“ von Forschung, bei dem adäquate Mengen von relativ hochwertigen Studien verfügbar sind, deren Ergebnisse aber nicht offensichtlich sind, können durchschnittliche Effektstärken mittels statistischer Methoden berechnet und der Einfluss von Studiendesigns, Verzerrungen und anderen Faktoren auf die Ergebnisse herausgearbeitet werden. Damit kann frühzeitig vermieden werden, dass weitere Studien (möglicherweise mit Placebokontrollen) in Bereichen durchgeführt werden, in denen kein Erkenntniszuwachs zu erwarten ist und Patienten Risiken oder unwirksamen Therapien ausgesetzt werden.

Wichtig ist, die Ergebnisparameter sinnvoll zu operationalisieren. Hier macht z. B. die Veränderung des Scores einer validierten und unter echten Doppelblindbedingungen psychiatrischen Skala wie der MADRS oder die Nutzung eines einheitlichen (ab-

soluten) Remissionskriteriums oder der Vergleich von Abbruchraten in Studien Sinn, nicht aber das Poolen von Response-, also Ansprechraten, zumal letztere aufgrund vielfältiger Verfälschungen problematisch sind [9].

Was gibt es zu beachten?

Was aber, wenn große (und nicht industriell finanzierte) Studien den Ergebnissen von Metaanalysen widersprechen? An dieser Stelle muss die CATIE-Studie genannt werden, die einen „harten“ Ergebnisparameter, nämlich die Abbruchraten nach 18 Monaten unter verschiedenen sogenannten atypischen und einem typischen Antipsychotikum, verwendet hat [11]. Man ist geneigt, hinsichtlich der Versorgungsrelevanz dieser großen und trotz einzelner methodischer Probleme bisher besten Studie zu Atypika mehr zu vertrauen als den Metaanalysen (z. B. [12,13]). Warum? Hier zieht das Garbage-in-Argument: Wir haben sowohl in den meisten Antidepressivastudien als auch in den Antipsychotikastudien deutliche Entblindingeffekte, was den (mittels Pharma-Marketing geförderten) Erwartungen der Therapeuten und Patienten ein starkes Gewicht gibt: die Kontrollintervention mit bisher oft zu hohen Dosierungen von Haloperidol verfälscht den Vergleich [14] und lässt die Rate an neurologischen Nebenwirkungen alter Antipsychotika zu hoch oder diejenige der neuen zu niedrig erscheinen [15]. Und die Ermessensspielräume klinischer Einschätzung und Risiken statistischer Auswertung sind bei Abbruchraten geringer als bei den Skalen. So wurden bei Leucht et al. [12] die nicht verblindeten Studien zu Atypika entgegen dem ursprünglichen Protokoll ausgeschlossen, weil sie systematisch die Atypika favorisierten. Vom gleichen Trend ist auszugehen, wenn Entblindungen in den Studien auftreten. Ob man der CATIE-Studie (einer sog. *Effectiveness*-Studie in der Routineversorgung) gerecht wird, wenn man sie in Metaanalysen gleichrangig behandelt wie andere Antipsychotikastudien, bleibt fraglich: eine Qualitätsgewichtung der in die Analysen eingehenden Studien würde hier vermutlich zu anderen Ergebnissen führen, wüsste man, welches echte Qualitätsmerkmale der Studien sind: mit der Randomisierung ist es offensichtlich alleine nicht getan.

Bei Antidepressivastudien spielt der sog. „Wish-bias“ eine Rolle: In einem systematischen Review aller veröffentlichten doppelblinden randomisierten Studien zu Fluoxetin konnte die Gruppe um Barbui zeigen, dass die Wirksamkeit dieses SSRI ganz unterschiedlich war, je nachdem ob das Medikament als zu prüfende Testsubstanz oder als Vergleichssubstanz für andere Antidepressiva zum Einsatz kam [16]. Fluoxetin war als experimentelle Substanz gegen andere Antidepressiva überlegen und als Vergleichssubstanz oft unterlegen. Ein Erklärungsansatz fand sich in der Wahl der Dosierungen: Wenn Fluoxetin als experimentelle Substanz zur Anwendung kam, war die Dosierung deutlich höher als in den Studien, in denen andere Hersteller von Antidepressiva Fluoxetin als Vergleichssubstanz auswählten. Der gleiche Befund zeigte sich in einer Reevaluation [17] einer viel zitierten Studie zur Wirksamkeit von Venlafaxin im Vergleich zu SSRI [18]: Die Remissionsraten waren insgesamt höher in den Studien, in denen Venlafaxin als Testsubstanz verwendet wurde – und nur dann zeigte sich eine signifikante Überlegenheit im Vergleich zu den SSRI.

