

fundierte Orientierung an einer Verbesserung der Gesundheit ist nicht erkennbar. Es ist leider nicht nachvollziehbar, welche Forschungsfragen im Vorfeld vom Gesundheitsforschungsrat erörtert wurden (und welche nicht) und wie die expliziten Kriterien lauten, nach denen Themen ein- und ausgeschlossen und schließlich in eine Rangfolge gebracht wurden.

Glasgow et al. (2007, S. 354) kritisieren die Gesundheitsforschungsagenda der amerikanischen National Institutes of Health als einseitig technologisch ausgerichtet, auf die pharmakogenetische Behandlung von individuellen Patienten mit genetischen Risiken zielend, das sog »4 p's«-Konzept: predictive, personalized, preemptive, participatory (prädiktiv, personalisiert, präventiv, partizipativ). Für einen gesamtgesellschaftlichen Nutzen müssten auch die Determinanten der Bevölkerungsgesundheit und des Gesundheitsverhaltens berücksichtigt werden. Als Prüfsteine dafür, dass Erkenntnisse gewonnen werden, die zum Nutzen der Gesellschaft umgesetzt werden, empfehlen sie 4 anders geartete Fragen (»4 W's«): »Who pays (and how much)? Who benefits? Who suffers? Who profits?«.

Reformen in der Zielsetzung, Planung und Umsetzung der Gesundheitsforschung sind erforderlich, wenn diese einen größeren Beitrag zur Verbesserung der Gesundheit und der gesundheitlichen Versorgung leisten soll. Notwendig erscheinen insbesondere die Erhöhung der Transparenz sowie Entscheidungen, die weniger von den Eigeninteressen wissenschaftlicher Gemeinschaften und der Wirtschaft geprägt sind. Im Zentrum des Interesses sollten der Patientennutzen medizinischer Leistungen und die Effektivität von Versorgungssystemen stehen. Kriterien sollten entwickelt werden für die Entscheidungen darüber, welche Fragen unter dem Gesichtspunkt der Begrenztheit der Ressourcen in der Forschung und in der Versorgung Vorrang haben.

## Weiterführende Literatur

Nelson DE (2011). Health in Research Agendas and Funding. In: Gigerenzer G, Muir Gray JA. Better Doctors, Better Patients, Better Decisions: Envisioning Health Care 2020. Cambridge, MA: MIT Press *Guter Überblick über das Thema Forschungsförderung.*

## 3.8 Evidenzbasierte berufliche Praxis

### 3.8.1. Einführung

#### Merksatz

Evidenzbasierte berufliche Praxis bezeichnet ein Handlungsprinzip in den Gesundheits- und Sozialberufen mit dem Ziel, wissenschaftliches Wissen (Evidenz) und Präferenzen der Patienten bzw. Klienten zur Grundlage von Handlungsentscheidungen zu machen.

Alle Berufsgruppen im Gesundheits- und Sozialwesen sind ständig mit der Frage befasst, welche von mehreren möglichen Interventionen das gesundheitliche, psychische oder soziale Problem (oder eine Kombination daraus) eines Patienten bzw. eines Klienten am besten löst. Das Konzept der Evidenzbasierten beruflichen Praxis (EBP) zielt darauf ab, diese Entscheidungen für individuelle, einmalige Patienten bzw. Klienten auf die gegenwärtig beste externe, Evidenz zu gründen.

#### Hintergrund des Konzeptes einer Evidenzbasierten beruflichen Praxis

Die Frage nach wissenschaftlichem Wissen (Evidenz) als Grundlage für problemlösendes berufliches Handeln stellt sich allen Berufsgruppen im Gesundheits- und Sozialwesen. Für die Medizin wurde das Konzept Evidenzbasierte Medizin (EBM) in den 1980er und 1990er Jahren von klinisch tätigen Ärzten, Gesundheitswissenschaftlern und Epidemiologen entwickelt. Diese hatten festgestellt, dass Patienten nicht die bestmögliche Behandlung erhielten, weil Ärzte

ihre Behandlungsentscheidungen auf überholtes Wissen, ungeprüfte Therapien und falsche Theorien gründeten anstatt auf die beste verfügbare Evidenz. Dieser Gedanke ist weder neu noch spezifisch für die Medizin. Für die Soziale Arbeit legte Mary Ellen Richmond bereits 1917 (Richmond 1917 S. 38 ff.) detailliert und weitsichtig verschiedene Formen und Anwendungsgebiete von Evidenz dar. Auch erkannte sie die Bedeutung des fairen Vergleichs (Richmond 1917, S. 40) zu einem Zeitpunkt, als der Begriff randomisierte kontrollierte Studie noch nicht geprägt war.

