



Health Technology Assessment (HTA) in der österreichischen Sozialversicherung

1 Einleitung

In dem Artikel „Evidence-based-Medicine-Orientierung: EBM und ihre Kritiker“¹, erschienen in der Zeitschrift Soziale Sicherheit im April 2012, wurde bereits auf den Zusammenhang zwischen „wissenschaftlichem Nachweis“ und einer Beurteilung der Angemessenheit für das System der sozialen Krankenversicherung eingegangen. Dabei lag ein Schwerpunkt der Betrachtung auf den ethischen Aspekten.

In der laufenden Arbeit ergibt sich nunmehr die Notwendigkeit, die Methodik und das Vorgehen einer Beurteilung von „Technologien“ auf der Ebene des Gesundheitssystems mit besonderer Berücksichtigung von erkenntnistheoretischen und rechtlichen Gesichtspunkten zu beleuchten. Dabei ergibt sich die Gelegenheit, die Erfahrungen des Hauptverbandes der österreichischen Sozialversicherungsträger als Partner im Europäischen Netzwerk für Health Technology Assessment – EUNetHTA² – nutzbar zu machen und zu berichten, wie dieses

Thema international gesehen und bearbeitet wird. Als Beispiele aus der laufenden Arbeit werden dabei die Aufarbeitung des Themas „physikalische Therapie“³ und die Erfahrungen mit einem laufenden Gerichtsverfahren wegen einer Bewertung einer Behandlungsmethode für Krebspatienten⁴ herangezogen. Damit kann das Spannungsfeld zwischen den Interessen der Anbieter und einer methodisch nachvollziehbaren Bewertung als eine Grundlage für Entscheidungsträger im Gesundheitssystem beleuchtet werden.

Im Hauptverband der österreichischen Sozialversicherungsträger wurde HTA im Jahre 2002 als eigener Aufgabenbereich institutionalisiert. Die derzeitige Arbeitsstruktur und die Methodik gibt es seit 2004, und auch das Arbeitsteam ist – mit wenigen personellen Veränderungen – bereits seit damals tätig. Damit bietet das „zehnjährige Jubiläum“ ebenfalls Anlass zu einer generellen Betrachtung und Reflexion über bisher Geleistetes und Gelerntes. Einen Überblick bietet die Internetseite www.hauptverband.at/EBM_HTA



Dr. Gottfried Endel ist Allgemeinmediziner und leitet den Bereich „Evidence-based Medicine“ (EBM) im Hauptverband der österreichischen Sozialversicherungsträger.

1 Evidence-based-Medicine-Orientierung: EBM und ihre Kritiker. http://www.hauptverband.at/portal27/portal/hvbportal/channel_content/cmsWindow?action=2&p_menuid=73019&p_tabid=2&p_pubid=653973 (abgefragt am 2.1.2014).

2 <http://www.eunetha.eu/>.

3 Quellen: Wilbacher, I., Brandstätter, S., Kruzic, N., Gerner, G., Kiesel, U.: Österreichischer Muster-Leistungskatalog Physikalische Medizin – empirische Grundlagen zur Wirksamkeit der Inhalte. GMS Health Technology Assessment 2013, 9, Doc04. <http://www.egms.de/static/en/journals/hta/2013-9/hta000110.shtml> (abgefragt am 2.1.2014); [wenn möglich, hier neue Zeile beginnen:] Evidenzbasierte Medizin in der Physikalischen Medizin und allgemeinen Rehabilitation. European Journal of Translational Myology – Basic Applied Myology 2013, 23 (4): 131–136.

4 Wegen des laufenden Verfahrens werden hier keine Quellen genannt. Die Erkenntnisse werden mit Bezug auf die Methodik dargestellt.

2 Was ist Health Technology Assessment?

Wie in vielen Bereichen der Wissenschaft sind Definitionen heute international vereinbart und meist auf Englisch verfügbar.

Die Definition im sogenannten HTA Glossary der Gesellschaft „Health Technology Assessment international – HTAi“⁵ lautet: „*The systematic evaluation of the properties and effects of a health technology, addressing the direct and intended effects of this technology, as well as its indirect and unintended consequences, and aimed mainly at informing decision making regarding health technologies. Note: HTA is conducted by interdisciplinary groups that use explicit analytical frameworks drawing on a variety of methods.*“

Ähnlich lautet die Definition beim „International Network of Agencies for Health Technology Assessment – INAHTA“⁶: „*Healthcare technology is defined as prevention and rehabilitation, vaccines, pharmaceuticals and devices, medical and surgical procedures, and the systems within which health is protected and maintained. Technology assessment in health care is a multidisciplinary field of policy analysis. It studies the medical, social, ethical, and economic implications of development, diffusion, and use of health technology.*“

Aus diesen beiden Definitionen sind folgende Eigenschaften und Aufgaben von HTA ersichtlich:

- interdisziplinäres/multidisziplinäres Arbeiten;
- systematisches methodisches Vorgehen;
- umfassende Berücksichtigung von Effekten und Einflüssen;
- umfassende Sicht von Technologie – von einzelnen Produkten bis zu Aspekten der Organisation eines Gesundheitssystems;
- Erstellung von Informationen für Entscheidungsträger.

Dies findet auch in der österreichischen nationalen HTA-Strategie 2009⁷, an welcher Vertreter des Hauptverbandes mitgearbeitet haben, seinen Ausdruck. Dort wird auf Seite 3 (Seite 13 des PDF-Files) Folgendes festgehalten: „*HTA steht den Entscheidungs- und Finanzierungsträgern als ein unterstützendes Werkzeug neben anderen Grundlagen zur Verfügung. Ein gesundheitspolitisches Ziel ist es, mit Hilfe von HTA den Patientinnen und Patienten im Rahmen des verfügbaren öffentlichen Budgets ein optimales Leistungsangebot zugänglich zu machen. Die relevanten Entscheidungs- und Finanzierungsträger mit einzubeziehen ist wichtig nicht nur,*

damit es überhaupt zu einem systematischen Einsatz von HTA im Gesundheitswesen kommt, sondern auch im HTA-Erstellungsprozess (z. B. Themenfindung, Präzisierung der Fragestellung) selbst.“

Für das EBM/HTA-Team des Hauptverbandes leiten sich damit die Rahmenbedingungen für Beurteilungen aus der gesetzlichen Aufgabenstellung der österreichischen Sozialversicherung ab. Dies und auch die methodischen Vorgangsweisen sind auf der Einstiegsseite der Internetpräsenz publiziert.⁸ Um dem Leser dieses Artikels den Vergleich zwischen den Definitionen der internationalen Fachgesellschaften, der nationalen HTA-Strategie 2009 und der Ausrichtung und Vorgangsweise des Hauptverbandes zu erleichtern, wird ein Teil hier zitiert:

