

Evidenzbasierte Medizin und Psychotherapie: die Frage nach den angemessenen Erkenntnismethoden

Norbert Schmacke

Evidence-Based Medicine and Psychotherapy: What are Adequate Research Methods?

Zusammenfassung

Die evidenzbasierte Medizin (EbM) hat sowohl als Methode zur systematischen Bewertung von Nutzen und Risiken als auch als Entscheidungshilfe für die Praxis im vergangenen Jahrzehnt international große Verbreitung gefunden. Als Goldstandard in der Therapieforschung hat sich die randomisierte, kontrollierte Studie (RCT) etabliert. Diese Ansätze stoßen in der medizinischen Fachwelt gleichwohl nach wie vor auf Skepsis, weil befürchtet wird, dass RCTs das Feld der Versorgung nicht ausreichend genau abbilden können. Diese Bedenken sind insbesondere in Psychiatrie und Psychotherapie stark vertreten. Der Artikel entwickelt auf zwei Ebenen eine die EbM unterstützende Argumentation: weder lassen sich Gründe finden, warum RCTs für die Verbesserung der Versorgung psychisch Kranker nicht nutzenstiftend sein können, noch erschöpft sich relevante Versorgungsforschung in den unverzichtbaren RCTs. Die Bedeutung qualitativer Methoden ist eine mindestens so gewichtige Black Box der Gesundheitsforschung wie der sinnhafte Einsatz von RCTs.

Abstract

Evidence-based Medicine (EbM) was accepted during the last decade both as a method for systematically evaluating risks and benefits and a decision aid for therapists in practice. Randomized controlled trials (RCT) have become the gold standard for medical research. The medical community is sceptical about RCTs because many therapists are not convinced that care can be mirrored adequately enough, though. The doubts are especially strong among psychiatrists and psychotherapists. The paper argues in favour of EbM in two ways. Firstly, from a scientific point of view there are no reasons why RCTs could not contribute to improve the standards of care. Secondly, health services research doesn't amount to much more than RCTs. The importance of qualitative methods is a black box as large within health services research as the meaningful use of RCTs.

Key words

Evidence based medicine · psychotherapy · qualitative research

Scientists will be denounced for trying to introduce cold-blooded reason into an area in which warm-blooded humanity is supposed to reign supreme. But warm emotion frequently gives way to hot-headed anger and even bigotry [1].

Wenn man sich in der Tradition der evidenzbasierten Medizin (EbM) auf den Weg macht, nach „belastbaren Daten“ für medizinische Empfehlungen zu suchen, dann sieht man sich mit einer Fülle von Argumenten konfrontiert, warum insbesondere in der Psychotherapie der Ansatz der EbM zu kurz greife. Die meisten dieser Argumente tragen defensive Züge: sie wollen häufig im-

Institutsangaben

Fachbereich Human- und Gesundheitswissenschaften, Universität Bremen

Korrespondenzadresse

Prof. Dr. med. Norbert Schmacke · Arbeits- und Koordinierungsstelle Gesundheitsversorgungsforschung · Fachbereich Human- und Gesundheitswissenschaften · Universität Bremen · Bibliothekstraße 1 · 28359 Bremen · E-mail: Schmacke@uni-bremen.de

Eingegangen: 20. Oktober 2005 · **Angenommen:** 5. Dezember 2005

Bibliografie

Psychother Psych Med 2006; 56: 202 – 209 © Georg Thieme Verlag KG Stuttgart · New York
DOI 10.1055/s-2006-932609
ISSN 0937-2032

mer noch den Zunftcharakter des jeweiligen Therapieverfahrens schützen. Das Schulendenken ist in Psychiatrie, Psychotherapie und somatischer Medizin in der alltäglichen Praxis tief verankert, auch wenn die Aufarbeitung differenzieller therapeutischer Effekte die Rigidität der traditionellen Schulen schon lange infrage gestellt und eher das Thema auf die Tagesordnung gesetzt hat, in welcher Weise bei einem hohen Maß universeller Wirkmechanismen dann doch unterschiedliche Verfahren für spezifische therapeutische Aufgaben etwa bezüglich Qualifikation der Therapeuten oder notwendiger Therapiedauer prädestiniert sind [2]. Metaanalysen schwanken dabei offenkundig noch zu stark zwischen der Überzeugung, die Überlegenheit spezifischer Therapieverfahren für Krankheitsentitäten nachweisen zu können (so z.B. Leichsenring [3]) und einer zurückhaltenderen Interpretation (so z.B. Reisner), die zwar nicht dem „Common Factors Model“ alle Erklärungskraft zuweist, aber doch auf die überragende Bedeutung allgemeiner therapeutischer Prinzipien wie hoher Empathie hinweist [4]. Es bleibt im Feld der Psychiatrie und Psychotherapie ein enormes Defizit an versorgungsrelevanter Forschung zu beklagen: so fehlen Studien zum Vergleich unterschiedlicher Behandlungsmethoden bzw. -konzepte mit relevanten klinischen Endpunkten und ausreichender Laufzeit. Bei einer Vielzahl von methodisch prinzipiell akzeptablen randomisierten, kontrollierten Studien (RCTs) im Bereich der Pharmakotherapie mangelt es – anders gewendet – ebenfalls sowohl an angemessen differierten Interventionen (häufig fehlt in Arzneimittelstudien als Kontrolle der gültige Standard, noch häufiger aber mangelt es an nicht-medikamentösen Interventionen) wie an klinisch relevanten Endpunkten und ausreichend langen Beobachtungszeiträumen. Es gibt insofern verständliche Ressentiments gegenüber dem Goldstandard des Forschungsansatzes der EbM, den RCTs. Dies mag bei der Adressierung psychischer Erkrankungen bzw. Probleme vielleicht auch noch etwas anders begründet sein als in der Somatik, da zum einen unterschiedliche Traditionen und überkommene gesellschaftliche Rollenzuweisungen an Ärzte und Psychotherapeuten eine große Bedeutung haben und zum anderen die Vorstellung des Gebots der Individualisierung jedweder Therapie bei psychischen Krankheiten und Störungen noch stärker verwurzelt sein könnte als dies in der Körpermedizin der Fall ist. Dass andere Untersuchungsdesigns, insbesondere Beobachtungsstudien, in der Medizin zu gravierenden Fehleinschätzungen geführt haben, hat Mühlhauser in dieser Ausgabe eindrucksvoll belegt. Es kann am Ende, so die Kernthese, gar nicht um die Frage gehen, ob (methodisch hochwertige) RCTs zur Ermittlung eines therapeutischen Fortschritts notwendig sind. Es geht vielmehr darum, wie darüber hinaus das Spektrum angemessener Methoden zur Weiterentwicklung eines wissenschaftlich begründbaren und sozial akzeptablen Versorgungssystems aussehen muss.

