

Jahrgang 14

Nr. 14

28. Juli 1992

Meta-Analysen (M. Egger & G. Davey Smith)..... 53

Meta-Analysen haben in letzter Zeit stark an Bedeutung gewonnen. Dabei werden die Daten verschiedener Studien vereinigt, um aussagekräftigere und präzisere Ergebnisse zu erhalten. Meta-Analysen können es ermöglichen, kleine, aber klinisch relevante Effekte einer Therapie zu erfassen. Die Qualität einer Meta-Analyse hängt stark von der Auswahl der Studien ab. Die Wahl der geeigneten statistischen Methoden ist ebenfalls wichtig, aber doch von sekundärer Bedeutung. Meta-Analysen müssen mit einem kritischen klinischen Blick betrachtet werden, bevor ihre Resultate in die Praxis übertragen werden können.

Übersicht

Meta-Analysen

M. Egger & G. Davey Smith

Manuskript durchgesehen von U. Helfenstein, M.M. Kochen,
M. Notter & J. Schmidt

Meta-Analysen klinischer Studien werden in den letzten Jahren immer häufiger durchgeführt. Dabei werden die Ergebnisse ähnlicher Studien zusammengefasst, um statistisch aussagekräftigere und für eine grössere Patientengruppe gültige Angaben über die Wirksamkeit einer Therapie zu erhalten.

Die erste «medizinische» Meta-Analyse wurde 1955 veröffentlicht.¹ Aus 15 Studien, in denen die behandelten Personen für ganz verschiedene Beschwerden ein Placebo erhalten hatten, wurde die durchschnittliche Wirksamkeit des Placebos ermittelt: bei 35% der insgesamt über tausend Patienten hatte das Placebo scheinbar die erwünschte Wirkung erbracht.

Mit der Entwicklung von geeigneteren statistischen Methoden haben Meta-Analysen seit dem Anfang der achtziger Jahre zunehmend an Bedeutung gewonnen. Der meta-analytische Ansatz blieb jedoch von heftiger Kritik nicht verschont. Während die einen die Meta-Analyse als «objektive, quantitative Methode» rühmen, wird das Verfahren von anderen als «statistischer Trick» bezeichnet, der «ungerechtfertigte Annahmen macht und zu unzulässigen Verallgemeinerungen führt».²

Warum meta-analysieren?

Oft sind Studien nicht gross genug, um kleinere, aber klinisch bedeutende Unterschiede zwischen zwei Therapien sicher und präzise zu erfassen. Eine Studie kann *falsch negativ* ausfallen, das heisst keinen Unterschied aufzeigen, wenn in Wirklichkeit ein solcher besteht. In diesem Fall spricht man von einem *Fehler der zweiten Art*, dessen Wahrscheinlichkeit (β) berechnet werden kann. Eine Auswertung von 71 «negativen» randomisierten Therapiestudien hat gezeigt, dass ein klinisch relevanter Unterschied zwischen Kontrolltherapie und experimenteller Therapie selten mit Sicherheit ausgeschlossen werden kann: β betrug in 62 der 71 Studien mehr als 20%.³ Besser bekannt ist der *Fehler der ersten Art*: Die Resultate einer Studie können einen Unterschied zeigen, der ausschliesslich durch den Zufall zustande gekommen ist und keinem echten (biologischen) Unterschied entspricht. Diese Wahrscheinlichkeit (α) wird mit dem p-Wert ausgedrückt; liegt α unter 5% ($p < 0,05$), so spricht man in der Regel von einem *signifikanten* Resultat.

