

Problematik der Evidence Based Medicine bei Naturheilverfahren

ANME - Konferenz, Frankfurt, 11.11.2006

Die Frage, in wie weit evidenzbasierte Medizin (EbM) bei Naturheilverfahren aussagekräftig ist wird kontrovers diskutiert. Dabei ist zunächst einmal festzuhalten, dass der Disput über das Thema „EbM und Naturheilverfahren“ meist auf einem sehr niederen wissenschaftlichen Niveau stattfindet. Das hat sich von Seiten der Naturheilverfahren in den letzten Jahren deutlich geändert, zugegebenermaßen nicht ganz freiwillig. Von Seiten der orthodoxen Medizin ist das Niveau aber leider unverändert nieder. Das hat die Diskussion um den Lancet-Artikel zur Homöopathie, den Sie ja vermutlich alle kennen, deutlich gezeigt hat (10).

Das sollte die Vertreter der Naturheilverfahren nicht abhalten, die eigenen wissenschaftlichen Kenntnisse weiter zu verbessern.

Wovon ich dringend abraten möchte, ist es, die Willfährigkeit der Statistik für wirtschaftliche und andere Interessen als Argument zu verwenden, nach dem Prinzip: „Jeder erfindet sich seine Statistik, wie er gerade will“. Natürlich ist das, was ich den „menschlichen Faktor“ nenne, also die Manipulation von Ergebnissen, ein erhebliches Problem in der Medizin. Insbesondere die vielfältige Einflussnahme der Pharmaindustrie führt zu einer Situation, die weit von jedem evidenzbasierten Zustand entfernt ist. Aber es sind ja gerade die Vertreter der EbM, die am intensivsten versuchen, dieses Problems Herr zu werden. Der „menschliche Faktor“ ist zwar ein großes Problem für die EbM, aber es ist nicht der wesentliche Kritikpunkt, wenn es um Naturheilverfahren geht.

Um Ihnen die eigentlichen Probleme darzulegen, zumindest soweit ich sie verstehe, möchte ich Ihnen ein paar Beispiele geben. Jedes dieser Beispiele weist auf tieferliegende wissenschaftstheoretische Schwierigkeiten hin, auf die ich in der Kürze der Zeit jedoch nicht eingehen kann.

1) *Wer heilt hat nicht unbedingt Unrecht*

Oft kommen Patienten mit einer ganzen Plastiktüte voll evidenzgeprüfter Medikamente zu uns, Medikamente, die sie jahrelang nahmen, ohne dass sich an ihrer Erkrankung irgendetwas Grundlegendes geändert hätte. Wenn nun solche Patienten

nach einem Jahr naturheilkundlicher Therapie keine Medikamente mehr nehmen und gesund sind, so ist das zwar kein Beweis für die Wirksamkeit der Naturheilverfahren. Es ist aber ein Beweis dafür, dass es ein Problem mit der Wirksamkeitskontrolle gibt.

2) *Die Unmöglichkeit individueller Aussagen*

Statistiker berichten von speziellen Schwierigkeiten mit randomisierten Studien.

Davon ist das Simpson-Paradox vielleicht das lustigste. Es besagt, dass

Wahrscheinlichkeiten, die für zwei getrennte Populationen richtig sind für die zwei Populationen zusammen nicht richtig sein müssen (7). Das bedeutet, dass etwas, das in einer Studie für Frauen und in einer anderen Studie für Männer gut ist, trotzdem schlecht für die Menschen sein kann (1). Oder wenn in einer Studie Therapie B besser ist als Therapie A und in einer anderen Studie Therapie C besser als Therapie B, so ist es ein Fehlschluss zu glauben, dass Therapie C der Therapie A überlegen ist (2). Ein weiteres der vielen statistischen Problem ist, dass Statistiken nur Wahrscheinlichkeiten für definierte Populationen angeben, jedoch nichts über den Einzelfall aussagen können. Es ist sogar so, dass unsere heutige Medizin gar keine Möglichkeit besitzt, eine individuelle Prognose auf wissenschaftlicher Grundlage zu erstellen. Kritiker des statistischen Ansatzes haben es einmal so ausgedrückt: „Große Zahlen liefern ein statistisch gesehen genaues Ergebnis, von dem man nicht weiß, auf wen es zutrifft. Kleine Zahlen liefern ein statistisch gesehen unbrauchbares Ergebnis, von dem man aber besser weiß, auf wen es zutrifft. Schwer zu entscheiden, welche dieser Arten von Unwissen die nutzlosere ist.“ (4)