Ist EbM Kochbuchmedizin?

In der Schlichtversion der Kritik wird EbM als Kochbuchmedizin an den Pranger gestellt, die das einzelne Patientenschicksal nicht mehr kennt. Die Gegenüberstellung von individuellen Therapiezielen und EbM ist aber grundsätzlich unsinnig. Ehe ein Therapeut mit seinem Patienten ein individuelles Therapieziel vereinbaren kann, muss er wissen, für welche Ansätze es Belege für eine akzeptable Nutzen-Risiko-Bilanz gibt. Von einem guten

Standard kann ein Therapeut nur begründet abweichen, wenn er diesen kennt – oder er kann realisieren, dass es einen durch hochwertige Studien ermittelten Standard für seine Fragestellung noch nicht gibt – oder vielleicht auch nicht geben kann. Beide Extreme setzen aber voraus, dass eine systematische Recherche der vorliegenden Evidenz und deren kritische Bewertung stattgefunden hat [5,6]. Man könnte salopp auch entgegnen: ein gutes Kochbuch ist sehr hilfreich, um ein gutes Menü zu produzieren – aber jeder Koch wird sich nicht sklavisch an jedes Detail aufgeschriebener Rezepte halten wollen und können. Auch der Versuch, gegenüber der Orientierung auf gute Standards die Heilkunde als Kunst zu kennzeichnen, wirkt nicht wesentlich überzeugender: Kunst setzt Können voraus, und EbM ist in praxi natürlich immer „nur“ Sichtung wie Bewertung von Wissensbeständen für eine konkrete Behandlungssituation. Handeln ist überspitzt gesagt auch im Zeitalter der EbM nicht zu ersetzen.

Ist EbM unethisch?

Ein weiteres Standardargument gegen die EbM lautet: RCTs sind entweder gar nicht durchführbar oder ethisch unverantwortbar. So scheint es in einem weiten Feld, nämlich chirurgischen Eingriffen, auf den ersten Blick tatsächlich hoch problematisch zu sein, eine Kontrollgruppe einem Scheineingriff (Plazebo) zu unterziehen, um den Nutzen eines vielleicht sogar plausiblen Operationsverfahrens methodisch hochwertig zu ermitteln. Eine Studie aus dem Jahr 2002 [7] zeigt, wie vorsichtig man mit dem Ethikargument sein sollte. In der Orthopädie hatte sich seit den 90er-Jahren mit der weiten Verbreitung der Spiegelung von Kniegelenken eingebürgert, nicht nur defekte Meniski zu entfernen, sondern immer dann zeitgleich mit dem diagnostischen Eingriff eine „Glättung“ des Knieinnenraums vorzunehmen, wenn der Untersucher zwar keinen Meniskusschaden im eigentlichen Sinne feststellte aber deutliche degenerative Veränderungen am Knorpel. Man ging davon aus, dass dadurch das Ausmaß beklagter Beschwerden bei Bewegung deutlich abnehmen müsste: das Gelenk war doch wieder freier geworden. Eine RCT, bei der die Kontrollgruppe lediglich einen Hautschnitt gleicher Art wie die Interventionsgruppe erhielt, kam zu völlig überraschenden Ergebnissen: 24 Monate nach dem Eingriff war die Interventionsgruppe eben doch nicht beschwerdefreier als die Plazebogruppe. Dieses Beispiel ist in vielfacher Hinsicht lehrreich. Es zeigt zunächst, wie neue Verfahren zur Routine werden können, ohne dass jemals nach der Nutzen-Risiko-Bilanz gefragt worden ist. Die Endoskopie an sich (Spiegelung) wurde zu einer ohne weiteres ambulant durchführbaren Methode, in der Hand des Geübten mit geringen Risiken behaftet (natürlich nicht risikolos), und es entspricht dem gesunden Menschenverstand, bei Kniebeschwerden auch ohne Hinweise auf Meniskusschäden in das Knie hineinzuschauen („ist bestimmt besser als Anamnese und/oder Röntgen“) und den Knorpel zu glätten, wenn Abfaserungen zu sehen sind („dann reibt nichts mehr aneinander“). Je höher die Dichte an niedergelassenen Orthopäden und Chirurgen, umso höher die Zahl der „explorativ-therapeutischen“ Kniegelenkspiegelungen. Dass diesen keine ausreichende Wirksamkeit zugesprochen werden kann, wäre ohne die genannte RCT niemals herausgekommen. Ist es ethisch verantwortbar, Menschen invasiven Behandlungsmethoden auszusetzen, ohne dafür gute Nutzenbelege zu haben? Möglicherweise wären sogar an-

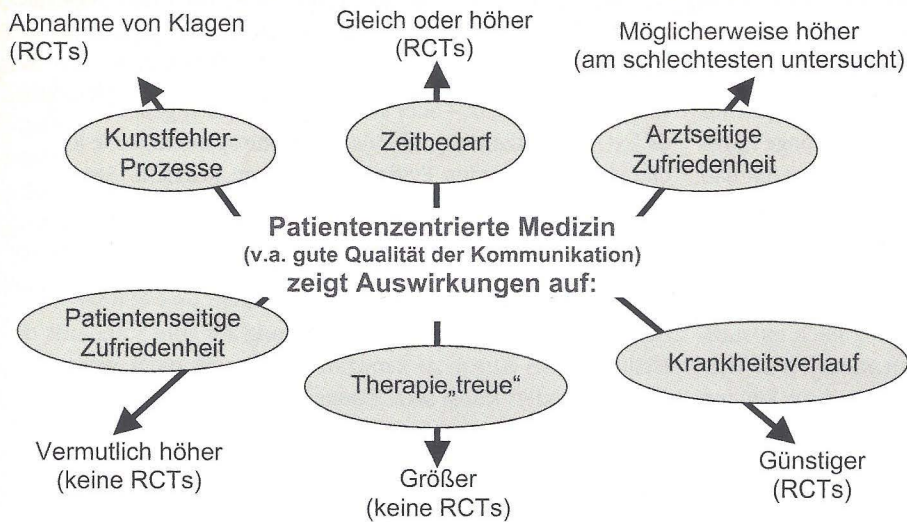


Abb. 1 Studienlage zum Outcome patientenzentrierter Medizin (Darstellung Schmacke nach Stewart et al. [8]).

derer Studiendesigns zur Bewertung dieses Behandlungsansatzes denkbar gewesen. Das Beispiel verdeutlicht aber unmissverständlich, wie wichtig es ist, nach der Entstehung von Behandlungsmethoden zu fragen.