Ein Beispiel: Ein Medikament, das die Sterblichkeit nach Myokardinfarkt um 10% reduziert, könnte allein in der Schweiz jedes Jahr Hunderten von Patienten das Leben retten. Wieviele Patienten mit Herzinfarkt müssten in eine Studie aufgenommen werden, um einen solchen Effekt mit einer Sicherheit von 90%, das heisst mit einem β von 10%, zu erfassen? Berechnungen ergeben, dass hierfür fünfstellige Patientenzahlen notwendig wären. Als Alternative zu derart grossen, teuren, und praktisch oft kaum durchführbaren Studien bietet sich die Meta-Analyse an. Dabei werden alle Personen, die in ähnlichen, kleineren Studien mit demselben Medikament behandelt wurden, meta-analytisch zusammengefasst. So werden die notwendigen grossen Patientenzahlen erreicht und es kann gelin-

gen, «kleinere» Wirkungen mit einiger Zuverlässigkeit festzustellen oder auszuschliessen.

Während eine Einzelstudie nur Schlüsse auf eine verhältnismässig eng umschriebene Auswahl von Patienten zulässt, lassen sich die Ergebnisse von korrekt durchgeführten Meta-Analysen oft auf einen grösseren Personenkreis anwenden. Meta-Analysen ergeben zudem Hinweise auf den in Zukunft am meisten Erfolg versprechenden oder auf den dringendsten Forschungsansatz. Sie erlauben auch eine genauere Berechnung der dafür notwendigen Anzahl Patienten. Im Vergleich mit konventionellen «Übersichten» zeichnen sich Meta-Analysen durch die Anwendung von quantitativen Methoden und damit möglicherweise durch eine objektivere Beurteilung der publizierten Daten aus.

Prinzipien der Meta-Analyse

Wie jeder Studie muss auch der Meta-Analyse eine klare Definition der zu untersuchenden Fragestellung vorangehen. Besonders wichtig ist die Definition von Einschlusskriterien für die zu berücksichtigenden Studien. Die Suche nach relevanten Studien erfolgt heute am besten mit Hilfe der elektronischen Datenbanken. Computerdatenbanken allein identifizieren jedoch weniger als zwei Drittel der relevanten Studien, weshalb zusätzlich andere Quellen («Current Contents», Übersichtsartikel, Standardwerke, Experten, Register von unpublizierten Studien usw.) herangezogen werden müssen.⁴ Jede Studie muss dann in allen Einzelheiten analysiert werden; besonders wichtig sind Studienanlage, Therapieschema, Patientencharakteristika und Erfassung des Therapieerfolgs.

Sodann geht es darum, die Resultate von Studien, welche die Einschlusskriterien erfüllen, in eine standardisierte Form zu bringen, damit ein Vergleich zwischen den Studien möglich wird. Zu diesem Zweck wird oft ein Wert berechnet, der mit «Odds Ratio» bezeichnet wird. Der Begriff «Odds» lässt sich deutsch ungefähr mit «Chancen» wiedergeben. Ein Beispiel: Wenn die Chancen 4 zu 6 stehen, dass innerhalb einer bestimmten Zeit unter einem Antibiotikum Durchfall auftritt, so entspricht dies einem «Odds»-Wert von 4/6. Die «Odds Ratio» (für Durchfall) berechnet sich dann als Quotient aus den «Odds» in der Antibiotika-Gruppe (4/6) und den «Odds» in einer Vergleichsgruppe ohne das betreffende Antibiotikum (z.B. 1/9). Die «Odds Ratio» beträgt hier 6: Die «Chance», an Durchfall zu erkranken, ist in der Antibiotikagruppe 6mal grösser als in der Vergleichsgruppe. Für die «Odds» und die «Odds Ratio» geht man also von einem bestimmten Ereignis aus.

Im Gegensatz dazu errechnet sich das *Risiko* eines bestimmten Effektes aus dem Verhältnis der Betroffenen zur Gesamtheit der Behandelten. In unserem Beispiel wäre daher das Risiko, unter dem Antibiotikum an Durchfall zu erkranken, 4/10 oder 40%. In der Kontrollgruppe ergäbe sich ein Risiko von 1/10 oder 10%. Das relative

Risiko, unter dem Antibiotikum an Durchfall zu erkranken, ist genau 4mal grösser als ohne Antibiotikum. Hier geht man nicht vom Ereignis, sondern von der Population aus.