3) *Was kurzfristig hilft, kann langfristig schaden*

Den nächsten Punkt möchte ich an ein paar Studien verdeutlichen:

- Bei der Therapie der Epicondylitis konnte gezeigt werden, dass eine Injektion mit Kortison kurzfristig eine deutliche Verbesserung des Krankheitsbildes mit sich bringt. Langfristig ist sie aber der Physiotherapie oder dem reinen Abwarten deutlich unterlegen (22)
- Eine Studie über Lauflernhilfen zeigt, dass sogenannte Baby-Walkers Kleinkinder eher in eine aufrechte Position bringen, dass aber diese Kinder später alleine stehen und gehen lernen (14)
- Kinder, die in den ersten 6 Monaten ihres Lebens mit Antibiotika behandelt werden haben eine höhere Wahrscheinlichkeit Allergien zu entwickeln (16). Wobei dieses Ergebnis nicht unumstritten ist (5).

- Frühchen, die eine angereicherte Nahrung erhalten, entwickelten im Jugendalter häufiger eine Insulinresistenz als solche, die mit Muttermilch oder einer üblichen Formula-Milch gefüttert wurden. Das heißt, das schnelle Erreichen eines Referenzwertes ist nicht gesünder (21).

Wenn das Prinzip „hilft kurzfristig, schadet langfristig“ weiter verbreitet wäre, wofür es eindeutige theoretische Modelle gibt, dann wären solche Studien, die nur einige Monate umfassen (also die meistem) und die die Wirksamkeit einer Therapie beweisen, geradezu eine Anleitung zur Schädigung.

4) *Je beeinträchtigender der Kontext desto spezifischer die Wirksamkeit*

Ein anderes Phänomen soll am Beispiel der Balneologie demonstriert werden: Wer sich einmal in alten Kurorten aufhielt und vielleicht sogar einige Anwendungen mitgemacht hat, kennt die heilsame Wirkung einer schönen Architektur und der dazugehörenden Atmosphäre. All dies vermittelt ein Gefühl von innerer Würde. Heutzutage werden aber balneologische Anwendungen vielfach in Funktionsbauten durchgeführt, die den Charme verschwitzter Socken besitzen. In einem solchen Setting werden spezifische Teilaspekte evidenzbasiert geprüft, beispielsweise welchen Effekt bestimmte Übungen auf die Mortalität von Patienten mit linksventrikulärer systolischer Funktionsstörung besitzen (12). Was ist der Wert solcher Studien? Sind sie wirklich unabhängig vom äußeren Rahmen? Und vor allem, was ist von einer Medizin zu halten, die Schönheit und Würde allenfalls als Placebo-Effekt betrachtet? Dies ist kein Seufzer eines Ästheten. Es handelt sich hierbei um ein erhebliches wissenschaftliches Problem. So konnte bei Hamstern gezeigt werden, dass Wunden dann schneller heilen, wenn sie sich in ihrem normalen sozialen Kontext befanden. Wurden sie in ein Glasröhrchen gesteckt, verschlimmerten sich die Wunden zunächst einmal (11). Würde also in einem solchen Setting die Wirksamkeit einer wundheilenden Substanz geprüft, so könnte sie sich im ersten Fall als Placebo erweisen, im zweiten Fall jedoch als spezifisch wirksames Medikament. Das bedeutet, dass eine Substanz (spezifisch) umso wirksamer erscheinen wird, je schädlicher der therapeutische Rahmen ist. In der Medizin wurde der Widerspruch zwischen krankenhausbasierter Forschung (Glasröhrchen) und hausärztlicher Anwendung zwar schon vielfach diskutiert, aber nicht wirklich zu Ende gedacht. Für die Naturheilverfahren stellt sich das Problem noch ausgeprägter,

da die meisten dieser Verfahren (anerkannterweise) ein positives therapeutisches Klima erzeugen.