Sind RCTs in „weichen“ Bereichen nicht durchführbar?

Ein weiterer Einwand lautet: RCTs sind für einfach zu beschreibende, eindimensionale Therapieverfahren nützlich, komplexe Behandlungsansätze oder übergreifende Ziele lassen sich aber damit nicht bewerten. Das Spitzenargument lautet: Medizin ist nicht reine Wissenschaft, sondern immer auch Kunst. Also ist die Frage, ob sich ein „weiches“ Konzept wie Patientenorientierung mittels hochwertiger randomisierter kontrollierter Studien untersuchen lässt, von besonderem Interesse. Bei dieser Überlegung ist der kritische Einwand berücksichtigt, Patientenorientierung müsse man gar nicht untersuchen, weil sie als Ziel jeder medizinischen Schulrichtung gewissermaßen selbstverständlich sei. Einen ausgezeichneten Überblick über die einschlägigen Forschungsprobleme wie -ergebnisse liefern Stewart et al. [8]. Die nachfolgende Grafik (eigene Darstellung) zeigt, zu welchen Endpunkten im Feld Patientenorientierung überhaupt Studien vorliegen, wieweit es sich hierbei um RCTs handelt, und welche Aussagen sich mithin heute auf RCTs stützen lassen. So lässt sich feststellen, dass patientenzentrierte Medizin keineswegs, wie gern behauptet wird, zeitaufwändiger ist: dies ist bekanntlich eines der Standardargumente gegen eine Stärkung der „sprechenden“ Medizin. Am interessantesten ist vielleicht die Studienlage zum Einfluss auf den Krankheitsverlauf: soweit dies mittels RCTs untersucht worden ist, zeigen sich günstigere Verläufe: ein derartiger Befund würde im Bereich der Pharmakotherapie auf allergrößte Beachtung stoßen. Und es könnte vielleicht sehr hilfreich sein, wenn bekannter würde, dass eine eindeutige Patientenorientierung in den untersuchten Studien dazu geführt hat, dass die Kunstfehlerklagen abnahmen: dies könnte ein starkes Motiv für die Ärzteschaft sein, einem entsprechenden Wandel in den Alltagsroutinen mehr Aufmerksamkeit zu schenken. Es ist umgekehrt mehr als verwunderlich, dass andere Erfolg versprechende Zielparameter wie Patientenzufriedenheit und Therapietreue in ihrem Zusammenhang mit einer dezidierten Patientenorientierung bislang nicht mittels RCT untersucht worden sind. Am aller-

erstaunlichsten ist vielleicht, dass der Zusammenhang zur arztseitigen Zufriedenheit generell am schlechtesten untersucht ist (s. Abb. 1).

Jedenfalls gibt es keinen Grund zu der Annahme, RCTs eigneten sich ausschließlich für die gewissermaßen praxisfernen, monodirektionalen Ansätze aus Arzneimittel-Zulassungs-Studien und könnten komplexe Vorgänge wie die Arzt-Patienten-Interaktion nicht vernünftig abbilden. Und: diese Forschung liefert nebenbei starke Hinweise darauf, dass gerade die am besten mittels RCT untersuchte Arzneimittelbehandlung alles andere als eindimensional verläuft; nur dass der Begriff des Placebos hier herkömmlich auf das verabreichte Substrat in Kontrollgruppen verkürzt wird.

Erübrigen bestehende Standards Belege durch RCTs?

Die Medizin lebt in weiten Bereichen von Standards, welche die jeweiligen Communities (Fachgesellschaften wie Berufsverbände) für gesetzt halten. Viele Lehrbücher sind letztlich Kompilationen derartiger Standards, die häufig Ausfluss medizinischer Schulen sind. In weiten Bereichen spiegeln sie die Konsensbildung zwischen führenden Autoritäten wider. Wenn man sich im Bereich der Psychopharmakabehandlung einmal den Einsatz der Antidepressiva ansieht, so findet man durchgängig die Zuordnung der verschiedenen Wirkstoffklassen zu verschiedenen Depressionstypen und -verläufen. Der Nutzen der medikamentösen Depressionsbehandlung wird generell unterstellt, erscheint m.a.W. nicht weiter begründungspflichtig. Noch kann man davon sprechen, dass die kritische Bewertung der vorliegenden Studien zur Wirksamkeit der pharmakologischen Depressionsbehandlung eines der bestgehüteten Geheimnisse der modernen Therapieforschung ist. Als Kirsch u. Sapirstein 1998 ihre Metaanalyse der von ihnen gefundenen 19 RCTs zur Wirkung oraler Antidepressiva publizierten [9], fand diese Veröffentlichung in der psychiatrisch-psychotherapeutischen Fachwelt in Deutschland kaum Aufmerksamkeit. Die Autoren hatten gezeigt, dass Placebos im Durchschnitt 75 % des Effektes des aktiven Medikaments erreichten. Noch irritierender war ihre Hypothese, die verbleibenden 25 % an „echter Wirkung“ könnten einem aktiven Placeboeffekt geschuldet sein, wonach Patienten das Verum an den