„*Bei unserer Arbeit halten wir uns im Interesse der Patienten an folgende Maximen:*

- *Unsere Arbeit wird entsprechend den Rahmenbedingungen des ASVG durchgeführt.*
- *Der Wissensgewinn erfolgt weisungsunabhängig und frei von parteilichen Einflussnahmen. Unsere Mitarbeiter zeichnen sich durch ihre fachliche Offenheit aus.*
- *Wir definieren uns als interner Dienstleister im Gesundheitssystem und erleichtern durch unseren kundenfreundlichen Service die fachliche Kommunikation nach innen und außen.*
- *Unsere Vorgangsweise ist nachvollziehbar und steht für höchste Zuverlässigkeit.*
- *Während unsere Arbeit auf publizierten wissenschaftlichen Erkenntnissen basiert, versuchen wir so weit wie möglich, auch noch unveröffentlichtes Wissen in Anspruch zu nehmen. Dazu haben wir ein internationales Netzwerk an Kontakten und nehmen an entsprechenden Kooperationen teil.*“

Hier wird klargestellt, dass die Interessen der Patienten als Solidargemeinschaft im Sinne des ASVG gewahrt werden und nicht der Anspruch erhoben wird, die Behandlung einzelner Individuen im Sinne der klinischen Medizin zu untersuchen. Die Ergebnisse dienen der Entscheidungsunterstützung innerhalb der Sozialversicherung und in weiterer Folge des gesamten Gesundheitssystems in Österreich. Damit wird auch nachvollziehbar gemacht, dass die Themenauswahl und Priorisierung von Entscheidungsträgern der Sozialversicherungsträger ausgeht. Dies war auch bei den gewählten Beispielen der Fall, womit sie gut geeignet sind, den HTA-Prozess innerhalb der österreichischen Sozialversicherung darzustellen.



Mag. Dr. Ingrid Wilbacher ist Mitarbeiterin des Bereiches „Evidence-based Medicine“ (EBM) im Hauptverband der österreichischen Sozialversicherungsträger.

5 <http://htaglossary.net/health+technology+assessment+%28HTA%29> (abgefragt am 2.1.2014).

6 <http://www.inahta.org/HTA/> (abgefragt am 2.1.2014).

7 <http://www.goeg.at/de/BerichtDetail/Nationale-HTA-Strategie198.html> (abgefragt am 2.1.2014).

8 http://www.hauptverband.at/portal27/portal/hvbportal/channel_content/cmsWindow?action=2&p_menuid=72429&p_tabid=5 (abgefragt am 2.1.2014).

Abbildung 1: Kriterien für die Katalogwartung

Aufnahmekriterien	
Mindestfrequenz pro Jahr	5x pro 10.000 Einwohner (niedergelassen) bzw. 3x pro 10.000 Einwohner (spitalsambulant)
Großgerätespezifische Leistungen	Erbringung nur an Spezialeinrichtungen mind. 12x im Jahr Umsatz pro Jahr > 24.000 Euro
Medizinisch relevante Leistungen	Kosten > 100 Euro mind. 1,25x pro Jahr pro 10.000 Einwohner
-> Mindestens eines der genannten Einschlusskriterien muss erfüllt sein.	
Ausschlusskriterien	
Keine Aufnahme von Medikamenten	analog stationärem Katalog
Keine Aufnahme von Bagatelleleistungen/Einzelhandgriffen	
Verhältnismäßigkeit	Kosten < 10 Euro
Verletzung von Katalogprinzipien	z. B. nicht für sich allein erbringbare Leistung
negatives HTA-Ergebnis	
-> Keines der Ausschlusskriterien darf erfüllt sein, sonst keine Aufnahme in den Katalog (K.o.-Prinzip).	

http://bmg.gv.at/cms/site2/attachments/1/3/0/CH1064/CMS1240821423857/dokuambulant_leistungskatalog_bericht_2012_stand_pla_20130123_eb_hp-bmg_korr.pdf

Einzelne Leistungen und umfassende Leistungskataloge werden untersucht.

3 Themenwahl

Die gewählten Beispiele repräsentieren zwei sehr unterschiedliche Informationsnotwendigkeiten innerhalb der Sozialversicherung.

Im Falle der Behandlung im Bereich der Onkologie entstand die Frage durch einen Leistungsantrag bei einem Krankenversicherungsträger. Da eine Kostenerstattung in der gewünschten Höhe einerseits nicht nachvollziehbar und andererseits von den tariflichen Rahmenbedingungen her nicht möglich war, ergab sich die Notwendigkeit, die Zweckmäßigkeit der Intervention genau zu untersuchen. Damit wird versucht, eine für eine Leistungsentscheidung und ein eventuell folgendes Leistungsstreitverfahren ausreichende Informationsbasis für eine bisher unbekannte, eventuell neue, seltene und besonders teure Therapie bereitzustellen. In diesem Fall liegt die Schwierigkeit im Mangel an publizierten Informationen und im Zeitdruck für eine rasche Bearbeitung des Leistungsantrages durch den Krankenversicherungsträger.

Die Überarbeitung des Musterkataloges für physikalische Medizin hingegen ist ein Beispiel für die systematische Betreuung und Wartung des Gesundheitssystems. Ein derartiger Prozess für die Wartung der in den Honorarordnungen der Gesamtverträge für ärztliche Hilfe enthaltenen Leistungspositionen wurde innerhalb der Sozialversicherung in unterschiedlicher Ausprägung schon immer betrieben. Instrumente dafür sind der Tarifausschuss als Beratungsgremium für die Träger

und den Hauptverband sowie Leistungskataloge für bestimmte thematische Bereiche. Die Wartung der Leistungskataloge ergibt sich als Aufgabe vor allem daraus, dass der Hauptverband für den Abschluss der Gesamtverträge für die Sozialversicherungsträger mit den jeweiligen Landesvertretungen über ärztliche Behandlung, Mutter-Kind-Passleistungen, Arbeitsmedizin, Vorsorgeuntersuchungen, Heilmittelversorgung, Psychotherapie usw. verantwortlich ist.