5) *Anpassungsfähigkeit versus Rigidität*

Wechselnde Kontexte spielen aber auch in anderer Hinsicht eine Rolle. So sind Beta-Blocker bei verschiedenen Erkrankungen evidenzgeprüft wirksam, vom Hypertonus über das Glaukom bis zu Erregungszuständen. Als Nebenwirkung besteht jedoch eine verminderte Anpassungsfähigkeit an Hitze, die mit einer höheren Tendenz zum Hitzeschlag einhergeht (8). Die positive Wirkung, die in Studien gefunden wurde, wäre in dem Augenblick hinfällig, in dem der Klimawandel zu heißeren Sommern führen würde. Das heißt, ein gut messbarer Effekt, der mit einer gewissen Rigidität erkauft wird, kann unter sich ändernden Umständen zu erheblichem Schäden führen. Demgegenüber kann ein kaum messbarer Effekt, der zu einer Besserung der Anpassungsfähigkeit führt, unter sich ändernden Umständen von großem Nutzen sein. Was im einen Kontext nützlich ist, kann in einem anderen Kontext schaden. Das bedeutet, randomisierte Studien können nicht zwischen Rigidität und Anpassungsfähigkeit unterscheiden.

6) *Lernen höherer Ordnung*

Ich möchte Ihnen ein logisches Instrument präsentieren, das geeignet ist, diesen Prozess genauer darzustellen, nämlich die verschiedenen Ordnungen des Lernens, wie sie von Bateson eingeführt wurden (3).

- Lernen 0. Ordnung besteht, wenn eine Einheit keine oder nur eine geringe Änderung in ihrem Reaktionsmuster auf wiederholte Reize aufweist. Diese Form des Lernens findet sich in einfachen maschinellen Regelkreisen, oder wenn ein Organismus überstimuliert ist, oder wenn die Reizantwort strukturell festgelegt ist
- Lernen 1. Ordnung ist das, was wir typischerweise als Lernen verstehen. Es ist das Pawlowsche Lernen, wie es unter Laborbedingungen geprüft werden kann. Auch das Erlernen einer Sprache ist Lernen 1. Ordnung.
- Lernen 2. Ordnung ist das Lernen des Lernens. Wer eine Sprache erlernt, lernt auch gleichzeitig, andere Sprachen leichter zu lernen. Lernen 2. Ordnung findet sich überall, ist aber nicht immer positiver. Wer beim Lernen einer Sprache unangenehme Erfahrungen macht, mag einen Widerwillen gegen das Erlernen von Sprachen entwickeln, so dass er/sie die zweite Sprache schlechter oder gar nicht erlernt. Es stellt sich das Problem, dass Lernen 2. Ordnung nicht durch einfache Messungen entdeckt werden kann. Man benötigt dazu eine Serie von Messungen *und* eine entsprechende Theorie

- Lernen 3. Ordnung ist charakterisiert durch eher grundlegende Änderungen, die über das eigentlich zu Lernende hinausgehen. Ein Wechsel von einem linearen, kausalorientierten Denken zu einem systemischen, lösungsorientierten Denken wäre Lernen 3. Ordnung. Es geht dabei nicht mehr um Lerninhalte sondern um die grundlegende Bewertung der Lerninhalte
- Lernen 4. Ordnung wäre eine Änderung der lernenden Struktur und umfasst morphologische und genetische Einflüsse

Im medizinischen Kontext ist beispielsweise der Erwerb einer spezifischen Immunität Lernen I. Wenn durch eine Infektion gleichzeitig die Widerstandsfähigkeit gegen andere Infekte zu- oder abnimmt, haben wir Lernen II. Wenn es durch einen solchen Infekt zu einer Besserung oder Verschlechterung einer chronischen Krankheit kommt, liegt Lernen III vor. Lernen IV wären die epigenetischen Folgen einer solchen Infektion. Das wesentliche Problem ist, dass die derzeitige Medizin nur Lernen 0. und 1. Ordnung kennt. Lernen höherer Ordnung, obwohl absolut spezifisch, wird als unspezifischer Effekt aus dem Beobachtungsrahmen eliminiert (15).