charakteristischen Nebenwirkungen erkennen und ihm damit eine höhere Potenz zuweisen konnten als einem inerten Placebo. Aber auch der zu ganz ähnlichen Ergebnissen kommende Bericht in der Cochrane Library aus 2004 [10] fand keine der Radikalität der Aussagen entsprechende Resonanz. Bilanzierten die Autoren doch, dass die Unterschiede zwischen den Effekten von Antidepressiva und Placebos erstaunlich klein seien und dass dringend weitere mehrarmige Studien unter Einschluss aktiver Placebos durchgeführt werden müssen, welche die Nebenwirkungen der Antidepressiva so gut es geht imitieren. Eine Reaktion haben diese Metaanalysen allemal verdient, auch wenn man wie Hollon et al. [11] – bei Bestätigung der Ergebnisse – für die Praxis daraus nicht die Konsequenz zieht, Antidepressiva seien deshalb etwa grundsätzlich entbehrlich. Parker et al. [12] geben in der Antidepressivadebatte zu bedenken, dass es eine hohe Rate an nicht publizierten RCTs gibt – zumeist solche mit negativen Ergebnissen. Sie beklagen weiter, dass in RCTs in der Regel zu wenig Patienten mit schweren Verläufen rekrutiert würden. Und sie kritisieren, dass die Outcomeindikatoren oft keine ausreichende Relevanz aufwiesen. Parker schlussfolgert an anderer Stelle [13], dass aus diesen Gründen die gültigen Leitlinien zur Depressionsbehandlung relativ wertlos seien, zumal sie häufig nur die Botschaft des „anything goes“ verkündeten. Warum all dies gegen gut gemachte RCTs sprechen soll, bleibt unklar. Man muss wohl eher umgekehrt argumentieren: wenn es nicht einmal für den harten Endpunkt „Suizid“ Studien mit angemessenen Wirksamkeitsbelegen gibt [14], dann ist damit zwar immer noch nicht das letzte Wort über die Sinnhaftigkeit der pharmakologischen Depressionsbehandlung gesprochen – aber wie lange können sich Psychiatrie und Psychotherapie mit dieser Situation zufrieden geben? Die systematischen Schwächen vieler RCTs haben vor allem etwas mit Publikationsstrategien und mangelnder Bereitschaft zur Veröffentlichung aller Behandlungsprotokolle und fehlender Registrierung klinischer Studien zu tun [15].

Zwischenbilanz: EbM als Chance für die Psychotherapieforschung

Es sei die Behauptung gewagt, dass es wenig Sinn stiftet, über die (fraglos vorhandene und zum Teil irritierende) Arzneimittellastigkeit in der Gesundheitsforschung, die Schwächen vieler publizierter RCTs und die Mühen der Finanzierung und Durchführung versorgungsrelevanter RCTs zu klagen. Stattdessen sollten möglichst viele zukunftsweisende Studiendesigns mit direktem Vergleich verschiedener Therapiekonzepte (mehrarmige Studien mit und ohne Arzneimittel) entwickelt werden. Es mag sein, dass die Rekrutierung von Probanden in solchen Studien schwieriger ist als in traditionellen Arzneimittelstudien, es kann wohl auch in der näheren Zukunft so bleiben, dass öffentliche und private Sponsoren größere Zurückhaltung zeigen und lieber die eingespielten Traditionen der Forschungsförderung bedienen. Aber eine Umkehr ist nötig. Es soll hier nur auf zwei Studien hingewiesen werden, die aus Sicht des Autors relevante Fragen mit angemessener Methodik beforscht haben bzw. noch weiter untersuchen. Zunächst auf eine dreiarmlige Studie zur Depressionstherapie, welche zwei überwiegend von Psychologinnen/Psychologen durchgeführte psychotherapeutische Verfahren – nicht direktive Beratung nach Rogers und kognitive Verhaltenstherapie – gegen die Standardbehandlung in der Allgemeinmedizin getestet hat

[16]. Danach zeigt sich, dass in dieser Untersuchung psychotherapeutische Verfahren der in der Allgemeinmedizin üblichen Standardtherapie mindestens ebenbürtig und – entgegen weit verbreiteter Annahmen – nicht teurer sind. Und zweitens eine Studie, die momentan im Rahmen des Kompetenzzentrums Schizophrenie durchgeführt wird [17]. Die Forschergruppe geht erstens der Frage nach, wie groß der Fortschritt durch den Einsatz von so genannten atypischen Neuroleptika in der Behandlung schizophrener Patienten tatsächlich ist: laut einschlägiger Fachgesellschaften wird diese Innovation bekanntlich als Meilenstein der Psychiatriegeschichte gefeiert. In dieser Studie wird für die Erstmanifestation einer Schizophrenie ein so genanntes Atypikum gegen ein niedrig dosiertes Standardmedikament (Haldol) getestet. Für die Phase der ersten Remission wird dann angesichts der hohen „Non-Compliance-Rate“ der Patienten der Ansatz einer psychotherapiegestützten pharmakologischen Eskalation untersucht. Diese Studie erscheint aus zwei Gründen vorbildlich: es wird ein Manko der Atypikaforschung adressiert, da viele Überlegenheitsstudien bisher mit inadäquaten Dosen an Haldol im Kontrollarm gearbeitet haben, und es wird der Tatsache Rechnung getragen, dass Patienten verständlicherweise nach der ersten Remission keine Neigung haben, die Neuroleptika ohne weiteres unbefristet einzunehmen.

Die Psychotherapie hat m.a.W. allen Grund, ihren Stellenwert in der Behandlung seelisch Kranker im Rahmen angemessener Studien untersuchen zu lassen. Noch einmal zurück zur Depressionsbehandlung. Das National Institute of Mental Health in den USA hatte bereits 1989 wegen einer unzureichenden Studienlage eine differenziertere Forschungsstrategie (Subgruppen, deutlicher abgrenzbare Interventionen unterschiedlicher Schulrichtungen, angemessene Laufzeit) gefordert [18]. Zehn Jahre später stellt Andrews in einem Rückblick auf 50 Jahre RCT-Forschung fest, dass die Psychiatrie offenbar nach wie vor große Berührungssängste gegenüber Wirksamkeitsnachweisen habe [19]. White und Walden haben im selben Jahr gezeigt, wie groß die Widerstände unter Psychotherapeuten in Australien waren, sich an einer auf zwei Jahre begrenzten RCT zur Frage der Wirksamkeit einer tiefenpsychologisch fundierten Psychotherapie von depressiven Patienten mit bisherigem Therapieversagen zu beteiligen [20]. Schließlich sei die These gewagt, dass ein weiteres Defizit der Psychotherapieforschung darin besteht, dass die Einbindung der hausärztlichen Versorgung in patientenzentrierte Therapieforschung zu wenig thematisiert und erprobt wird: und das, obwohl es gute Hinweise dafür gibt, dass eine Systematisierung der Zusammenarbeit zwischen Hausärzten und Fachärzten bessere Versorgungsergebnisse mit sich bringt [21,22]. RCTs sind nach allen bisherigen Erfahrungen kein Hindernis für die Weiterentwicklung der Psychotherapie in der Versorgungslandschaft.