Der Begriff Evidence Based Medicine wurde 1990 von Gordon Guyatt geprägt (Guyatt et al. 2015, Kindle Position 497) und schnell von Wissenschaftlern und Praktikern anderer Disziplinen aufgegriffen: Evidence-based Nursing and Caring (<http://ebn-zentrum.de>), Evidence-based Social Work (Meng 2006), Evidence-based Practice in Psychology (APA 2005), Evidence-based Physiotherapy (CEBT Maastricht <http://www.cebp.nl>).

Bildungspolitisch wird in Deutschland die Bedeutung der Evidenzbasierung aller Berufe in der Gesundheitsversorgung anerkannt und damit einhergehend die (partielle) Akademisierung gefordert.

Die EBP bedeutet eine deutliche Veränderung der bis dahin gültigen Denk- und Vorgehensweisen. Guyatt hat die neue und die alte Sicht im Vergleich von zusammengefasst (Daly 2005, S. 88 f., eigene Übersetzung):

Die Annahmen der Evidenzbasierten Medizin

1. Klinische Erfahrung hat einen niedrigen Stellenwert für die Bewertung diagnostischer Tests, für die Frage, ob Therapien funktionieren und für die Prognose.
2. Medizinische Ausbildung und gesunder Menschenverstand sind eine sehr unzulängliche Richtschnur für die Beurteilung der Validität von Studienergebnissen. Erforderlich sind Evidenzregeln, insbeson-

dere aus der klinischen Epidemiologie.

3. Annahmen über die Wirksamkeit auf Grundlage der Physiologie sind ohne empirische Überprüfung häufig irreführend.
4. Die Meinung von Experten hat einen niedrigen Wert.

Die alten Annahmen:

1. Die klinische Erfahrung ist ein valider Weg, um Wissen über die Prognose, den Nutzen diagnostischer Tests und Untersuchungen und Therapie zu gewinnen.
2. Die richtige Behandlung lässt sich allein auf Grundlage von Physiologie und physiologischer Mechanismen ermitteln. Das Wissen darum, wie ein Medikament die Physiologie beeinflusst, lässt den klinischen Effekt vorhersagen.
3. Eine gute medizinische Ausbildung und gesunder Menschenverstand reichen für die kritische Bewertung von Studien aus.
4. Die Meinung von Experten hat einen hohen Wert.

Hierzu ist anzumerken, dass dem niedrigen Wert der Expertenmeinung ein überholtes Verständnis von »Experte« zugrunde liegt, nämlich solche Ärzte, die ihre auf Erfahrung gründende Meinung auch ohne Berücksichtigung der Studienlage für gültig erachten. Ironisch werden sie auch als »Eminenzen« bezeichnet. Ein wirklicher Experte integriert Studienwissen und reflektiert seine Erfahrung.

### Wie der Begriff Evidence based Medicine entstand

Der neuen Vorgehensweise wollte Guyatt im Jahr 1990 für seine Lehre an der Medizinischen Fakultät den Namen »scientific medicine« geben. Seine Kollegen in der medizinischen Fakultät liefen dagegen Sturm, weil »scientific medicine« suggeriere, sie würden eine unwissenschaftliche Medizin betreiben. Sein zweiter Versuch lautete »evidence based medicine«.

Diesen Begriff akzeptierte die Fakultät (Guyatt et al. 2015, Kindle Position 493). Evidenzbasierte Medizin ist also nicht mehr und nicht weniger als ein im Jahr 1990 noch neues Konzept von Wissenschaftlichkeit in der Medizin.