Randomisierte Studien bei Naturheilverfahren

Randomisierte Studien sind so konzipiert, dass sie die Wirksamkeit einer spezifischen Intervention in einem definierten Setting gegen eine Scheinintervention prüfen. Das entspricht aber nicht dem Setting der Naturheilverfahren. Das Schlagwort heißt Individualisierung. Ich möchte dies am folgenden Beispiel verdeutlichen: Bereits eine einfache Massage muss, um erfolgreich zu sein, bei jedem Patienten anders gestaltet und dosiert werden. Was dem einen nützt, kann dem anderen durchaus schaden. Was ist nun von einer Studie zu halten, die den Effekt einer standardisierte Physiotherapie (Massage, Wärme, Kälte u.a.) mit der einfachen standardisierte Anweisung vergleicht, weiter aktiv zu bleiben (13)? Das Ergebnis, dass die Physiotherapie dem einfachen Ratschlag nicht überlegen ist, mag wohl wenig Relevanz besitzen. Viel entscheidender ist jedoch, dass beide Vorgehensweisen (standardisierte Therapie und standardisierter Ratschlag) schlechte Medizin sind. Man könnte für diese Form der Wirksamkeitsprüfung in den Naturheilverfahren folgende Faustregel aufstellen: Eine Studie ist in ihrer Methodik umso besser, je schlechter die praktizierte Medizin ist. Im Grunde messen die üblichen Studien eine Form der naturheilkundlichen Therapie, die nicht den tatsächlichen Gegebenheiten entspricht. Das Konzept des ‚spezifischen‘

Wirkfaktors in einer ‚spezifische‘ Situation ist bei Verfahren, die meist eine unspezifische Aktivierung der Eigenregulation anstreben, weitgehend unangebracht. Formallogisch können Studien nur zu Aussagen über das führen, was sie prüfen. Die besagte Physiotherapiestudie belegt demnach nur, dass diese Form der standardisierten Therapie einem einfachen Rat nicht überlegen ist. Auch wenn potenzierte Hausstaubmilbe bei Hausstauballergie kein besseres Ergebnis als ein Placebo zeigt (17), so sagt das weder etwas über potenzierte Substanzen aus, noch etwas über die Homöopathie. Genau so wenig ist eine Studie, die keinen positiven Effekt für Acetylsalicylsäure bei Brustkrebs findet (9), in der Lage, die Wirksamkeit von Acetylsalicylsäure generell zu beurteilen, und erst recht gibt sie keine Aufschlüsse über die Effektivität der chemischen Arzneimitteltherapien insgesamt. Dennoch sind solche Fehlschlüsse an der Tagesordnung. In der Logik wird dies als ‚Verletzung der logischen Typisierung‘ bezeichnet (19). Wenn eine Klasse (Heilverfahren) mit dem Element der Klasse (spezifische Intervention) in die gleiche Kategorie gepackt wird, so entsteht eine paradoxe Situation, die keine verlässliche Aussage mehr zulässt. Die Vorstellung, man könne ein ganzes Verfahren aufgrund einer Studie prüfen, ist daher ein erhebliches Problem.

Natürlich wurde versucht Studien dem naturheilkundlichen Setting anzupassen. Beispielsweise erfüllte die Münchner Kopfschmerzstudie (23) die a priori Voraussetzungen von Statistiker und Homöopathen. Langjährige Kopfschmerzpatienten (durchschnittliche Krankheitsdauer von 23 Jahre) wurden individualisiert behandelt. Das Ergebnis war, dass das individuell bestimmte homöopathische Mittel keine bessere Wirkung aufwies als das Placebo. Natürlich wurde im Nachhinein (wie immer wenn eine Studie nicht das gewünschte Ergebnis zeigt) von homöopathischer Seite eine ganze Reihe von Einwänden gemacht. Die Studie sei zu kurz gewesen, die Fälle für das einfache Schema der Studie zu schwer, usw. Es in diesem Zusammenhang gleichgültig, ob diese Argumente stichhaltig sind. Wesentlich ist, dass die homöopathische Therapie an sich (Verum und Placebo) bei fast einem Viertel der Patienten wirksam war. Dies ist insofern bemerkenswert, als eine ganze Reihe von Studien zur Akupunktur zu einem ganz ähnlichen Ergebnis kamen (6, 18, 20): Placebo- und Verumtherapie sind beide effektiv, aber die Verumintervention ist dem Placebo nicht oder nur geringgradig überlegen. Das

heißt, Homöopathie und Akupunktur sind effektive Therapieformen, die in der Lage sind langjährige Schmerzzustände kostengünstig und nebenwirkungsarm zu kontrollieren. Sie machen es aber nicht „spezifisch“.