Erster Perspektivwechsel: Das Lehrstück der Grounded Theory

Nun haben RCTs im Rahmen der Strukturierung der medizinischen Versorgung sicher keinen Monopolanspruch für die Generierung versorgungsrelevanter „Daten“. Wenn kaum nachvollziehbar ist, warum RCTs in der Medizinforschung immer noch einen so schweren Stand haben, dann ist mindestens so unver-

ständig, dass die Bedeutung der qualitativen Sozialforschung für das Verständnis therapeutischer Zusammenhänge bis heute weitestgehend unbekannt geblieben ist. Dies ist deshalb so schmerzhaft, weil allemal im Zeitalter der chronischen Erkrankungen das traditionelle Krankheitsmodell der Akutmedizin („ein entzündeter Blinddarm muss entfernt werden, weil der Patient sonst mit hoher Wahrscheinlichkeit stirbt“) an massive Grenzen stößt. Dass das Erheben von Primärdaten in ihren sozialen Kontexten mittels teilnehmender Beobachtung, narrativer Interviews und anderer Methoden häufig unverzichtbar ist, um komplexe Wirkungszusammenhänge besser zu verstehen, das passt offenbar in die Welt ärztlichen Denkens und Handelns schwer hinein – so sehr die Ärzte auch gern berichten, ihre Profession definiere sich nicht allein als Wissenschaft, sondern mindestens genauso sehr als Kunst. Die Grounded Theory von Glaser und Strauss [23] ist der etablierten Medizin bis heute ein Buch mit sieben Siegeln geblieben. Wie bedauerlich (oder auch: wie verständlich?) dies ist, lässt sich an der jüngst publizierten Dissertation von Anja Hermann zeigen [24]. Hermann hat in ihrer multimodalen Feldforschung gezeigt, dass in einer Spezialklinik für die Behandlung von Patientinnen und Patienten mit weit fortgeschrittenem Knochenkrebs das Thema der radikalen Begrenzung der Lebenserwartung durch die Erkrankung in der Arzt-Patient-Beziehung schlicht keinen Platz hat – und dass eine Kommunikation über die Begrenztheit der therapeutischen Ansätze auch im Gespräch zwischen Patientinnen/Patienten mit ihren Angehörigen keinen Raum findet: das Arrangement der Hoffnung, wie die Autorin es nennt, ist so perfekt, dass Illusionen die Realität erdrücken. Um welchen Preis? Man könnte ketzerisch formulieren: Die Studie von Hermann liefert keine unmittelbar verwertbaren, aber hoch relevante Ergebnisse zum Thema „Verhinderung gelingender Kommunikation“ oder „Palliativmedizin zwischen integrierter Versorgung und rhetorischen Alibis“. Man muss unwillkürlich an die Studie „Awareness of Dying“ von Glaser und Strauss denken [25].

Das vielleicht härteste Vorurteil: qualitative Methoden sind für Evidence-based Health Care nicht von Gewinn

EbM ist „nicht mehr“ als die kritische Bewertung vorhandener publizierter Evidenz. Evidenzbasierte Versorgung (Evidence-based Health Care = EbHC) leitet sich erst daraus ab: die Frage, welche Leistungen ein Gesundheitssystem unter Berufung auf vorhandene Evidenz aufnimmt oder ablehnt, unterliegt einer ganzen Reihe von fachlichen, ethischen und ökonomischen Gründen [3]. Ein aktuelles Beispiel soll dies verdeutlichen. So wird die Akupunktur im Rahmen der Behandlung bestimmter chronischer Schmerzzustände ab Herbst 2005 im zuständigen gemeinsamen Bundesausschuss (www.g-ba.de) beraten werden. Nachdem bei schwer zu interpretierender Studienlage vom damaligen Bundesausschuss ab 2000 befristet hochwertige Studien zur Klärung des Nutzens, der Risiken und der Wirtschaftlichkeit der Akupunktur zugelassen worden waren, liegen jetzt peu à peu die Ergebnisse der RCTs vor, welche zeigen sollten, ob das Setzen von Körpernadeln an definierten Orten gemäß der Lehre der Traditionellen Chinesischen Medizin (TCM) eine spezifische Wirksamkeit gegenüber dem Setzen derselben Nadeln an anderen, TCM-gemäß „unspezifischen“ Punkten erbringt. Bei einer gewissen Spreizung der Ergebnisse zeichnet sich als entschei-

gendes und ausgesprochen überraschendes Ergebnis ab (dies ist einstweilen die persönliche Einschätzung des Autors), dass die „spezifische“ Akupunktur der „unspezifischen“ Akupunktur gar nicht oder allenfalls geringfügig überlegen ist, dass aber das Nadeln insgesamt der medikamentösen Standardtherapie überlegen ist [26]. Diese Ergebnisse zeigen zum einen, dass eine angeblich jahrtausende alte bewährte Methode, die dann in den 60er-Jahren in den reichen Industrienationen neu „entdeckt“ und adaptiert wurde, bislang unzulänglich auf ihre Wirksamkeit hin untersucht worden war. Nun wird sich aber zeigen, dass mit diesem wichtigen Kernbefund der großen RCTs zur Akupunktur bei chronischen Schmerzzuständen die Frage nach dem Stellenwert der Akupunktur noch nicht beantwortet ist. Welche Form der Akupunktur soll von Akupunkteuren welchen Ausbildungsstandes wie häufig und in welcher Einbindung in andere Therapie-schemata angewandt werden? Ist die Akupunktur überhaupt sinnvollerweise ein Baustein ärztlicher Behandlung? Hierfür fehlen vor allem Daten, welche Erwartungen Patientinnen und Patienten in ihren jeweiligen biografischen und krankheitsbedingten Kontexten mit der Akupunktur verknüpfen. Derartige Forschung ist rar (eine Ausnahme stellt die Studie von Froberg, Schütze, Voigt, Detka, Dorendorf u. Malina dar [27]) und muss gerade im immer wieder verkündeten Zeitalter der Patienten-zentrierung in der Medizin systematisch gefördert werden. Erst über die Integration von quantitativer und qualitativer Forschung wird sich in dieser Situation in einem vertretbaren Zeitraum der Stellenwert der Akupunktur im Versorgungsgeschehen evidenzbasiert wie gesellschaftlich akzeptiert bestimmen lassen.