Überhaupt zeigen erneute Auswertungen und Aktualisierungen von Metaanalysen, dass die Methodik durchaus eine Rolle spielt: die Effektstärke von 1,3 in einer Metaanalyse zu Methylphenidat

bei ADHD im Erwachsenenalter [19] reduzierte sich in einer neueren Auswertung [20] deutlich, da die neuere keine Subgruppenergebnisse verwendete und einen Fehler in der Wahl der Standardabweichung korrigierte. Die methodischen Einflüsse, die durchaus zu unterschiedlichen Ergebnissen führen können, erschließen sich allerdings erst, wenn man in die Detailarbeit einsteigt und die biometrische Auswertung und die Einzeldaten der Metaanalysen hinterfragt. Genau dies kann vom klinischen Leser kaum verlangt werden – wird aber offensichtlich auch vom Peer Review in den Zeitschriften nicht immer geleistet. Metaanalysen sollten daher von unabhängigen Forschergruppen repliziert werden und erfordern im Grunde ein vorher zugänglich gemachtes Protokoll.

Fazit

1. Metaanalysen zur Psychopharmakotherapie machen durchaus Sinn, wenn sie nicht vorgeben, generell die „Wahrheit“ über Effektstärken herauszufinden, sondern uns Hinweise geben, warum Studien zu unterschiedlichen Ergebnissen kommen (Heterogenität), welches die Einflussfaktoren auf die Effektstärken sind, wie sich Ergebnisse ändern, wenn verschiedene Ergebnisparameter oder Auswertungsstrategien verwendet werden. Eine, wenn nicht die entscheidende Frage, die Metaanalysen beantworten könnten, ist die nach der Ursache für Heterogenität: Hier kann der Einfluss der Kontrollbedingung, der Region, des Settings, der Methodik, des Sponsorings und vieles andere herausgearbeitet werden.
2. Die Frage der Qualitätsgewichtung von Studien, die in Metaanalysen einbezogen werden, ist noch nicht gelöst. In jedem Fall sollten Sensitivitätsanalysen prüfen, was passiert, wenn methodisch problematische Studien ausgeschlossen werden. Transparenz ist hier das oberste Prinzip. Im Grunde wird der Einbezug auch unveröffentlichter Studien begrüßt: aber deren Ergebnisse kann der Leser nicht nachvollziehen, wenn der Hersteller die Studie nicht publizieren lässt. Vor dem Hintergrund der Diskussion zu Interessenkonflikten und Ermessensspielräumen in der Auswertung von Studien ist daher bei Metaanalysen, die von pharmazeutischen Herstellern der zu prüfenden Substanz unter Verwendung unpublizierter Studien (*data on file*) finanziert werden, eher Vorsicht geboten.
3. Metaanalysen sollten nur vergleichbare Studien poolen („separate but similar experiments“). Wenn „Apfel“ das interessierende Merkmal ist, sollten nur Apfelstudien, wenn „Obst“ das interessierende Merkmal ist, können Apfel- und Birnenstudien zusammengeworfen werden.
4. Es ist Zeit, Metaanalysen gezielt nur für diejenigen Fragestellungen einzusetzen, für die sie Sinn machen. Die inflationäre Verwendung von Metaanalysen zum Nachweis der „Wirksamkeit“ von Therapien kann unter bestimmten Umständen für Patienten schädlich sein. Zudem wird mit einem starken Fokus auf Metaanalysen die Aufmerksamkeit auf Bereiche (wie die Psychopharmakotherapie) gerichtet, in denen alleine aufgrund der Natur der Intervention ein hohes Maß an Standardisierung möglich ist. Andere werden vernachlässigt. Genau dies passiert mit psychosozialen Interventionen bei schweren psychischen Erkrankungen. Zwar gibt es auch hier gute Evidenz und Metaanalysen – aber vor allem für Interventionen, die standardisierbar, aber in vielen Ländern aufgrund der Rahmenbedingungen nicht generell verfügbar sind

wie Assertive Community Treatment und kognitive Verhaltenstherapie [21]. Wenn wir wie gebannt nur auf Metaanalysen schauen, verlieren wir den Blick für viele relevante Aspekte psychiatrischen Handelns.