Besteht *kein Unterschied* zwischen zwei Vergleichsgruppen, so beträgt der Wert für die «Odds Ratio» und das relative Risiko genau 1. Solange das untersuchte Ereignis relativ selten auftritt, ergeben «Odds Ratio» und relatives Risiko ähnliche Werte. Tritt das Ereignis aber häufiger auf, so ergibt sich – wie das Beispiel gezeigt hat – für die «Odds Ratio» ein deutlich grösserer Wert als für das relative Risiko. Im Vergleich mit dem relativen Risiko werden also vor- oder nachteilige Ereignisse mit der «Odds Ratio» überschätzt.

Der grosse Vorteil der «Odds Ratio» für Meta-Analysen liegt aber darin, dass damit mathematisch einfach ein Summationswert aller Studien berechnet und auf seine statistische Signifikanz hin getestet werden kann. Die «Odds Ratios» werden zusammen mit dem zugehörigen *Vertrauensintervall* graphisch dargestellt. Die Ergebnisse von Studien werden bekanntlich durch das Spiel des Zufalls beeinflusst. Das Vertrauensintervall zeigt den Bereich, in dem der wahre Therapieeffekt mit einer gegebenen Wahrscheinlichkeit liegen muss. Das gebräuchliche Vertrauensintervall von 95% gibt den Bereich an, in dem der wahre Therapieeffekt mit einer Wahrscheinlichkeit von 95% zu finden ist. Wenn das 95%-Vertrauensintervall den Wert 1 einschliesst, dann ist der Unterschied in der Wirksamkeit der beiden Therapien statistisch nicht signifikant, das heisst, p ist grösser als 0,05.

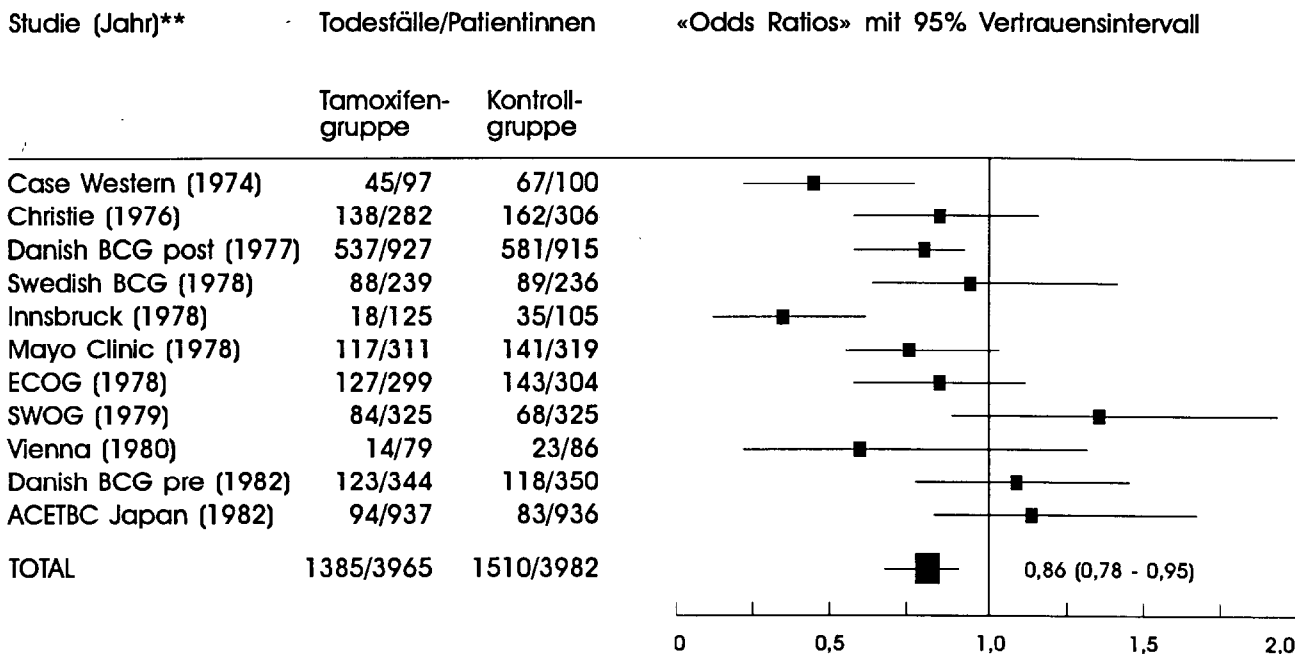
Der letzte Schritt besteht in der Berechnung des *Summationswertes* mit Vertrauensintervall, in dem alle Studien vereint werden. Dafür stehen verschiedene statistische Methoden zur Verfügung, die jedoch meistens zu sehr ähnlichen Resultaten führen.⁵ Allgemein werden dabei Studien mit grösseren Patientenzahlen, die mehr Information beisteuern, stärker gewichtet als kleine Studien. Dieser Ansatz ist offensichtlich der Berechnung eines einfachen arithmetischen Mittelwertes überlegen.