*»Through Dave Sackett's leadership, critical appraisal had evolved into a philosophy of medical practice based on knowledge and understanding of the medical literature (or lack of such knowledge and understanding) supporting each clinical decision. (...) The term suggested to describe this new approach was scientific medicine. Those already hostile were incensed and disturbed at the implication that they had previously been »unscientific«. My second try at a name for our philosophy of medical practice, evidence-based medicine, turned out to be a catchy one. (...) Residents are taught an attitude of 'enlightened scepticism' towards the application of diagnostic, therapeutic, and prognostic technologies in their day-to-day management of patients.«* (Guyatt et al. 2015, Kindle Position 498)

*Durch David Sacketts Initiative hatte sich die kritische Bewertung in eine Philosophie der medizinischen Praxis entwickelt, die sich auf Kenntnis und Verständnis der medizinischen Literatur gründete (bzw. den Mangel eines solchen Wissens und Verständnisses) zur Unterstützung jeder Entscheidung. Der Begriff, den ich vorschlug, um diese neue Herangehensweise in der Medizin zu beschreiben, lautete »wissenschaftliche Medizin«. Die mit einer bereits ablehnenden Haltung waren erobert und beunruhigt über die Implikation, dass sie bisher »unwissenschaftlich« waren. Mein zweiter Versuch der Namensfindung für unsere Philosophie der medizinischen Praxis, Evidenzbasierte Medizin, erwies sich als eingängig und prägnant. (...) Ärzten in der Aus- und Weiterbildung wird die Haltung eines »aufgeklärten Skeptizismus« gegenüber der Anwendung diagnostischer, therapeutischer und prognostischer Technologien gelehrt, die sie in ihrem klinischen Arbeitsalltag anwenden.*

### **Der Wissenschaftsrat zur evidenzbasierten Praxis in den Gesundheitsberufen**

In seinen »Empfehlungen zu hochschulischen Qualifikationen für das Gesundheitswesen« aus dem Jahr 2012 bezeichnete es der Wissenschaftsrat für zunehmend wichtig, dass *die mit besonders komplexen Aufgaben betrauten Angehörigen der Gesundheitsfachberufe ihr eigenes pflegerisches, therapeutisches oder geburtsbelferisches Handeln auf der Basis wissenschaftlicher Erkenntnis reflektieren, die zur Verfügung stehenden Versorgungsmöglichkeiten hinsichtlich ihrer Evidenzbasierung kritisch prüfen und das eigene Handeln entsprechend anpassen können«* (Wissenschaftsrat 2012, S. 8).

Die folgenden Darlegungen am Beispiel der EBM sollen helfen, der notwendigerweise zu führenden Diskussion um eine EBP eine sachliche und gut informierte Grundlage zu geben. Bislang leiden manche Auseinandersetzungen unter Missverständnissen und Fehlinterpretationen. Wie immer man zur EBP steht, ein Wissen um das Konzept, die Begrifflichkeiten sowie die Möglichkeiten und Grenzen erscheinen heutzutage unabdingbar.

### **Definition Evidenzbasierte Medizin**

*»EBM ist der gewissenhafte, ausdrückliche und vernünftige Gebrauch der gegenwärtig besten externen wissenschaftlichen Evidenz für Entscheidungen in der medizinischen Versorgung individueller Patienten. Die Praxis der EbM bedeutet die Integration individueller klinischer Expertise mit der bestverfügbaren externen Evidenz aus systematischer Forschung.«* (Website Deutsches Netzwerk Evidenzbasierte Medizin <http://tinyurl.com/q4szwyk>)

Diese Definition orientiert sich an der häufig zitierten Definition einer kanadisch-englischen Arbeitsgruppe aus dem Jahr 1996:

*»Evidence based medicine is the conscientious and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence based medicine means in-*

*tegrating individual clinical expertise with the best available external clinical evidence from systematic research.*« (Sackett et al. 1996)

Der Fokus der EBM richtet sich auf Entscheidungen, die zwischen Arzt und Patient zu treffen sind. Dabei geht es um das Zusammenspiel ärztlicher Erfahrung auf dem Gebiet der Krankenbehandlung (»individuelle klinische Expertise«) mit dem aktuellen Stand der Wissenschaft (»beste externe wissenschaftliche Evidenz aus systematischer Forschung«), um zu einer Entscheidung zu gelangen, die den individuellen und persönlichen Bedürfnissen des Patienten entspricht (S. 145 ff.).