Zweiter Perspektivenwechsel: Kernergebnisse zum Qualitätsmanagement in der Medizin

Ad 1: Qualitätsmanagement muss verstanden werden

Einmal unterstellt, es gäbe einen Konsens in den verschiedenen medizinischen Communities, dass evidenzbasierte Leitlinien und Standards im Prinzip hilfreich und dem traditionellen Lehrbuch-„kompass“ überlegen sind (eine deutlich abweichende Position vertritt z.B. der Präsident der Bundesärztekammer [28]), dann bleibt immer noch die Frage offen, wie es gelingt, derartige Leitlinien in den Versorgungsalltag zu integrieren, tatsächlich routinefähig zu machen. In einer kritischen Übersichtsarbeit hat Grol gezeigt, dass es bislang keine wirklich überzeugende, leicht definierbare Qualitätsstrategien gibt, evidenzbasierten Leitlinien dergestalt zum Leben zu verhelfen. Qualitätssteigernd wirkt offenbar immer nur ein ganzes Bündel an Maßnahmen, welches vor allem mit Remindern, Rückspiegelung relevanter Qualitätsindikatoren und dem Einsatz von Peers arbeitet. Professions- und organisationsbezogene Ansätze müssen ineinander greifen. Am Ende scheint von überragender Bedeutung zu sein, ob Therapeuten der Auffassung sind, dass Leitlinien mit ihrem professionellen Selbstverständnis vereinbar sind und die ärztliche Autonomie und die Individualität der Arzt-Patient-Beziehung gewahrt bleiben. Diese Grundfragen therapeutischen Selbstverständnisses haben offenkundig eine eigene Bedeutung, ganz unabhängig von anderen Kernfragen, nämlich ob Ärzte überhaupt Fortbildungsmaßnahmen wahrnehmen oder ob sie wirtschaftliche Gründe gegen die Übernahme von Leitlinienempfehlungen ins Feld führen. Grols generelle Schlussfolgerung [29]: es müssen Brücken geschlagen werden zwischen dem professionellen

Selbstverständnis, externen Qualitätsindikatoren, dem Interesse der Kostenträger und den Patientenbedürfnissen. Damit wird erneut ein gewaltiges Feld für qualitative Forschung erkennbar. Zeigt sich doch, dass Grundfragen der Realisierung von wünschenswerter Qualität und der Vertrauensbindung in der Medizin bisher nicht angemessen verstanden werden.

Ad 2: International wurden Forschungsfelder für Qualitätsmanagement definiert

Im Jahr 2000 legte das Institute of Medicine seinen Bericht „To Err is Human“ [30] vor. Die Kernaussage, dass vermeidbare behandlungsinduzierte Erkrankungen und Todesfälle im Gesundheitssystem der USA inzwischen zu einer großen Belastung geworden sind (und es gibt keinen Grund zu der optimistischen Annahme, dies sei in anderen Systemen grundsätzlich anders), stieß auf große Empörung und Erschrecken. Inzwischen verschafft sich aber immer stärker die Haltung Respekt, derartige Analysen nicht als Diffamierung zu verstehen, sondern als Ansporn zur Entwicklung einer neuen Kultur der Fehlerfreundlichkeit. Im Gefolge des genannten IOM-Berichtes von 2001 wurde u. a. das Harvard Interfaculty Program on Health Systems Improvement (PHSI) gegründet [31]. Es definierte fünf Forschungsfelder:

1. **Organizational Supports:** Was muss sich im institutionellen Ablauf ändern, damit Fehler rascher erkannt und vermieden werden?
2. **Evidence-based Care:** Wie kann der Erfahrungsschatz zu Nutzen und Risiken von Verfahren so rasch und wirksam wie möglich in die Praxis gelangen?
3. **Clinical Information Technology:** Wie können moderne Informationstechnologien im Sinne der Fehlererkennung und -vermeidung intelligent genutzt werden?
4. **Payment Policies:** Welche ökonomischen Anreize müssen gesetzt werden, um eine Fehlerkultur zu unterstützen?
5. **Preparing the Workforce:** Welche Aus-, Fort- und Weiterbildungswege müssen verstärkt werden, um die Beschäftigten für die gewachsenen Belastungen in der Praxis auszurüsten?

Für alle Felder werden sowohl quantitative wie qualitative Methoden als unverzichtbar gehalten: Immer geht es sowohl um die Produktion von verlässlichen Indikatoren wie das Verstehen förderlicher wie hemmender Faktoren. Das Versorgungssystem selber ist m. a. W. zum zentralen Gegenstand der Qualitätsforschung geworden.

Dritter Perspektivwechsel: Die HTA-Community diskutiert über angemessene Methoden der Bewertung von Behandlungsverfahren

In allen entwickelten Gesundheitssystemen haben sich sowohl inneruniversitär wie in den Versorgungssystemen Arbeitsgruppen und zuständige Institutionen zur Bewertung von Untersuchungs- und Behandlungsverfahren für das Versorgungssystem entwickelt – in Deutschland sind dies für die GKV der zuletzt 2004 reorganisierte Gemeinsame Bundesausschuss (G-BA) und das neu gegründete Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Die Aufmerksamkeit dieses Health Technology Assessment (notabene: technology schließt alle Verfahren einschließlich Psychotherapie ein und meint nicht

technische Verfahren in der Medizin!) liegt stark bei der Suche nach RCTs und darauf fußender Metaanalysen. Es zeichnet sich jetzt aber bereits ab, dass vor allem für die Frage nach der Optimierung von Versorgungsstrategien künftig die Heranziehung qualitativer Studien an Bedeutung gewinnen wird. Einen ersten Überblick über den Stellenwert qualitativer Forschung für HTA findet sich bei Murphy et al. [32]. Danach gelten folgende Felder als Domäne qualitativer Studien ab:

1. die Entwicklung von Forschungsfragen,
2. die Interpretation quantitativer Forschungsergebnisse für die Strukturierung der Praxis,
3. die Erläuterung von Kontexten von Forschungsergebnissen für Umsetzung/Politik,
4. die Implementations- und Impactforschung zu HTA selbst.

Die Frage nach der Angemessenheit von Methoden ist m. a. W. im professionellen Kern der Bewertungsprozesse selber angekommen – auch wenn die Übermacht an quantitativer Forschung darüber vielleicht noch eine Zeit hinwegtäuschen mag – und wenn auch in Zukunft viele Ressourcen notwendigerweise weiter in die Planung und Durchführung hochwertiger RCTs fließen müssen: da am Ende die Frage nach der Überlegenheit anderer Verfahren ohne RCTs schlicht nicht zu beantworten ist.