Insbesondere für den Bereich von Leistungskatalogen wurden methodische Grundlagenarbeiten durch eine Zusammenarbeit mit dem Ludwig Boltzmann Institut für HTA (LBI-HTA) geschaffen. Grundsätze für die Katalogwartung⁹ sowie die speziellen Fragestellungen im Zusammenhang mit diagnostischen Maßnahmen¹⁰, welche zunehmend den Umfang der Leistungskataloge dominieren, wurden ausgearbeitet und können nun die Wartung von Leistungskatalogen in Österreich unterstützen. Dies ist am Beispiel des Kataloges ambulanter Leistungen (KAL) erkennbar, welcher in einer Zusammenarbeit zwischen dem Bundesministerium für Gesundheit (BMG), den Bundesländern sowie der österreichischen Sozialversicherung entwickelt wurde.

Ein fehlender Wirkungsnachweis im Sinne eines „negativen HTA-Ergebnisses“ stellt ein Ausschlusskriterium dar und verhindert eine Aufnahme einer neuen Leistung in den Katalog, bei einer „alten“ Leistung ist die Entfernung aus dem Katalog anzustreben. In diesem Kontext ist auch die Beurteilung der Wirksamkeit von physikalisch-therapeutischen Maßnahmen zu sehen. Der Musterkatalog wurde als Unterstützung für die Krankenversicherungsträger erarbeitet und sollte auch zu einer vergleichbaren Vertragsgestaltung im Bereich der ärztlichen Hilfe beitragen.

4 Methodik

Der Prozess der Katalogwartung wurde dabei erstmals umfassend durch eine methodische Erhebung der Wirksamkeitsnachweise begleitet. Im Vergleich zur Beantwortung eines einzelnen Leistungsantrages ein umfangreiches Projekt, da viele Therapieformen für viele Indikationen zu untersuchen waren. Die Abarbeitung einer Matrix mit mehr als 10.000 möglichen Einzelfragen verlangte die Anpassung der Methodik zur Bearbeitung von Einzelfragen für diese Fülle von Fragestellungen. Auch die für die Bearbeitung benötigte Zeit und Anzahl der Mitarbeiter ist mit der Erstellung von Assessments für einzelne Technologien – Sin-

9 Breyer, E. (2008): Aktualisierung von Leistungskatalogen. Eine vergleichende Analyse von Refundierungsprozessen für ärztliche Leistungen in ausgewählten Ländern. HTA-Projektbericht 022. <http://eprints.hta.lbg.ac.at/817/> (abgefragt am 2.1.2014).

10 Nachtnebel, A. (2010): Evaluation diagnostischer Technologien – Hintergrund, Probleme, Methoden. HTA-Projektbericht 36. <http://eprints.hta.lbg.ac.at/898/> (abgefragt am 2.1.2014).

Der Prozess der Katalogwartung wurde dabei erstmals umfassend durch eine methodische Erhebung der Wirksamkeitsnachweise begleitet. Im Vergleich zur Beantwortung eines einzelnen Leistungsantrages ein umfangreiches Projekt, da viele Therapieformen für viele Indikationen zu untersuchen waren.

gle Technology Assessment – nicht vergleichbar. Die Vorgangsweise ist in den Dokumenten „Vorgangsweise Anfragebearbeitung“ und „Handbuch für EBM-Berichte“ veröffentlicht.¹¹ Für das Thema aus dem Bereich der Onkologie war eigentlich eine Bearbeitung nach der Methode des Health Evidence Network (HEN) der WHO vorgesehen. Mangels geeigneter Sekundärliteratur und da sich auch nur sehr wenig primäre Literatur fand, musste – im Sinne eines Systematic Review – alles herangezogen werden, was auffindbar war.

Für die Wartung des Musterkataloges wurde hingegen von vornherein ein Systematic Review als Methode und nur für Teilbereiche mit ausreichend Sekundärliteratur der HEN-Level-Ansatz gewählt. Dazu wurden die verschiedenen Sozialversicherungsträger operativ und projektbegleitend einbezogen und ein multidisziplinäres Team in der systematischen Literaturlaufbereitung eingesetzt. Da durch eine Katalogwartung Leistungsanbieter jedenfalls betroffen sind, versuchten diese ihre Interessen und ihre Sichtweise entsprechend zu positionieren. Dies einerseits durch den grundsätzlich begrüßenswerten Versuch, Belege für die Wirksamkeit von Therapieformen vorzulegen, und andererseits mit einer die Vorgangsweise des HTA grundlegend ablehnenden Argumentation.¹² Die Argumentation für den Wirkungsnachweis entspricht allerdings der in der Literatur (siehe Fußnote 14) in Tabelle 1 als Nummer 1 gelisteten Vorgangsweise: „*1 Selective, non-systematic reviews. Cite just those reports that support your product, proposal, or policy (and slag all your competitors).*“

Daher wurde das Ergebnis der Katalogwartung einem externen Peer-Review-Verfahren unterzogen und publiziert.¹³

5 Methodenkritik

5.1 Einreichung einer Klage

Im Falle des Single Technology Assessments in der Onkologie erfolgte eine Klage auf Unterlassung der Veröffentlichung letztendlich eines einzelnen Satzes des Assessments, welcher eigentlich Bestandteil einer zusammenfassenden Darstellung der Beschrei-

bung der Methode auf der Website des Leistungsanbieters ist. Diese Website wurde dabei als Quelle ausgewiesen. Die Zusammenfassung, so der Vorwurf, verletze das Persönlichkeitsrecht und das Recht am ausgeübten Gewerbebetrieb und sei kreditsschädigend. Die Problematik eines fehlenden Wirkungsnachweises wird davon nicht berührt, es wird damit nur versucht, diese Information zu unterdrücken.

Der weitere Verlauf des Verfahrens wird zeigen, ob eine Kurzdarstellung einer Information mit Angabe der Quelle geeignet ist, den Ersteller und Eigentümer der Information zu schädigen. Es spricht prinzipiell nichts gegen eine Übernahme des gesamten Originaltextes. Dabei ist nur zu bedenken, dass dann eine Urheberrechtsverletzung beklagt werden könnte und die Lesbarkeit von derart erstellten Dokumenten deutlich eingeschränkt wird. Die Kerninformation des Assessments, dass nämlich bisher kein Wirkungsnachweis für die Methode erbracht wurde und diese trotzdem teuer vermarktet wird, bleibt vom Inhalt des beklagten Satzes allerdings unberührt.