»Individuelle klinische Expertise« bezeichnet die Erfahrung der Ärzte im medizinischen Alltag. Erfahrung im Sinne wiederholter Beobachtung vergleichbarer Fälle dient zur Bildung von Mustern und Heuristiken (S. 112). Wenige und teils unbewusst wahrgenommene Merkmale ermöglichen, erleichtern und beschleunigen das Erfassen, Einordnen und Verstehen von Symptomen, die Formulierung zielführender Hypothesen und die Anwendung von Studienwissen zum Nutzen des Patienten. Ärzte und Angehörige anderer Gesundheitsberufe können jedoch auf Grundlage ihrer Erfahrung und Expertise kein sichereres Urteil über die Wirksamkeit einer Behandlung treffen – die Beobachtung von Einzelfällen führt, wie bereits dargestellt, fast regelhaft in die Irre. Die klinische Expertise kann also das Studienwissen, die externe Evidenz, nicht ersetzen (S. 171 ff.). »Beste externe wissenschaftliche Evidenz aus systematischer Forschung« – systematische Forschung ist gleichzusetzen mit wissenschaftlichen Untersuchungen (die ja stets einem Plan, einer Systematik folgen). Das Wort »extern« in Zusammenhang mit dem Wort »Beweis« bezieht sich auf Wissen, das durch Studien gewonnen wird. Das Wort »bestmöglicher« in Zusammenhang mit »Beweis« hebt hervor, dass vorhandene Evidenz kritisch auf ihre Validität bewertet wird. Während heute für viele Fragen

die beste Evidenz in aufbereiteter Form zu finden ist, musste sich der Arzt zum Zeitpunkt der Veröffentlichung dieser Definition die Evidenz zur Lösung eines Patientenproblems noch mit Hilfe der 5-schrittigen Methode (S. 131 ff.) beschaffen. »Beste wissenschaftliche Evidenz« sind die Ergebnisse aus klinisch relevanter, patientenzentrierter Forschung zur Wirksamkeit von therapeutischen, rehabilitativen und präventiven Maßnahmen sowie zur Genauigkeit von diagnostischen Untersuchungen und zur Vorhersagekraft prognostischer Marker.

Die Integration von individueller Expertise und der bestverfügbaren externen Evidenz soll verdeutlichen, dass die Evidenz alleine als Grundlage für Entscheidungen nicht ausreicht. Weitere Komponenten sind das Wissen und die Erfahrung des Arztes aber auch des Patienten. Evidenz ist eine notwendige, aber nicht hinreichende Voraussetzung für bessere Entscheidungen. Patienten entwickeln entsprechend ihrer Lebenssituation und ihrer subjektiven Bewertung der Evidenz und der Erwünschtheit von Behandlungszielen sowie der Nutzenwahrscheinlichkeiten und Schadensrisiken durchaus unterschiedliche Präferenzen. Die Individualisierung bzw. Personalisierung von Entscheidung ist ein Kernelement wohlverstandener Evidenzbasierter Medizin. Die dazu geeignete Form der Arzt-Patient-Kommunikation ist Shared Decision Making (S. 145 ff.).

Die Begründer der Evidenzbasierten Medizin selbst weisen unmissverständlich darauf hin, dass die klinische Praxis von der Evidenz tyrannisiert zu werden droht: *»Good doctors use both individual clinical expertise and the best available external evidence, and neither alone is enough. Without clinical expertise, practice risks becoming tyrannised by evidence for even excellent external evidence may be inapplicable to or inappropriate for an individual patient.*« (Sackett et al. 1996) *Gute Ärzte nutzen sowohl die klinische Expertise als auch die beste verfügbare Evidenz – keines von beiden ist*