Es ist hier vielleicht der Ort, an eine Sentenz zweier EbM-Pioniere zu erinnern. Sackett u. Wennberg [33] betitelten 1997 ein Editorial wie folgt: „Choosing the best research design for each question: It's time to stop squabbling over the 'best' methods.“ In dieselbe Richtung zielt Frommer mit seiner Anmerkung zum Forschungsstreit in der Psychosomatik/Psychotherapie [34]: „Der Dichotomie zwischen qualitativer und quantitativer Forschung kommt...methodisch eine zentrale Bedeutung zu. Ein einfacher Kompromiss dahingehend, dass biologische Merkmale quantitativ-erklärend und psychosoziale Merkmale qualitativ-verstehend zu untersuchen seien, ist bereits seit der Etablierung der Experimentellen Psychologie vor über 100 Jahren nicht mehr denkbar“. Frommer, Langenbach und Streeck weisen noch auf einen weiteren vielleicht besonders erstaunlichen Befund hin, dass nämlich in der Psychotherapieforschung die Biografieforschung ein stiefmütterliches Dasein fristet: man könnte dies in einem unabweislich auf hermeneutische Ansätze angewiesenen Therapieverfahren als eine noch unerklärte Paradoxie bezeichnen, deren Auflösung jedenfalls nur durch eine bessere Verzahnung von quantitativer und qualitativer Forschung denkbar erscheint [35].

Vision: Netzwerk Methodenforschung in den Gesundheitswissenschaften

Mit der Formulierung von Visionen soll der Mensch bekanntlich vorsichtig sein, und doch muss man heute die Etablierung einer leistungsfähigen Infrastruktur für angemessene Methodenforschung im Gesundheitswesen noch als Vision bezeichnen. Angesichts der Tatsache, dass sowohl hochwertige biometrische als auch hermeneutische Verfahren große Akzeptanzprobleme in weiten Teilen der Scientific Community, aber auch unter Entscheidungsträgern haben, wäre es lohnend, ein Netzwerk zu gründen, welches die Vorteile der Methodentriangulation jenseits überholten Schulendenkens in Politik, Öffentlichkeit und last not least in die wichtigen Gutachtergremien der Förderinsti-

tutionen anhand anschaulicher Beispiele hineinträgt. Der dogmatische Streit „qualitativ versus quantitativ“ sollte baldmöglichst auch in der Gesundheitsforschung der Geschichte angehören. Im *British Medical Journal* hat die Soziologin Ann Oakley 1998 auf die große Tradition von RCTs für die Bewertung komplexer sozialer Interventionen in den USA im Zeitraum von etwa 1960 bis 1985 hingewiesen [36] – eine wirkliche Erfolgsgeschichte, die aber von politischer Seite offenbar nicht zuletzt deshalb beendet wurde, weil die RCTs häufig nicht die erhofften positiven Effekte gezeigt oder komplexere Lösungswege für erforderlich gehalten hatten. Warum die Psychotherapie schwerer mittels aller zur Verfügung stehenden methodischen Instrumente beforschbar sein sollte als hoch komplexe Unterstützungssysteme für Familien in prekären Lebenslagen, das wird in Kenntnis der Genese der evidenzbasierten Medizin und Sozialwissenschaften insofern eine rein rhetorische Frage. Oakley steht als Forscherin für die Integration von qualitativer und quantitativer Forschung [37,38]. Es geht, so beschließt sie ihren Aufsatz im BMJ, um den Schutz der Öffentlichkeit vor potenziell gefährlichen, unkontrollierten Experimenten und um ein besseres Verständnis, wie professionelle Interventionen Nutzen stiften können.

Schlussbetrachtung

In allen Feldern klinischer Forschung, so auch zu Wirksamkeit und Effizienz psychotherapeutischer Verfahren, werden nach wie vor Bedenken gegen den Anspruch der evidenzbasierten Medizin vorgebracht, wonach sich Aussagen zu Nutzen und Schaden von Behandlungsverfahren vor allem auf Ergebnisse methodisch hochwertiger quantitativ ausgelegter Studien, in der Regel RCTs stützen müssen. Sosehr die hierbei vorgebrachten Gründe vor allem angesichts der Privilegierung der Pharmakotherapie in der Forschungslandschaft auch zu verstehen sein mögen, so wenig können sie überzeugen. Entsprechende Forschungsdesigns mögen wesentlich aufwändiger sein als in der Arzneimittel-forschung; es ist auch mehr als bedauerlich, dass Mittel für derartige Forschung oft schwer zu akquirieren sind. Langfristig wird sich die Psychotherapie ihren Platz in Gesamtkonzepten angemessener Versorgung ohne stärkere Berücksichtigung von RCTs aber nicht sichern können. Ein zweiter, die somatische wie die psychosoziale Versorgung gleichermaßen betreffender Mangel ist demgegenüber in der chronischen Unterschätzung qualitativer Forschungsmethoden zu sehen. Dies ist wohl nicht zuletzt einem der Zahlengläubigkeit verfallenen Zeitgeist zu verdanken, der nicht zur Kenntnis nehmen will, das zur Generierung intelligenter Fragestellungen, zur Interpretation quantitativer Forschungsergebnisse und zur Gewinnung größeren Verständnisses für Krankheits- und Gesundungsprozesse hermeneutische Verfahren unverzichtbar sind.

Fazit

Die evidenzbasierte Medizin (EbM) gilt vielen Klinikern und Forschern nach wie vor als zu anspruchsvoll und zu wenig praxisorientiert. So wird in der Psychotherapie insbesondere bezweifelt, ob der Goldstandard der Nutzenbewertung, näm-

lich die randomisierten kontrollierten Studien (RCT), ähnlich sinnhaft eingesetzt werden können wie in der somatischen Medizin. Diese grundsätzlichen Bedenken sind ebenso wenig begründet wie sich EbM auf RCT reduzieren lässt. Bedenken richten sich offenbar gleichermaßen gegen den Stellenwert hermeneutischer Verfahren für ein besseres Verständnis der vielschichtigen und häufig nicht transparenten Therapielandschaft in der Psychotherapie. Es zeichnet sich ab, dass das traditionelle Schulendenken in der Psychotherapie durch die Anwendung angemessener Forschungsmethoden noch stärker infrage gestellt werden wird als dies heute schon der Fall ist.