Nach der üblichen Logik sollte man fordern, die medikamentöse Standardtherapie aus dem Leistungskatalog der gesetzlichen Krankenkassen zu streichen, da sie schlechter abschneidet als eine unspezifische Placebothherapie! Natürlich würde eine solche Forderung niemand ernstnehmen. Sie ist auch unsinnig. Denn sie basiert auf Fehlschlüssen. Aber auch die umgekehrte Schlussfolgerung, nämlich die der Unwirksamkeit der Naturheilverfahren, ist genau so unzulässig. Dass Therapien wirksam sind obwohl sie unwirksam sind, ist ein typisches Phänomen, das dann auftritt, wenn die logische Typisierung verletzt wird, wenn also die Klasse (Verfahren) mit dem Element der Klasse (spezifische Therapie) verwechselt wird.

Das heißt, die Ergebnisse dieser Studien beweisen weder die Wirksamkeit noch die Unwirksamkeit der Naturheilverfahren. Sie beweisen allenfalls, dass diese Art von Wirksamkeitsmessung aus logischen Gründen nicht geeignet ist, zu einem soliden Ergebnis zu kommen.

Schlussfolgerung

Randomisierte Studien gelten als der Goldstandard der evidenzbasierten Medizin. Doch trotz des immensen Siegeszuges der EbM gibt es eine Vielzahl von Bedenken, die auf die Beschränktheit dieses Ansatzes hinweisen. Grundlegende Themen wie Adaptation, individuelle Übertragbarkeit, Beurteilung komplexer und/oder langfristige Prozesse können mit dieser Form der Wirksamkeitsmessung nicht oder nur unzureichend beurteilt werden.

Randomisierte Studien sind das geeignete Mittel für eine Medizin, die im Wesentlichen eine Symptombeseitigung anstrebt. Für regulative Prozesse, bei denen ein passender Reiz für ein Individuum gewählt wird, ist das Hauptverfahren der EbM wenig tauglich, denn in so genannten Netzwerkpathologien kann derselbe Reiz eine Besserung oder Verschlechterung auslösen. Die Frage, ob das eine oder andere Mittel oder Verfahren bei Kopfschmerz besser wirkt, ist dabei eher von zweitrangiger Bedeutung.

Dieses völlig andere Verständnis von Gesundheit und Krankheit findet sich nicht nur bei den Naturheilverfahren. Die neueren Forschungen zur Immunologie weisen ja allenthalben darauf hin, dass das hergebrachte Krankheitsverständnis, das auf simplen pathophysiologischen Beschreibungen oder gar auf dem Abzählen von Kriterien beruht, ein Auslaufmodell ist.

Wie die Wirksamkeitsprüfung in einem solchen regulativen Gesundheitsmodell (im Gegensatz zu einem pathophysiologischen Krankheitsmodell) aussehen könnte, ist schwer zu sagen. Sie muss meiner Ansicht nach ein dynamisches Verfahren sein, das Änderungsprozesse modellieren kann. Ob man dann eine solche Form der Wirksamkeitsprüfung, die sich grundlegend von der heutigen unterscheiden wird, weiterhin EbM nennen mag, ist Geschmacksache. Ich persönlich tendiere eher dazu, ein moderneres Verständnis von Medizin nicht mit alten Begriffen zu belasten.