Ein Beispiel: Tamoxifen bei Mammakarzinom

Vor kurzem wurden die Ergebnisse einer grossangelegten Meta-Analyse aus der Onkologie veröffentlicht.⁶ Es ging um die Bedeutung der *adjuvanten Therapie* nach der chirurgischen Entfernung des Primärtumors bei Mammakarzinom ohne Fernmetastasen. Insgesamt wurden 133 randomisierte Studien analysiert, in denen die Wirkung von verschiedenen systemischen Therapien (Hormon-, Chemo- oder Immuntherapie) geprüft worden war.

Das folgende Beispiel greift einen Teilaspekt dieser Meta-Analyse heraus. Es betrifft Studien, in denen die Wirkung einer einjährigen Therapie mit dem Anti-Östrogen Tamoxifen (Nolvadex® u.a.) untersucht worden war. In sieben Studien war die Kombination *Chemotherapie/Tamoxi-*

Abbildung: Tamoxifenbehandlung des Mammakarzinoms (1 Jahr)*



* Modifiziert nach ⁶

** Die genauen Bezeichnungen und Autoren der Studien finden sich in der Originalarbeit.⁶

mit *Chemotherapie* allein und in vier Studien *Tamoxifen* mit *keiner* adjuvanten Therapie verglichen worden.

Die Abbildung auf dieser Seite zeigt für die verschiedenen Studien die Sterblichkeit als «Odds Ratio» mit einem Vertrauensintervall von 95%. Acht Studien ergaben keinen statistisch signifikanten Effekt (die Vertrauensintervalle enthalten 1). Nur drei Studien («Case Western», «Danish BCG post», «Innsbruck») zeigten eine der Kontrolltherapie signifikant überlegene Wirkung von Tamoxifen («Odds Ratio» unter 1, die 95%-Vertrauensintervalle enthalten den Wert 1 nicht).

Man könnte nun leicht zum Schluss kommen, dass eine einjährige adjuvante Behandlung mit Tamoxifen keine Vorteile bietet, da die überwiegende Mehrheit der Studien keinen eindeutigen Effekt ergaben. Eine Meta-Analyse aller Studien zeigt jedoch eine signifikante Wirkung mit einer «Odds Ratio» von 0,86. Die einjährige Therapie mit Tamoxifen kurz nach Diagnosestellung hat somit die «Odds», in den folgenden Jahren an einem metastasierendem Mammakarzinom zu sterben, um 14% reduziert. Dies entspricht einer Reduktion des Risikos um etwa 10%, eine kleine, aber dennoch wichtige Verbesserung der Therapie des Mammakarzinoms. Eine noch bessere Wirkung (Reduktion der Sterblichkeit um etwa 20%) wird durch eine länger dauernde Tamoxifen-Therapie erreicht.⁶

Meta-Analysen von Tamoxifen-Studien haben schon vor Jahren zu einer Änderung der therapeutischen Richtlinien geführt, die ohne eine derartige quantitative Auswer-

tung aller Studienresultate kaum zustande gekommen wäre.⁷ Die Richtigkeit dieser Entscheidung wird durch die neueren Daten klar bestätigt.