Weitere Besonderheiten können darin gesehen werden, dass die Unterlassung der Veröffentlichung (im Internet) beschränkt auf das Gebiet der Bundesrepublik Deutschland ausgesprochen wurde. Auch das Verfahren findet in der Bundesrepublik Deutschland statt. Grundsätzlich ist diese Vorgangsweise auch schon 2003 in der Literatur beschrieben¹⁴ und dort in Tabelle 2 folgendermaßen dargestellt worden: „*SHARKS (Striking Horror And Retreat through Killer Solicitors). Hire all the really good lawyers. Use them to threaten nay-sayers, members of drug review boards, etc with frivolous but expensive libel lawsuits. Suppress negative health technology assessment reports until you've met your sales targets.*“

5.2 Kritik an der „Evidenzbasierung“ von Entscheidungen

Diese Form einer Kritik wird immer wieder vorgebracht. Als Beispiel einer philosophisch, ethisch argumentierten Kritik sei auf den Anlass des in der Einleitung genannten Artikels aus 2012¹⁵ verwie-

Unerwünschte Informationen können mit verschiedenen Methoden unterdrückt werden.

¹¹ http://www.hauptverband.at/portal27/portal/hvbportal/channel_content/cmsWindow?action=2&p_menuid=72429&p_tabid=5 (abgefragt am 2.1.2014).

¹² Siehe Literaturangabe in Fußnote 3.

¹³ Siehe ebenfalls Fußnote 3 mit Angabe der URL.

¹⁴ Sackett, D. L., Oxman, A. D.: HARLOT plc: an amalgamation of the world's two oldest professions. *BMJ* 2003, 327: 1442–1445.

¹⁵ Die Anwendung von Evidence-Based-Medicine (EBM) zur Priorisierung von Gesundheitsleistungen in Österreich als ethisches Problem. Ein Positionspapier von Univ.-Prof. em. Dr. Günter Virt, a. o. Prof. DDr. Hans Schelkshorn et al.



© apops - Fotolia.com

Sichtweise der Patienten, Erfahrung der Behandler und wissenschaftliche Belege stehen gleichberechtigt nebeneinander.

sen. Die wissenschaftstheoretischen Überlegungen wurden als Antwort und zur Begründung der Notwendigkeit einer Evidenzbasierung von Entscheidungen dargestellt.

Die Autorengruppe aus dem Bereich der physikalischen Medizin argumentiert vor allem damit, dass für (klinische) Entscheidungen drei Informationsquellen heranzuziehen sind:

- die „individuelle klinische Erfahrung des Arztes“,
- die „Patientenbedürfnisse“ und
- die beste „externe Evidenz“.

Die gewählte Reihenfolge impliziert mit der weiteren Argumentation möglicherweise auch eine Wertung im Sinne einer Priorität. Nimmt man diese drei Säulen in einer Metapher als Stützen eines Tisches, so wird klar, dass die Tischplatte nur dann horizontal sein wird, wenn alle drei Kategorien im gleichem Maße berücksichtigt werden. Damit besteht hier vollständige Übereinstimmung, dass nämlich eine umfassende Patienteninformation und in der Folge ein „Informed Consent“ des Patienten ebenso wie die Fachexpertise eines Gesundheitsdiensteanbieters (GDA), aber auch ein ausreichender wissenschaftlicher Nachweis der Zweckmäßigkeit für eine individuelle Therapieentscheidung erforderlich sind.

Damit ist auch erkennbar, wo ein Unterschied zwischen Evidence-based Medicine (EBM) und HTA entsteht. Entscheidungsträger auf der Ebene des Gesundheitssystems ist nicht eine einzelne Person, sondern es sind Entscheidungen im Interesse der Versicherten als Solidargemeinschaft zu treffen.

Bleibt man beim Bild der „EBM-Triade“, so wird die „individuelle klinische Erfahrung des (einzelnen) Arztes“ durch multiprofessionelle Erfahrungen und Sichtweisen auf das Gesundheitssystem ersetzt werden müssen. Der Beitrag von Fachgesellschaften, Fachgruppen und (gesetzlichen) Interessensvertretungen aus dem Bereich der GDA steht dabei gleichberechtigt neben den Beiträgen anderer Fachdisziplinen.

Patientenbedürfnisse und -Präferenzen werden durch Zielsetzungen der Solidargemeinschaft – oder umfassender der Gesellschaft – ersetzt. Wie hier die Gewichtung und der Ausgleich zwischen der Sichtweise Betroffener verschiedener Krankheiten, von GDA und „Gesunden“ ist dem Bereich der politischen Meinungsbildung und Entscheidungsfindung zuzuweisen.

Einzig unverändert verbleibt die Frage der besten „externen“ oder richtiger „wissenschaftlichen“ Evidenz für die Abschätzung der Folgen einer Entscheidung. Daher resultiert möglicherweise auch der Eindruck einer Dominanz dieser Säule aus dem Bereich der EBM im Bereich der HTAs.

Neben der Argumentation, dass mehrere Informationsquellen herangezogen werden müssen – welcher wie dargestellt angepasst auf die Aufgaben des HTAs vollständig gefolgt wird – wird auch ein Versuch einer „Beweislastumkehr“ unternommen. Dies macht es erforderlich, nochmals auf wissenschaftstheoretische Grundlagen zurückzukommen. Teilweise wird dabei die Argumentation von 2012 übernommen. Sie wird aber durch zusätzliche Darstellungen der Problematik einer Feststellung von Ursache-Wirkungs-Beziehungen ergänzt. In der weiteren Folge werden einige der 2012 nur angeschnittenen mathematischen Grundlagen graphisch dargestellt, da die Erfahrungen aus der Vermittlung der notwendigen Grundkenntnisse gezeigt haben, dass diese vielen Personen intuitiv nicht zugänglich sind. Die Beweislastumkehr wird dabei folgendermaßen dargestellt: *„Eine bisher fehlende Evidenz für einzelne Therapieformen beweist nicht deren Unwirksamkeit!“*¹⁶ *„Absence of evidence is not evidence of absence. [1] Eine fehlende externe Evidenz für einzelne Therapieformen ist nicht der Beweis für deren Unwirksamkeit. Trotzdem ziehen viele Verantwortliche diesen falschen Schluss. [4, 29]“*¹⁷

Zu den Begriffen „Wissen“, „Erfahrung“ und „Erfahrungswissen“ zur Klarstellung ein Zitat aus 2012: *„Dies würde nun zu der Frage führen, was*

16 Evidenzbasierte Medizin in der Physikalischen Medizin und allgemeinen Rehabilitation. European Journal of Translational Myology – Basic Applied Myology 2013, 23 (4): 132.]

17 Ebenda, 134. Die Literaturangaben in eckigen Klammern beziehen sich auf:

[1] Alderson, P.: Absence of evidence is not evidence of absence. BMJ 2004, 328 (7438): 476–477.

[4] Brandstätter, S. (2009): Rückenleiden. Gesund werden. Gesund bleiben 2. Holzhausen Verlag GmbH.