Literatur

- (1) Baker SG, Kramer BS (2002): *The transitive fallacy for randomized trials: If A bests B and B bests C in separate trials, is A better than C?* BMC Medical Research Methodology 2002 2:13, available <http://www.biomedcentral.com/1471-2288/2/13>
- (2) Baker SG, Kramer BS (2002a): *Good for Women, Good for Men, Bad for People: Simpson's Paradox and the Importance of Sex-Specific Analysis in Observational Studies.* Journal of Women's Health & Gender-Based Medicine 10; 9: 867 - 872
- (3) Bateson G (1990): *Ökologie des Geistes*, Suhrkamp, Frankfurt, S: 362-399
- (4) Beck-Bornholdt HP, Dubben HH: *Der Schein der Weisen.* Reinbek bei Hamburg: Rowohlt 2003
- (5) Benn CS, Melbye M, Wohlfahrt J, Björkstén B, Aaby P (2004): *Cohort study of sibling effect, infectious diseases, and risk of atopic dermatitis during first 18 months of life,* BMJ 328:1223
- (6) Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC (2004): *Effectiveness of Acupuncture as Adjunctive Therapy in Osteoarthritis of the Knee,* Annals of Internal Medicine 141; 12: 901-910
- (7) Bogomolny A (1996-2003): *Mediant fractions,* available <http://www.cut-the-knot.org/blue/Mediant.shtml#simpson>
- (8) Bouchama A, Knochel JP (2002): *Heat stroke,* N Engl J Med 346: 1978 - 1988
- (9) Cook NR, Lee IM, Gaziano JM, et. al (2005): *Low-Dose Aspirin in the Primary Prevention of Cancer: The Women's Health Study: A Randomized Controlled Trial,* JAMA 294: 47-55
- (10) Dellmour F (2006): *Klinische Studien und Metaanalysen in der Homöopathie,* Deutsche Zeitschrift für klinische Forschung 5/6: 52-60, available http://www.dzkg.de/heft/2006_05-06/12.pdf
- (11) Detillion CE, Craft TKS, Gasper ER, Prendergast BJ, DeVries AC: (2004) *Social facilitation of wound healing,* Psychoneuroendocrinology 29; 8: 1004-1011
- (12) ExTraMATCH Collaborative (2004): *Exercise training meta-analysis of trials in patients with chronic heart failure (ExTraMATCH),* BMJ 328:189
- (13) Frost H, Lamb SE, Doll HA, Carver PT, Stewart-Brown S (2004): *Randomised controlled trial of physiotherapy compared with advice for low back pain,* BMJ 329:708
- (14) Garrett M, McElroy AM, Staines A. *Locomotor milestones and babywalkers: cross sectional study.* BMJ 2002;324:1494
- (15) Ivanovas G, Tomaras V, Paritsis (2007): *Human adaptation and conscious purpose in contemporary medicin,* Kybernetes, to be published 2007
- (16) Johnson CC, Ownby DR, Alford SH, Havstad SL, Williams LK, Zoratti EM, Peterson EL, Joseph CL (2005): *Antibiotic exposure in early infancy and risk for childhood atopy,* J Allergy Clin Immunol. 115; 6:1218-24.
- (17) Lewith GT, Watkins AD, Hyland ME, et. al. (2002): *Use of ultramolecular potencies of allergen to treat asthmatic people allergic to house dust mite: double blind randomised controlled clinical trial.* BMJ 324:520
- (18) Linde K, Streng A, Jürgens S, Hoppe A, et. al.(2005): *Acupuncture for Patients With Migraine,* JAMA. 293:2118-2125
- (19) Russell, Bertrand (1930): *Introduction to Mathematical Philosophy.* Allen&Unwin, London
- (20) Scharf HP, Mansmann U, Streitberger K, Witte S, Krämer J, Maier C, Trampisch HJ, Victor N, PhD (2006): *Acupuncture and Knee Osteoarthritis,* Annals of Internal Medicine 145; 1: 12-20
- (21) Singhai A, Fewtrell M, Cole TJ, Lucas A (2003): *Low nutrient intake and early growth for later insulin resistance in adolescents born term.* Lancet 361: 1089-97

- (22) Smidt N, van der Windt DAWM, Assendelft WJJ, Devillé WLJM, Korthals-de Bos IBC, Bouter LM (2002): *Corticosteroid injections, physiotherapy, or a wait-and-see policy for lateral epicondylitis: a randomised controlled trial*. *Lancet* 2002; 359: 657-62
- (23) Walach H, Haeusler W, Lowes T, et. al. (1997): *Classical homeopathic treatment of chronic headaches*, *Cephalalgia* 1997; 17: 119-126

Georg Ivanovas, Milatos, GR72400 Neapoli, ivanovas@gmx.net