Literatur

- Koshland DE. Scientific Literacy. *Science* 1985; 230: 4724
- Lambert MJ, Ogles BM. The efficacy and effectiveness of psychotherapy. In: Lambert MJ (eds): *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*. New York: Wiley, 2004: 139–193
- Leichsenring F. Are psychodynamic and psychoanalytic therapies effective? A review of empirical data. *Int J Psychoanal* 2005; 86: 841–868
- Reisner AD. Common factors, empirically validated treatments, and recovery models of therapeutic change. *Psychological Record* 2005; 55 (3) http://www.findarticles.com/p/articles/mi_qa3645/is_200507/ai_n14849823 (überprüft 28.9.2005)
- Schmacke N. Wie viel Medizin verträgt der Mensch? Bonn: Kompart, 2005: 101–126
- Schmacke N. Innovationen in der Medizin: das magische Dreieck von Evidenz, Werten und Ressourcen. *ZBBS* 2004; 5: 331–353
- Moseley JB, O'malley K, Petersen NJ et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002; 347: 81–88
- Stewart M, Brown JB, Weston WW et al. *Patient-Centered Medicine. Transforming the Clinical Method*. Abingdon: Radcliffe Medical Press, 2003
- Kirsch I, Sapirstein G. Listening to prozac but hearing placebo: a meta-analysis of antidepressant medication. *Prevention & Treatment* 1998; 1: Article 0002a (<http://www.journals.apa.org/prevention/volume1/pre0010002a.html>)
- Moncrieff J, Wessely S, Hardy R. Active placebos versus antidepressants for depression. *The Cochrane Library*, Issue 1. Chichester: John Wiley & Sons, 2004
- Hollon S, DeRubeis RJ, Shelton RC, Weiss B. The emperor's new drugs: effect size and moderation effects. *Prevention & Treatment* 2002; 5: Article 28 (<http://www.journals.apa.org/prevention/volume5/pre0050028c.html>)
- Parker G, Anderson IM, Haddard P. Clinical trials of antidepressant medications are producing meaningless results. *Br J Psychiatry* 2003; 183: 102–104
- Parker G. Critique of the guidelines for the treatment of depression: flaws in the construction. *Aust N Z J Psychiatry* 2004; 38: 885–890
- Antidepressiva: Lebensgefährliche Placebos? *Arzneitelegamm* 2005 (5); 36: 45–47
- An-Wen C, Hróbjartsson A, Haahr MT et al. Empirical evidence for selective reporting of outcomes in randomized trials. *Comparisons of protocols to published articles*. *J Am Med Assoc* 2004; 291: 2457–2465
- Ward E, King M, Lloyd M et al. Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. *Br Med J* 2000; 321: 1383–1388
- Riesbeck M, Wilmsdorff M von, Krohmer R et al. Optimierung der pharmakologischen Langzeitbehandlung ersterkrankter schizophrener Patienten. *Psychoneuro* 2004; 30 (11): 615–619
- National Institute of Mental Health. Treatment of depression collaborative research program. General Effectiveness of treatments. *Arch Gen Psychiatry* 1989; 46: 971–982
- Andrews G. Randomised controlled trials in psychiatry: important but poorly accepted. *Br Med J* 1999; 319: 562–564

- ²⁰ White RT, Walden MA. Obstacles to a randomised controlled trial of intensive dynamic psychotherapy: an account of the New South Wales section of psychotherapy outcomes project. *Aust NZ J Psychiatry* 2000; 34: 271 – 278
- ²¹ Gilbody S, Whitty P, Grimshaw J, Thomas R. Educational and organizational interventions to improve the management of depression in primary care. *J Am Med Assoc* 2003; 289: 3145 – 3151
- ²² Dietrich AJ, Oxman TE, Williams JW et al. Re-engineering systems for the treatment of depression in primary care: cluster randomised controlled trial. *Br Med J* 2004; 329: 602 – 605
- ²³ Glaser B, Strauss A. *The discovery of grounded theory*. Chicago: Aldine, 1967
- ²⁴ Hermann A. *Das Arrangement der Hoffnung. Eine qualitative Untersuchung des kommunikativen und interaktiven Umgangs mit der Knochen- und Weichgewebesarkomen inhärenten Todesdrohung während der stationären chirurgischen Behandlung*. Frankfurt: Mabuseverlag, 2005-09-13
- ²⁵ Glaser B, Strauss A. *Awareness of Dying*. Chicago: Aldine, 1965
- ²⁶ www.gerac.de
- ²⁷ Froberg M, Schütze F, Voigt J et al. Kurzgefasste Zusammenfassung der Ergebnisse des Magdeburger Akupunkturprojektes zur Bedeutung der Akupunktur für AOK-Versicherte und ihre Ärztinnen und Ärzte im Rahmen des Bundesausschuß-Modellvorhabens“, Ms. September 2005
- ²⁸ Hoppe JD. Mehr Vertrauen in die ärztliche Urteilskraft. *Deutsches Ärzteblatt* 2005; 102 (14): A 943
- ²⁹ Grol R. Improving the quality of medical care. Building bridges among professional pride, payer profit and patient satisfaction. *J Am Med Assoc* 2001; 286: 2578 – 2585
- ³⁰ Kohn L, Korrigan JM, Donaldson MS (eds). *To err is human. Building a safer health system*. Committee on quality of health care in America. Institute of Medicine. Washington: National Academy Press, 2000
- ³¹ Fernandopulle R, Ferris R, Epstein A et al. A research agenda for bridging the „quality chasm“. *Health Aff* 2002; 22: 178 – 190
- ³² Murphy E, Dingwall R, Greatbatch D et al. Qualitative research methods in health technology assessment: a review of the literature. *Health Technol Assess* 1998; 2 (16): iii – ix, 1 – 274
- ³³ Sackett DL, Wennberg JE. Choosing the best research design for each question. *Br Med J* 1997; 315: 1636
- ³⁴ Frommer J. Psychosomatische Medizin – Eine Wissenschaft für das 21. Jahrhundert. *Zeitschrift für Psychosomatische Medizin und Psychotherapie* 2004; 50: 20 – 34
- ³⁵ Frommer J, Langenbach M, Streeck U. Qualitative Psychotherapy Research in German-Speaking Countries. *Psychotherapy Research* 2004; 14: 57 – 75
- ³⁶ Oakley A. Experimentation and social interventions: a forgotten but important history. *Br Med J* 1998; 317: 1239 – 1242
- ³⁷ Oakley A. Who's afraid of the randomised controlled trial? Some dilemmas of the scientific method and 'good' research practice. In: Roberts H (ed): *Women's Health Counts*. London: Routledge and Kegan Paul, 1990: 167 – 194
- ³⁸ Thomas J, Harden A, Oakley A et al. Integrating qualitative research with trials in systematic reviews. *Br Med J* 2004; 329: 1010 – 1012