[29] Wilbacher, I., Brandstätter, S., Kruzic, N., Gerner, G., Kiesel, U.: Österreichischer Muster-Leistungskatalog Physikalische Medizin – empirische Grundlagen zur Wirksamkeit der Inhalte. GMS Health Technology Assessment. 2013, 9: 1–14.

hier genau unter ‚Wissen‘, ‚Erfahrung‘ und ‚Erfahrungswissen‘ verstanden werden soll. Unter ‚Wissen‘ versteht man seit der Antike so etwas wie wahre, gerechtfertigte Behauptung oder Meinung. Eigentlich wissen wir aber nie, ob Wissen vorliegt – denn unsere Theorien können sich prinzipiell immer ändern. Darum spricht man in der Wissenschaft prinzipiell von ‚Hypothesen‘ und von ‚Arbeitshypothesen‘ (Popper). Eine Arbeitshypothese (oder bewährte Theorie) ist eine solche, die bislang nicht falsifiziert wurde – und somit vorläufig weiter genutzt werden kann. Es ist aber noch kein Wissen im strengen Sinn. Umgekehrt ist der Begriff ‚Erfahrungswissen‘ insofern undeutlich, als Wissen meistens auf Erfahrung zurückgeht, Wissen also meistens auch (wenn auch nicht ausschließlich) mit Erfahrung zu tun hat.“¹⁸

Auch Alderson stellt dies zutreffend fest: „When is it reasonable to claim that a study has proved that no effect or no difference exists? The correct answer is ‘never’, because some uncertainty will always exist. However, we need to have some rules for deciding when we are fairly sure that we have excluded an important benefit or harm.“¹⁹

Wenn also Feststellungen immer nur mit einer (statistisch bestimmbaren) Sicherheit getroffen werden können, so ist „Unwirksamkeit“ nicht beweisbar, sondern nur das Mislingen und Fehlen eines Wirkungsnachweises feststellbar. Besonders bei den als „klinische Erfahrung“ oft genannten Kasuistiken ist die Feststellung eines Kausalzusammenhanges besonders schwierig.

5.3 Kausalzusammenhang

Hypothesen und „Wissen“ sind damit immer nur „statistisch“ abgesichert. Die dadurch entstehende Unsicherheit für Entscheidungen (durch einen Patienten oder auf Systemebene) oder Empfehlungen (im Einzelfall oder generell) bereitet Unbehagen und führt zu unterschiedlichen Lösungsansätzen. Diese sind teilweise durch unklare Definition von Begriffen bedingt.

Zur Erläuterung sei auf Judea Pearls „Causality. Models, reasoning, and inference“²⁰ verwiesen. Pearl beschreibt hier verschiedene – mathematisch definierbare – Begriffe im Zusammenhang mit Kausalität (unter anderem):

- Probability of Necessity, PN: Ein Ereignis tritt nicht ein, außer eine Bedingung ist erfüllt.
- Probability of Sufficiency, PS: Wenn eine Bedingung erfüllt ist, wird auch das Ereignis eintreten.

In der Folge wird versucht, die in den Kapiteln 9 und 10 dazu enthaltenen Überlegungen zusammengefasst wiederzugeben.

Unterschieden wird dabei zwischen „aktuellen Ursachen“ für einen Einzelfall und „generellen Ursachen“ für Typen von Vorfällen. Diese wurden in der Philosophie getrennt und kontrovers diskutiert. Pearl schlägt mit der Methode des Structural Modelling eine Vorgangsweise vor, mit welcher beide Typen von Ursachen beschrieben werden können. Dazu wird „Counterfactual Reasoning“ – „den Tatsachen widersprechendes Argumentieren“, hypothetisches Argumentieren – zur Herstellung von Beziehungen zwischen den Geschehnissen verwendet. Die Beispiele aus „Causality“ sind folgende:²¹

- PN ist dabei vor allem eine Eigenschaft von Einzelfällen. Die Wahrscheinlichkeit, mit einem normalen Gewehr auf 1000 Meter ein Ziel zu treffen, ist gering. Wenn aber eine Person trotz dieser geringen Wahrscheinlichkeit auf diese Entfernung erschossen wird, so ist ein Schuss eine notwendige Voraussetzung.
- PN allein erklärt allerdings Ursache-Wirkungs-Beziehungen nur unvollständig. Würde PN allein betrachtet, so besteht die Möglichkeit, fälschlich allgemeine Umweltbedingungen als Ursachen zu definieren.
- PS ist daher eher für generelle Ursachen das geeignete Maß. Bricht in einem Raum ein Feuer aus, so ist eine Notwendigkeit das Vorhandensein von Sauerstoff. Ohne Sauerstoff kein Feuer – analog der Erklärung, ohne Schuss kein Erschossener. Das Entzünden eines Streichholzes oder das Umfallen einer Kerze werden hier aber die bessere ursächliche Erklärung abgeben. Denn Sauerstoff allein führt üblicherweise nicht zu einem Brand.

Ursache-Wirkungsbeziehungen nachzuweisen ist aufwendig.

Abbildung 2: Rahmenbedingungen für den Nachweis von Ursache-Wirkungs-Beziehungen

Exogeneity	Assumptions		Data Available		
	Monotonicity	Additional	Experimental	Observational	Combined
+	+	covariate control	ERR bounds	ERR bounds corrected ERR	ERR bounds corrected ERR
+	-		—	—	corrected ERR bounds
-	+		—	—	—
-	-		—	—	—

Note: ERR stands for the excess risk ratio. $1 - P(y|x)/P(y|x)$; corrected ERR is given in (9.31).

Aus: Pearl, J. (2000): Causality. Models, reasoning, and inference. Cambridge University Press, 344.

18 Siehe FN 1

19 Alderson, P.: Absence of evidence is not evidence of absence. BMJ 2004, 328 (7438): 476–477.

20 Pearl, J. (2000): Causality. Models, reasoning, and inference. Cambridge University Press. ISBN 0-521-77362-8 (hardback).