Probleme der Meta-Analysen

Obwohl Meta-Analysen somit die Erfassung von «kleinen» Wirkungen zu einem frühen Zeitpunkt ermöglichen können, sollte diese Methode keinesfalls unkritisch angewendet werden. Wie jede einzelne Studie, muss auch eine Meta-Analyse mehrerer Studien auf ihre Qualität hin untersucht werden.⁸⁻¹⁰

Die wichtigsten Punkte, die in diesem Zusammenhang überprüft werden müssen, lassen sich kurz folgendermaßen beschreiben:

Ist die Qualität der berücksichtigten Studien zufriedenstellend?

Zwischen verschiedenen Studien bestehen zum Teil erhebliche Qualitätsunterschiede, die in einer Meta-Analyse zunächst nicht berücksichtigt werden. Als Mindestanforderung gilt, dass nur korrekt *randomisierte Studien* mit vollständigen Angaben über alle am Anfang in die Studie aufgenommenen Personen meta-analysiert werden. Der Randomisierungsvorgang, der leider oft ungenügend beschrieben ist, verdient deshalb besondere Aufmerksamkeit.

Daneben ist es auch von Bedeutung, dass in den verschiedenen Studien tatsächlich gleiche oder vergleichbare Therapien und Beurteilungsverfahren eingesetzt wurden. Bei

vielen Medikamentenstudien muss z.B. gefordert werden, dass die Bewertung der Therapieergebnisse *blind* erfolgt.

Wurden alle relevanten Studien berücksichtigt?

Da «positive» Resultate mit grösserer Wahrscheinlichkeit veröffentlicht werden als Studien, die keinen Effekt zeigen,¹¹ sollten auch unpublizierte Studien berücksichtigt werden. Dies gilt natürlich nur, soweit die erwähnten Qualitätskriterien erfüllt sind. Wenn allenfalls in Frage kommende Studien nicht in eine Meta-Analyse aufgenommen werden, so sollten sie erwähnt und der entsprechende Entscheid begründet werden.

Haben die berücksichtigten Studien teilweise entgegengesetzte Resultate erbracht?

Der Variabilität der Ergebnisse wird oft nicht genügend Beachtung geschenkt. Das Problem der Heterogenität von Studienresultaten kann nicht durch die Anwendung eines statistischen Tests gelöst werden. Wenn stark voneinander abweichende Resultate vorliegen, geben klinische und biologische Überlegungen den Ausschlag, ob überhaupt eine Meta-Analyse sinnvoll ist.

Wie «robust» sind die Ergebnisse der Meta-Analyse?

Bei der Durchführung einer Meta-Analyse gibt es eine Reihe von teilweise *arbiträren* Entscheiden (Ausschluss von Studien, Wahl der statistischen Methoden, Interpretation bestimmter Resultate). Das Resultat einer Meta-Analyse sollte von solchen Entscheiden einigermaßen unabhängig sein. Wenn also die gleiche Meta-Analyse z.B. mit einer etwas anderen Studienausswahl oder mit einem anderen statistischen Verfahren durchgeführt wird, sollte sie ungefähr dasselbe Resultat ergeben. Unser Beispiel (Meta-Analyse der Tamoxifen-Studien) unterscheidet sich so von der publizierten Analyse⁶ und hat dennoch ein ganz ähnliches Ergebnis erbracht.

Bedeutung für Praxis und Klinik

Die Anwendung der Ergebnisse von Meta-Analysen in der klinischen Praxis kann aus verschiedenen Gründen problematisch sein.¹² «Durchschnittspatienten», die in klinische Studien aufgenommen werden, sind oft männlich und 40 bis 60 Jahre alt. In der primärärztlichen Praxis sind dagegen Frauen und ältere Menschen stark vertreten. Ähnlich können die Resultate von Studien, die mehrheitlich in spezialisierten Zentren durchgeführt wurden, nicht unbedingt auf das durchschnittliche Spital übertragen werden. Schliesslich ist die klinische Praxis einer raschen Entwicklung unterworfen. Studien, die einige Jahre zurückliegen, sind unter Umständen heute von begrenzter Relevanz, können jedoch in einer Meta-Analyse dominieren.