5. Wir sollten mehr auf die Qualität von Metaanalysen achten. Genauso wie klinische Studien entsprechend ihrer methodischen Qualität zu unterschiedlichen Ergebnissen kommen, tragen auch Metaanalysen ein Risiko für Bias in sich. Metaanalyseergebnisse müssen repliziert werden.
6. Wie bei den Studien sollte mehr Sorgfalt auf die klinische Signifikanz der in Metaanalysen berechneten Effektstärken gelegt werden – statistische Signifikanz ist nicht alles. Wir brauchen sicherlich mehr gute systematische Reviews, aber vielleicht auch weniger Metaanalysen!

Literatur

- 1 Smith GD. Meta-Analysis. Potentials and promise. *BMJ* 1997; 315 (7119): 1371–1374
- 2 Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ* 2001; 323 (7308): 334–336
- 3 Komossa K, Kissling W, Leucht S et al. Therapieentscheidung aufgrund von Metaanalysen. *Psychiatr Prax* 2008; 35 (8): 373–375
- 4 Maier W, Möller HJ. Metaanalysen: Methoden zur Evidenzmaximierung von Therapiestudien? *Nervenarzt* 2007; 78 (9): 1028–1036
- 5 Egger M, Smith GD, Sterne JA. Uses and abuses of meta-analysis. *Clin Med* 2001; 1 (6): 478–484
- 6 Freiman JA, Chalmers TC, Smith H Jr et al. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 „negative“ trials. *N Engl J Med* 1978; 299 (13): 690–694
- 7 Lieberman JA, Greenhouse J, Hamer RM et al. Comparing the effects of antidepressants: consensus guidelines for evaluating quantitative reviews of antidepressant efficacy. *Neuropsychopharmacology* 2005; 30 (3): 445–460
- 8 Leucht S, Heres S, Hamann J et al. Methodological issues in current antipsychotic drug trials. *Schizophr Bull* 2008; 34 (2): 275–285
- 9 Moncrieff J, Kirsch I. Efficacy of antidepressants in adults. *BMJ* 2005; 331 (7509): 155–157
- 10 Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ* 1998; 317 (7167): 1181–1184
- 11 Lieberman JA, Stroup TS, McEvoy JP et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med* 2005; 353 (12): 1209–1223
- 12 Leucht S, Corves C, Arbter D et al. Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *Lancet* 2009; 373 (9657): 31–41
- 13 Leucht S, Barnes TR, Kissling W et al. Relapse prevention in schizophrenia with new-generation antipsychotics: a systematic review and exploratory meta-analysis of randomized, controlled trials. *Am J Psychiatry* 2003; 160 (7): 1209–1222
- 14 Geddes J, Freemantle N, Harrison P et al. Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *BMJ* 2000; 321 (7273): 1371–1376
- 15 Miller DD, Caroff SN, Davis SM et al. Extrapyramidal side-effects of antipsychotics in a randomised trial. *Br J Psychiatry* 2008; 193 (4): 279–288
- 16 Barbui C, Cipriani A, Brambilla P et al. „Wish bias“ in antidepressant drug trials? *J Clin Psychopharmacol* 2004; 24 (2): 126–130
- 17 Weinmann S, Becker T, Koesters M. Re-evaluation of the efficacy and tolerability of venlafaxine vs SSRI: meta-analysis. *Psychopharmacology (Berl)* 2008; 196 (4): 511–520
- 18 Smith D, Dempster C, Glanville J et al. Efficacy and tolerability of venlafaxine compared with selective serotonin reuptake inhibitors and other antidepressants: a meta-analysis. *Br J Psychiatry* 2002; 180: 396–404
- 19 Faraone SV, Spencer T, Aleardi M et al. Meta-analysis of the efficacy of methylphenidate for treating adult attention-deficit/hyperactivity disorder. *J Clin Psychopharmacol* 2004; 24 (1): 24–29
- 20 Koesters M, Becker T, Kilian R et al. Limits of meta-analysis: methylphenidate in the treatment of adult attention-deficit hyperactivity disorder. *J Psychopharmacol* 2008 DOI: 10.1177/0269881108092338
- 21 Riedel-Heller S, Richter D. Psychosoziale Interventionen und soziale Inklusion. *Psychiatr Prax* 2008; 35 (5): 213–215