21 Ebenda 331.

Abbildung 3: Vergleich Schmerzbehandlung bei 30 % Wirksamkeit eines Placebos

	A	B		C		D		E		F		G		H		I		J	K
		Gruppe Placebo		Gruppe Intervention						CI									
1	Szenario	Personen	nicht gebessert	Personen	nicht gebessert	RRR	NNT	NNT min	NNT max	Chi Quadrat	Signifikant								
2	1	100	70	100	70	0%	#DIV/0!	-7,87	7,87	0,00	nein								
3	2	100	70	100	69	1%	100,00	-8,50	7,27	0,02	nein								
4	3	100	70	100	68	3%	50,00	-9,24	6,75	0,09	nein								
5	4	100	70	100	67	4%	33,33	-10,13	6,30	0,21	nein								
6	5	100	70	100	66	6%	25,00	-11,21	5,91	0,37	nein								
7	6	100	70	100	65	7%	20,00	-12,56	5,57	0,57	nein								
8	7	100	70	100	64	9%	16,67	-14,27	5,26	0,81	nein								
9	8	100	70	100	63	10%	14,29	-16,54	4,99	1,10	nein								
10	9	100	70	100	62	11%	12,50	-19,67	4,74	1,43	nein								
11	10	100	70	100	61	13%	11,11	-24,29	4,52	1,79	nein								
12	11	100	70	100	60	14%	10,00	-31,77	4,32	2,20	nein								
13	12	100	70	100	59	16%	9,09	-45,96	4,14	2,64	nein								
14	13	100	70	100	58	17%	8,33	-83,29	3,97	3,13	nein								
15	14	100	70	100	57	19%	7,69	-449,69	3,81	3,65	nein								
16	15	100	70	100	56	20%	7,14	131,80	3,67	4,20	ja								
17	16	100	70	100	55	21%	6,67	57,38	3,54	4,80	ja								

Notwendigkeit und „ausreichende“ Wirkungsstärke gemeinsam sind notwendig, um Ursache-Wirkungs-Zusammenhänge zu verstehen. In den weiteren Ausführungen werden Exogenität und Monotonie des Zusammenhanges definiert und entsprechende mathematische Beweise geführt. Die Begriffe werden in der weiteren Folge mit üblichen epidemiologischen Kennzahlen verglichen bzw. vergleichbar gemacht. In Kapitel 9.3.3 wird die Methodik auf ein epidemiologisches Fallbeispiel angewendet. In 9.3.4 wird der Nutzen der Kombination von experimentellen mit beobachtungsabhängigen Daten dargestellt. Die Zusammenstellung der Rahmenbedingungen und Möglichkeiten wird in Abbildung 2 dargestellt. Die Anwendung dieser Überlegungen auf die Frage eines ausreichenden Wirkungsnachweises zeigt, dass keineswegs nur experimentelle Studien – Randomized Controlled Trials (RCTs) – geeignet sind, als „Beweis“ verwendet zu werden. Viele Fragestellungen sind solchen Untersuchungen aus methodischen und ethischen Gründen nicht zugäng-

lich. Dies bedeutet aber als Konsequenz nicht, dass die „klinische Erfahrung des behandelnden Arztes“ nur im Sinne einer erzählten Erinnerung und nicht bestimmten Regeln der Dokumentation und der Nachvollziehbarkeit folgend als Entscheidungsbeurteilung bestimmend wird. Dies soll nun mit einer eigenen exemplarischen Berechnung, welche auf Interventionen grundsätzlich zutrifft und nicht nur auf Fragen der physikalischen Therapie anwendbar ist, erläutert werden.

In der Schmerztherapie ist ein Placeboeffekt von etwa 30 % bekannt.²² Die Berechnungen werden mit der „Vierfeldertafel“²³ für Interventionen durchgeführt. Dabei zeigt sich, dass bei Studien mit jeweils 100 Teilnehmern (Interventionsgruppe und Placebogruppe) ein relativer Unterschied in der Wirksamkeit von mindestens 20 % erforderlich ist, um eine statistisch signifikante Überlegenheit der Intervention nachzuweisen (Abbildung 3).

Eine weitere mathematisch begründete Tatsache ist damit gegeben, dass kleinere Unterschiede in der

Abbildung 4: Studiengröße und Wirksamkeitsunterschied

		Gruppe Placebo		Gruppe Intervention				CI											
20	Szenario	Personen	nicht gebessert	Personen	nicht gebessert	RRR	NNT	NNT min	NNT max	Chi Quadrat	Signifikant								
21	1	100	70	100	63	10%	14,29	-16,54	4,99	1,10	nein								
22	2	150	105	150	94,5	10%	14,29	-27,38	5,66	1,65	nein								
23	3	200	140	200	126	10%	14,29	-44,93	6,16	2,20	nein								
24	4	250	175	250	157,5	10%	14,29	-79,90	6,56	2,75	nein								
25	5	300	210	300	189	10%	14,29	-187,75	6,88	3,30	nein								
26	6	350	245	350	220,5	10%	14,29	3824,85	7,16	3,85	ja								
27	7	400	280	400	252	10%	14,29	209,84	7,39	4,40	ja								

22 Siehe z. B. Pincus, T., et al.: Patient Preference for Placebo, Acetaminophen (paracetamol) or Celecoxib Efficacy Studies (PACES): two randomised, double blind, placebo controlled, crossover clinical trials in patients with knee or hip osteoarthritis. Ann Rheum Dis 2004, 63: 931–939; Artus, M., et al.; Low back pain symptoms show a similar pattern of improvement following a wide range of primary care treatments: a systematic review of randomized clinical trials. Rheumatology 2010, 49: 2346–2356.

23 http://www.hauptverband.at/portal27/portal/hvbportal/channel_content/cmsWindow?action=2&p_menuid=72443&p_tabid=5

Wirksamkeit größere Studienpopulationen erforderlich machen. Daher ist vor einer klinischen Studie immer eine sogenannte Power Calculation durchzuführen, um abzuschätzen, ob mit der geplanten Studie überhaupt eine Chance besteht, den angenommenen Effekt nachzuweisen (Abbildung 4). Gleichzeitig hat eine vergleichende Untersuchung von Studien in der Indikation „Gelenksbeschwerden“ gezeigt, dass kleine Studien tendenziell die Wirksamkeit von Behandlungen überschätzen. Die Erklärungen liegen in Besonderheiten von Verzerrungsmöglichkeiten, welche bei kleinen Studienpopulationen schlechter ausgeschlossen werden können.²⁴ Werden immer kleinere Unterschiede mit größeren und teureren Studien als statistisch nachgewiesen wirksam belegt, entsteht eine neue Fragestellung, nämlich welches Ausmaß eines (nachweisbaren) Unterschiedes noch als bedeutsam oder „wertvoll“ beurteilt werden kann. Die Diskussion darüber soll durch folgende Darstellung unterstützt werden:

Bei einem relativem Wirkungsunterschied von 10 % wird die Schwelle eines statistischen Nachweises bei je 350 Patienten in den beiden Gruppen erreicht. Will man einen statistischen Wirkungsnachweis bei nur 1 % relativem Wirkungsunterschied nachweisen, so braucht man in jeder Gruppe 20.000 Patienten. Dabei wird allerdings ein Unterschied von 1 % in der Wirkung gegen Schmer-

zen als nicht relevant zu bezeichnen sein. Das Beispiel einer 20-prozentigen Wirksamkeit einer Therapie soll in der Folge auch mit den Werten von PN, PS und PNS diskutiert werden. Die entsprechenden Ergebnisse mit der Vierfeldertafel und der Erweiterung sind in Abbildung 5 enthalten.