*alleine genug. Ohne klinische Expertise droht die Praxis von der Evidenz tyrannisiert zu werden, weil selbst die beste externe Evidenz für den einzelnen Patienten nicht anwendbar oder unpassend sein kann.*

Evidenz ist die Grundlage für einen Prozess der Entscheidungsfindung, zu dessen Ingredienzen notwendigerweise auch die klinische Expertise (Erfahrung) des Arztes und die Klärung der Wünsche und Bedürfnisse (Präferenzen) des Patienten gehören. Diese Integration erfolgt im Shared Decision Making. Gelingt dieser Prozess, handelt es sich um eine individualisierte und passgenaue Entscheidung für einen einzigartigen Patienten.

Was unter »Evidenz« zu verstehen ist, gibt die folgende Definition einer amerikanischen Arbeitsgruppe wieder:

### Definition Evidenz

*»Evidence is generally considered to be information from clinical experience that has met some established test of validity, and the appropriate standard is determined according to the requirements of the intervention and clinical circumstance.« (IOM Roundtable on Evidence-Based Medicine 2007, S. 354) Evidenz wird allgemein als Information aus klinischer Erfahrung betrachtet, die durch einen anerkannten Test auf ihre Validität geprüft wurde. Der erforderliche Standard wird durch die Anforderungen der Intervention und durch die klinische Frage bestimmt.*

Informationen, die auf klinischer Erfahrung beruhen (»information from clinical experience«) – sowohl systematisch in Studien als auch individuell im Alltag gewonnene Erfahrungen – werden dieser Definition zufolge dadurch zu Evidenz, dass sie mit einer geeigneten Methode auf ihre Validität untersucht werden. Welche Methode als geeignet erachtet wird, hängt von der Fragestellung ab und von der für akzeptabel erachteten Irrtumswahrscheinlichkeit. Evidenz dient dazu, der Erörterung und Abwägung vor

einer Entscheidung eine rationale Grundlage zu geben, sie determiniert jedoch die Entscheidung nicht. Dazu Gordon Guyatt, der »Erfinder« des Begriffs Evidenzbasierte Medizin:

*»The »adequacy of evidence« is a matter of judgment or opinion, not an objective characteristic of the evidence itself.« (Guyatt et al. 2000) Welche Evidenz angemessen ist, ist kein objektives Merkmal der Evidenz, sondern eine Sache der Beurteilung oder Meinung.*

Am Beispiel von zwei Methoden zur Prävention der tiefen Beinvenenthrombosen soll verdeutlicht werden, dass unterschiedliche Ansprüche an die Evidenz gerechtfertigt sein können.

### Prävention der tiefen Beinvenenthrombose bei Langzeitflügen

Langzeitflüge erhöhen das Risiko für Beinvenenthrombosen. Die Studienlage für präventive Maßnahmen ist dürftig.

Besonders Gefährdeten empfiehlt das Arznteilegramm wiederholtes Anspannen der Wadenmuskulatur, wiederholtes Herumwandeln im Gang und bei Flügen ab sechs Stunden Dauer Unterschenkelkompressionsstrümpfe. In Studien wurden diese Maßnahmen bislang nicht geprüft. Es handelt sich um eine Expertenmeinung, die allein auf pathophysiologischen Erwägungen und Plausibilität gründet (arznteilegramm 2012). Die Evidenzstufe ist niedrig und die Irrtumswahrscheinlichkeit hoch. Da aber Schadensmöglichkeiten kaum denkbar sind, keine höherwertige Evidenz vorliegt und noch dazu keine oder nur geringe Kosten entstehen, ist es in diesem Fall hinnehmbar, dass das Thromboserisiko möglicherweise nicht gesenkt wird. Die Nutzen-Schaden-Bilanz kann somit nicht negativ werden.

Anders stellt sich die Situation für die Prävention durch blutgerinnungshemmende Mittel dar. Hier sind Schadensmöglichkeiten in Studien nachgewiesen – das Risiko für Blutungen wird erhöht. Bei ungewissem Nutzen und

nachgewiesenen Schadensrisiken ist es durchaus denkbar, dass der mögliche Schaden den möglichen Nutzen überwiegt. Hier steht also mehr auf dem Spiel, entsprechend sind die Ansprüche an die Güte der Evidenz höher (s.a. Evidenzstufen, S. 135).