Meta-Analysen können uns Hinweise auf wichtige therapeutische Entwicklungen vermitteln. Ob ihre Resultate aber für unsere eigenen Patientinnen und Patienten relevant sind, müssen wir in jedem Einzelfall sorgfältig prüfen.

Literatur

- 1 Beecher HK. JAMA 1955; 159: 1602-6
- 2 Thompson SG, Pocock SJ. Lancet 1991; 338: 1127-30
- 3 Freimann JA et al. N Engl J Med 1978; 299: 690-4
- 4 Dickersin K et al. Cont Clin Trials 1985; &: 306-17
- 5 Berlin JA et al. Statistics Med 1989; 8: 141-51
- 6 Early Breast Cancer Trialists' Collaborative Group. Lancet 1992; 339: 1-15
- 7 Consensus Conference. JAMA 1985; 254: 3461-3
- 8 Sacks HS et al. N Engl J Med 1987; 316: 450-5
- 9 Hartzema AG. Ann Pharmacother 1992; 26: 96-7
- 10 Kassirer JR. N Engl J Med 1992; 327: 273-4
- 11 Dickersin K et al. JAMA 1992; 267: 374-8
- 12 Wittes RE. Statistics Med 1987; 6: 281-6

Mitarbeiter dieser Ausgabe

Dr. M. Egger, Department of Epidemiology and Public Health, University College London, London WC1E 6EA (England)
Dr. G. Davey Smith, Department of Public Health, University of Glasgow, Glasgow G12 8RZ (Scotland)
Dr. U. Helfenstein, Institut für Sozial- und Präventivmedizin der Universität, CH-8006 Zürich
Prof. Dr. M.M. Kochen, Abteilung für Allgemeinmedizin, Georg-August-Universität, D-3400 Göttingen
Dr. M. Notter, Strahlentherapie, Kantonsspital, CH-5001 Aarau
Dr. J. Schmidt, Furrenmatte 4, CH-8840 Einsiedeln

Korrigendum

In meinem «ceterum censeo» zum Thema «Amerika: Hast du es besser?» (pharma-kritik 1992; 14: 43-4) ist der Kommentar zur Information über Ceftriaxon (Rocephin®) missverständlich. Man könnte meinen, im schweizerischen Arzneimittel-Kompendium finde sich kein Hinweis auf die Möglichkeit einer pseudomembranösen Kolitis. Dies trifft nicht zu. Auch in der Schweiz wird darauf hingewiesen, dass die pseudomembranöse Enterokolitis als «sehr seltene» Nebenerscheinung «mitgeteilt» worden sei. Im Unterschied dazu enthält jedoch die entsprechende Information in den USA eine ausführliche, eindringliche Warnung vor der Kolitis-Gefahr.

Etzel Gysling

pharma-kritik

Herausgegeben von Etzel Gysling (Wil)
unter Mitarbeit von Renato Galeazzi (St. Gallen) & Urs A. Meyer (Basel)
Redaktion: Marianne Beutler, Peter Ritzmann
Externe redaktionelle Mitarbeiter: Benedikt Holzer, Michael M. Kochen, Peter Koller, Eva Maurer, Johannes Schmidt
Verlagsmitarbeiterin: Judith Künzler

Bezugspreise: Jahresabonnement Fr. 86.- (Studenten Fr. 43.-),
Zweijahresabonnement Fr. 142.-, Einzelnummer Fr. 7.-
Infomed-Verlags-AG, Postfach 528, 9500 Wil 1, Telefon (073) 22 18 18
© 1992 Etzel Gysling Wil. All rights reserved.