Als Grundwerte zeigen 44 von 100 behandelten Personen einen Erfolg. Aber auch 30 von 100 nur scheinbar (Placebo) behandelten Personen zeigen einen Erfolg. Eine absolute Risikodifferenz (absoluter Unterschied in der Wirkung) von 14 % (CI 6,7 % bis 33,2 %) entspricht einer relativen Risikodifferenz (relativem Unterschied in der Wirkung) von 20 % (CI 0,76 % bis 27,24 %) und führt zu einer Number Needed to Treat (NNT) von 7,143 mit einem Konfidenzintervall, welches von 3,6 bis 131,7 reicht. Diese NNT bedeutet, dass im Schnitt 7,143 (möglich zwischen 3,6 und 131,7) Personen behandelt werden müssen, damit sich für eine Person ein Behandlungserfolg ergibt. Ein statistisch signifikantes Ergebnis einer wirksamen Therapie.

Wie sind nun die neuen Werte zu interpretieren? Die Intervention erklärt für 20 % der Personen tatsächlich den Erfolg. Damit ist PS gleich der Relative Risk Difference und 20 % derer, deren Gesundheitszustand sich verbessert hat, hätten daher ohne Intervention keine Verbesserung erfahren. Bei 31,8 % (das ist der Wert von PN) der erfolg-

Das „Bauchgefühl“ reicht für eine Beurteilung der Wirksamkeit nicht aus – man muss rechnen.

Abbildung 5: Werte bei 20 % Wirksamkeit einer Intervention

Ergebnis	Intervention	X ¹ -Placebo		
	X-Inrevention			
Y-Therapieversagen	56	70		
Y ¹ -Therapieerfolg	44	30		
PNS	$(Y_i / (Y_i + Y_{i1})) - (Y_p / (Y_p + Y_{p1}))$	0,140		
PN	$PNS / (Y_i / (Y_i + Y_{i1}))$	0,318		
PS	$PNS / (1 - (Y_i / (Y_i + Y_{i1})))$	0,200		
		Ergebnisse	CI	
Control event rate	$c / (c + d)$	0,700		
Experimental event rate	$a / (a + b)$	0,560		
Relative risk difference	$(CER - EER) / CER$	0,200	0,0676	0,3324
Risk difference	$CER - EER$	0,140	0,0076	0,2724
Number needed to treat	$1 / (CER - EER)$	7,143	131,7952	3,6709
Risk ratio	$(a / (a + b)) / (c / (c + d))$	0,800		
Odds ratio	$(a / b) / (c / d) = ad / bc$	0,545		
SE (RD, ARR)		0,068		
Chi Quadrat		4,204		
		Signifikant:	ja	

Ausgangswerte: grün
 Ergebnisse nach Pearl: blau
 Ergebnisse der Vierfeldertafel: hellgrün

²⁴ Nüesch, E., et al.: Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. BMJ 2010, 341: c3515.
 Nüesch, E., et al.: The Importance of Allocation Concealment and Patient Blinding in Osteoarthritis Trials: A Meta-Epidemiologic Study. Arthritis & Rheumatism (Arthritis Care & Research), Vol. 61, No. 12, December 15, 2009: 1633–1641.

Abbildung 6: Wirksamkeit von Kombinationen

Zusammen II „kombiniert“				Medikament		
Ergebnis	Intervention			Ergebnis	Intervention	
	X-Intervention	X ¹ -Placebo			X-Intervention	X ¹ -Placebo
Y-Therapieversagen	26	70		Y-Therapieversagen	40	70
Y ¹ -Therapieerfolg	74	30		Y ¹ -Therapieerfolg	60	30
PNS	0,440			PNS	0,300	
PN	0,595			PN	0,500	
PS	0,629			PS	0,429	
	Ergebnisse	CI		Ergebnisse	CI	
Control event rate	0,700			0,700		
Experimental event rate	0,260			0,400		
Relative risk difference	0,629	0,5042	0,7529	0,429	0,2971	0,5601
Risk difference	0,440	0,3157	0,5643	0,300	0,1685	0,4315
Number needed to treat	2,273	3,1679	1,7720	3,333	5,9340	2,3176
Risk ratio	0,371			0,571		
Odds ratio	0,151			0,286		
SE (RD, ARR)	0,063			0,067		
Chi Quadrat	38,782			18,182		
	Signifikant:	ja		Signifikant:	ja	

Erfahrung ohne rechnerische Untersuchung der Tatsachen führt leicht in die Irre.

reich Behandelten war die Behandlung tatsächlich für den Erfolg erforderlich. Aus PNS ergibt sich, dass für 14 % die Behandlung tatsächlich einen Unterschied bedeutet. Die restlichen 86 % haben unabhängig von der Entscheidung zwischen Intervention und Placebo den gleichen Krankheitsverlauf.

5.4 Erfahrung aus dem klinischen Alltag

Als Behandler erlebt man bei fast jedem zweiten Patienten (44 von 100) einen Erfolg. Kombiniert man die Behandlung mit einer zweiten Intervention (beispielsweise einem Medikament), bei welcher 60 von 100 einen Erfolg erleben, so ergibt die gemeinsame Wirksamkeit ein eindruckliches Ergebnis. Damit wird die Erfahrung aus dem klinischen Alltag nachvollziehbar. Bei einer Behandlung nur mit einem Medikament ist nur jede dritte Behandlung erfolgreich. Bei einer anderen Behandlung ist nur jede siebente erfolgreich. Gemeinsam ist (rechnerisch) jeder zweite Patient ein Erfolg. Bei diesen hohen Wirkunterschieden ist auch PNS mit 44 % eine starke Begründung für die Notwendigkeit einer solchen Therapiekombination, bedeutet dies doch, dass für 44 (60 % der Erfolge) von 100 Personen die Durchführung der Kombinationstherapie einen Unterschied im Krankheitsverlauf bewirkt hat. Spricht man solchen Rechenergebnissen allerdings nicht die Fähigkeit einer ausreichenden Begründung für medizinisches Handeln zu, so müsste man jede einzelne Kombination eines Wirkstoffes mit einer anderen Intervention – z. B. einer physi-

kalischen Therapie – in einer eigenen randomisierten Studie überprüfen.