Aus dem Beispiel folgt, dass eine Behandlung ohne Schadensrisiken, die möglicherweise das Risiko für eine Lungenembolie mindert, auch ohne höhergradige Evidenz für die meisten Menschen akzeptabel sein dürfte. Geht die Behandlung jedoch mit gravierenden Risiken einher, wie z.B. einer Gehirnblutung, sind die Ansprüche an den Nutznachweis naturgemäß höher. Die Umstände und die persönliche Bewertung der Risiken bestimmen somit, welche Evidenz für angemessen erachtet wird.

Einen andersartigen und bedeutsamen Kontext für Evidenz stellen Versorgungsentscheidungen auf Systemebene dar, also Entscheidungen darüber, welche Leistungen z.B. die GKV ihren Versicherten finanziert und welche nicht. Die Evidenzbasierte Medizin kann hier zu rationalen Entscheidungen und einer evidenzbasierten Gesundheitsversorgung (evidence based healthcare) beitragen. In Deutschland soll der Gemeinsame Bundesausschuss dieses Anliegen verwirklichen.

### Nicht evidenzbasiert?

Im medizinischen Alltag werden Behandlungsverfahren bisweilen als »evidenzbasiert« bzw. »nicht evidenzbasiert« bezeichnet. Dieser Sprachregelung liegt die falsche Vorstellung zugrunde, Evidenz sei entweder vorhanden oder nicht. Richtig ist hingegen, dass Evidenz immer vorliegt, jedoch die Frage nach der Qualität und Verlässlichkeit gestellt werden muss. Liegt keine Evidenz höherer Stufe vor, kann die Expertenmeinung durchaus die beste verfügbare Evidenz und somit trotz ihrer eher niedrigen Validität und hohen Irrtumswahrscheinlichkeit eine

rationale Grundlage für eine Behandlungsentscheidung sein. Die Protagonisten der EBM haben ausdrücklich festgestellt, dass jede empirische Beobachtung einen potentiellen Beweis darstellt. »*What is the nature of the evidence in EBM? We suggest a broad definition: any empirical observation about the apparent relationship between events constitutes potential evidence. (...) The hierarchy makes it clear that any statement to the effect that there is no evidence addressing the effect of a particular treatment is a non sequitur.*« (Guyatt et al. 2000) *Was ist die Beschaffenheit der Evidenz in EBM? Ich schlage eine weite Definition vor: Jede empirische Beobachtung über die offensichtliche Beziehung zwischen zwei Ereignissen stellt potenzielle Evidenz dar. (...) Die Evidenzhierarchie verdeutlicht, dass jede Feststellung derart, dass es keine Evidenz für den Effekt einer bestimmten Behandlung gibt, ein logischer Fehler ist. [lat. non sequitur: logischer Fehler innerhalb der Argumentation eines Beweises].*

Mit anderen Worten: Keine Evidenz gibt es nicht.

Hier noch ein Hinweis auf eine Merkwürdigkeit. Das Wort Evidenz bedeutet im Zusammenhang mit evidenzbasierter Medizin »Beweis, Nachweis, Beleg«, entsprechend dem englischen Wort evidence (»facts or signs that show clearly that something exists or is true« (Longman Dictionary of Contemporary English online <http://tinyurl.com/pk9f4m2>). Evidence« entspricht dem Beweis, wie er vor Gericht erbracht wird, um beispielsweise die Richter von der Schuld oder Unschuld eines Angeklagten zu überzeugen. Das deutsche Adjektiv »evident« hingegen bedeutet laut Duden »unmittelbar einleuchtend, keines Beweises bedürftig« – evident und Evidenz meinen also etwas eher Gegensätzliches.

### 3.8.2. Evidenzbasierte Praxis – das Handlungskonzept

Ein Problem erkennen, präzise beschreiben und auf Grundlage von Evidenz lösen – dies ist in wenigen Worten die Vorgehensweise einer EBP. Das Vorgehen lässt sich in 5 Schritte fassen: (Deutsches Netzwerk Evidenzbasierte Medizin <http://tinyurl.com/q4szwyk>)

1. Ableitung einer relevanten, beantwortbaren Frage aus dem klinischen Fall
2. Planung und Durchführung einer Recherche der klinischen Literatur
3. Kritische Bewertung der recherchierten Literatur (Evidenz) bezüglich Validität/ Brauchbarkeit
4. Anwendung der ausgewählten und bewerteten Evidenz beim individuellen Fall
5. Bewertung der eigenen Leistung.

Hier die etwas ausführlichere englische Version:  
 Step 1 Converting the need for information (about prevention, diagnosis, prognosis, therapy, causation, etc) into an answerable question  
 Step 2 Tracking down the best evidence with which to answer that question  
 Step 3 Critically appraising that evidence for its validity (closeness to the truth), impact (size of the effect), and applicability (usefulness in our clinical practice)

Step 4 Integrating the critical appraisal with our clinical expertise and with our patient's unique biology, values and circumstances

Step 5 Evaluating our effectiveness and efficiency in executing Steps 1 – 4 and seeking ways to improve them both for next time (Straus et al. 2010, S. 3).

Jeder dieser Schritte erfordert Fertigkeiten, die systematisch zu erlernen sind. Im Folgenden sollen die Schritte anhand von vier Beispielen skizziert verdeutlicht werden, in denen es jeweils um den Vergleich von zwei Interventionen bzw. Vorgehensweisen geht. Mit jeder dieser Fragen sind jeweils verschiedene Gesundheitsberufe befasst. Die Entwicklung

einer Informationsinfrastruktur ist bislang in der Medizin und in der Pflege am weitesten fortgeschritten.

#### Schritt 1 Vom Problem zur recherchierbaren und beantwortbaren Frage

1. Führt bei Patienten mit akutem unkompliziertem Kreuzschmerz Aktivierung im Vergleich zu Bettruhe schneller zur Schmerzfremheit?
2. Lässt sich durch ein strukturiertes Schulungsprogramm (structured intervention, multicomponent intervention) die Rate zwangsfixierter Bewohner (physical restraints) in Pflegeheimen (nursing homes) mindern?

Die Forschungsfrage sollte als PICO-Frage formuliert werden:

#### Die PICO-Frage

- P Population/Patientenproblem: Beschreibung des Patienten/des Problems  
 I Intervention: die Intervention (wobei der Begriff sehr weit gefasst auch eine Ursache, einen prognostischen Faktor oder eine Behandlung bezeichnen kann)  
 C Comparison: die Vergleichsmaßnahme  
 O Outcome: die Endpunkte, an denen man den Erfolg messen möchte (Guyatt et al. 2015, Kindle Position 1063).

Der Patient und sein Problem sowie die Intervention sollten präzise beschrieben werden.

Als Vergleichsintervention kommt je nach Fragestellung der bisherige Standard in Frage, eine Scheinintervention bzw. ein Placebo oder auch »keine Intervention«, wie z.B. eine Wartegruppe.

Als Outcome sollten patientenrelevante Endpunkte definiert werden, also solche, die den Nutzen für den Patienten ausdrücken, wie Verbesserung der Prognose bezüglich des Überlebens (Lebensdauer, Mortalität) oder Lebensqualität (Beschwerdeminderung). Zu ver-

meiden sind Surrogatendpunkte. Dabei handelt es sich um Endpunkte, die nicht in gesicherter Beziehung zu den genannten patientenrelevanten Endpunkten stehen, wie z.B. Blutzucker, Blutdruck.

### Schritt 2 Informationssuche

Informationen stehen in unterschiedlichen Formaten zur Verfügung. Ausgangspunkt können Originalstudien zum jeweiligem Problem sein. In wachsendem Umfang steht Studienwissen in aufgearbeiteter Form wie systematischen Übersichtsarbeiten und Leitlinien zur Verfügung, häufig in elektronischer Form. Eine wichtige Entwicklung insbesondere in der Medizin sind Online-Lehrbücher, die in kurzen Abständen vollständig aktualisiert werden.