5.5 Interpretation aus „Systemsicht“

Mit Beispielen für „evidenzbasierte“ Behandlungen wurde versucht, die mathematischen Zusammenhänge und ihre Auswirkungen auf das Erleben von „Wirkung“ im Alltag verständlich zu machen. Aber auch bei derart gut nachgewiesenen Therapien erscheint die „Systemsicht“ keineswegs als vorteilhaft. Eine NNT von sieben bedeutet schließlich, dass sechs Patienten nicht nur keinen Erfolg erleben, sondern auch die vollen Behandlungskosten anfallen. Beispiele von wissenschaftlich belegten und in Leitlinien empfohlenen Interventionen mit Werten der NNT von 30 oder 50 sind keine Seltenheit. Für eine Empfehlung einer Primärprävention bei erhöhten Blutfettwerten beträgt die NNT > 150.²⁴ Diese unterschiedliche Sichtweise führt zu einer Bewertung, die von der durch klinisches Erleben bedingten Erfahrung abweicht.

Da diese – aus klinischer Sicht „falsche“ – Bewertung mittels Statistik und Studien begründet wird, folgt nur allzu leicht eine kritische oder sogar ablehnende Einstellung zur Methode selbst. Wird die Methode des wissenschaftlichen Erkenntnisgewinnes abgelehnt – weil sie zu Ergebnissen führt, welche nicht zur Alltagsrealität passen –, wird ein Nachweis oder eigentlich eine kritische Untersuchung von Interventionen nicht als wesentlich angesehen werden.

24 Shepherd, J.: Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. NEJM 1995, 333 (20).

6 Zusammenfassung

Zwei aktuelle Anlässe einer Kritik an wissenschaftlichen Ergebnissen der Arbeit des seit nunmehr zehn Jahren (in dieser Form insgesamt seit zwölf Jahren) bestehenden Teams „Evidenzbasierte Medizin und Health Technology Assessment“ im Hauptverband der österreichischen Sozialversicherungsträger sind die Motivation und das Spannungsfeld zwischen „persönlicher klinischer Erfahrung“ und „wissenschaftlich abgesicherten Erkenntnissen“ zu untersuchen. Neben verschiedenen methodischen Überlegungen sind es besonders zwei Aspekte, welche diese Spannungen bedingen:

- Betrachtung des individuellen Falles und Sammlung solcher Fälle zu „Erfahrung“ versus Betrachtung statistischer Effekte mit Geltung für (größere) Gruppen. Damit wird die Frage nach „aktuellen Ursachen im Einzelfall“ gegenüber „generellen Ursachen für Falltypen“ und deren Erklärung eingefordert.
- Wert von rechnerischen (statistischen) Ergebnissen von (nicht dem Alltag entsprechenden) Experimenten gegenüber der durch erinnerte Einzelfälle entstehenden „klinischen Erfahrung“ von Experten. Dies vor allem dann, wenn die rechnerischen Ergebnisse mit der Alltagserfahrung nicht übereinstimmen. Dies erfordert neben einer Erklärung vor allem eine Methodenkritik aus erkenntnistheoretischer Sicht.

Die Anlassfälle zeigen aber auch eine unterschiedliche Umgangsweise mit der Tatsache einer unterschiedlichen Sichtweise. Einmal wird ein direktes Geschäftsinteresse gefährdet gesehen und diese Gefahr bei Gericht mit Rechtsmitteln bekämpft. Das Ergebnis wird dabei wohl durch ein Gericht festgelegt werden. Im anderen Fall wird eine Form eines wissenschaftlichen/gesellschaftlichen Diskurses gewählt. Dieser Diskurs zeigt Bereiche, in welchen durchaus Übereinstimmung besteht, und andere, welche unterschiedlich gesehen werden. Übereinstimmung besteht dahingehend, dass alle Bestrebungen nicht zu absoluter Sicherheit führen werden. Dazu sei nochmals aus Aldersons Artikel (siehe Fußnote 18) zitiert: *„When is it reasonable to claim that a study has proved that no effect or no difference exists? The correct answer is ‚never‘, because some uncertainty will always exist. However, we need to have some rules for deciding when we are fairly sure that we have excluded an important benefit or harm.“*

Ausgehend davon kann man untersuchen, welche Bedingungen für Entscheidungen unter Unsicherheitsbedingungen als ausreichend bzw. notwendig angesehen werden. Dies führt zu dem Versuch einer Erklärung des Zustandekommens unterschiedlicher Sichtweisen. Verständlich werden diese Un-



© Vladimir Voronin - Fotolia.com

terschiede dann durch differente Aufgabenstellungen und Notwendigkeiten. Wenn eine Therapiekombination in drei Viertel der Fälle erfolgreich ist, erscheint es aus klinischer Sicht (und wohl auch aus jener der betroffenen Patienten) als irrelevant, dass der Erfolg „nur“ zu 60 % durch das Handeln erklärt wird. Aus der Sicht des Gesamtsystems ist der Wert der Erfolgsrate von 74 % dadurch gemindert, dass für 56 % aller Betroffenen die Entscheidung für die Therapie den Verlauf nicht verändert, aber natürlich die Kosten des Systems deutlich erhöht, da die Behandlungen ja bezahlt werden müssen. Der rechnerische Vorteil von Kombinationen von Interventionen lässt sich leicht belegen. Ob dieser in der Realität erreichbar ist oder die Erwartungen sogar übersteigt, kann nach geltender Lehrmeinung nur durch experimentelle Untersuchungen nachgewiesen werden. Hier ist eine Weiterentwicklung der Methodik unbedingt erforderlich, da die Erfüllung dieser Forderung wegen des damit verbundenen Aufwandes und ethischer Beschränkungen für zulässige klinische Studiendesigns nicht zu erwarten ist.

Diese Form der Unsicherheit und Teilwirksamkeit führt zur Diskussion, ob das Glas halb voll oder halb leer ist. Der Zusammenhang zwischen Wirkungsunterschied und der Möglichkeit bzw. dem Aufwand für einen wissenschaftlichen (eventuell experimentellen) Nachweis ist ebenfalls mathematisch leicht darstellbar. Schwierigkeiten, diesen Nachweis zu erbringen, können aber nicht dazu führen, dass die Notwendigkeit eines positiven Nachweises grundsätzlich verneint wird.

Der kritische Diskurs ist jedenfalls ein Ansporn zu versuchen, eine komplexe Materie mittels einfacher und nachvollziehbarer Beispiele verständlich aufzubereiten. Alleine deshalb sollte dieser Weg gemeinsam weiter beschritten werden.