### Originalstudien

Mit einer geschickten Suchstrategie gilt es, die relevanten Studien zu einer Frage zu finden. Dies ist vergleichsweise aufwändig und sicher nicht für jede in der Praxis auftretende Einzelfrage durchführbar. Die Suche erfolgt in Datenbanken. Die größte biomedizinische Datenbank – PubMed (<http://www.pubmed.gov>) – wird von den National Institutes of Health geführt und umfasst im Jahr 2015 mehr als 24 Mio. Einträge. Eine Zusammenstellung von Links relevanter medizinischer, psychologischer und sozialwissenschaftlicher Datenbanken findet sich auf der Website zum Buch.

### Systematische Übersichtsarbeiten

*»A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made.«* (Higgins und Green 2011)

*Eine systematische Übersichtsarbeit versucht, die gesamte empirische Evidenz zusammenzufassen, die über vorab definierte Auswahlkriterien erfasst wird,*

*um eine spezifische Forschungsfrage zu beantworten. Einer systematischen Übersichtsarbeit liegen explizite, systematische Methoden zugrunde, die unter dem Gesichtspunkt ausgewählt werden, Bias zu minimieren. Dadurch stellen sie verlässlichere Ergebnisse zur Verfügung, aus denen Schlussfolgerungen gezogen und Entscheidungen getroffen werden können.*

Eine systematische Übersichtsarbeit ist eine Sekundärstudie, eine »Studie über Studien«. Sie soll die Essenz des Studienwissens zu einer bestimmten Frage darstellen. Durch die Zusammenfassung der Ergebnisse mehrerer Studien zur gleichen Frage sollen die Aussagekraft und die Ergebnissicherheit erhöht werden. Eine *Meta-Analyse* fasst im Rahmen einer systematischen Übersichtsarbeit die Ergebnisse mehrerer Einzelstudien zur gleichen Frage quantitativ zusammen – die Daten werden gepoolt, um ein Gesamtergebnis mit höherer Aussagekraft zu erhalten.

Die Kernelemente einer systematischen Übersichtsarbeit sind

- eine klar formulierte Forschungsfrage
- klare Kriterien für den Einschluss und Ausschluss von Studien
- eine explizite Suchstrategie (Suchbegriffe, Datenbanken)
- kritische Bewertung (critical appraisal) anhand festgelegter methodischer Qualitätsmerkmale
- Reproduzierbarkeit durch Dokumentation der Methodologie.

Die Cochrane Library umfasst im Jahr 2015 über 9000 Cochrane Reviews. Die Kurzfassungen (Abstracts) und die Zusammenfassungen in einfacher englischer Sprache (Cochrane Summaries) sind frei verfügbar. Cochrane Reviews unterliegen einem rigiden Qualitätssicherungsverfahren, das in einem umfangreichen Handbuch festgelegt ist (Online-Version <http://handbook.cochrane.org/v5.0.2>).

Eine Datenbank für systematische Übersichtsarbeiten zu (psycho-)sozialen Interventionen bietet die Campbell Collaboration Library



**Methodischer Hintergrund von Leitlinien:  
S-Klassifikation**

<b>S3</b>	<b>Evidenz- und Konsensbasierte Leitlinie</b>	Repräsentatives Gremium, Systematische Recherche, Auswahl, Bewertung der Literatur, Strukturierte Konsensfindung
<b>S2e</b>	<b>Evidenzbasierte Leitlinie</b>	Systematische Recherche, Auswahl, Bewertung der Literatur
<b>S2k</b>	<b>Konsensbasierte Leitlinie</b>	Repräsentatives Gremium, Strukturierte Konsensfindung
<b>S1</b>	<b>Handlungsempfehlungen von Expertengruppen</b>	Konsensfindung in einem informellen Verfahren

**S  
Y  
S  
T  
E  
M  
A  
T  
I  
K**

Abbildung 3.8 Stufenklassifikation nach dem Regelwerk.

Quelle: AWMF 2012, S. 9

of Systematic Reviews (<http://www.campbellcollaboration.org/lib>).

### Online-Bücher und Kurzzusammenfassungen

In der Medizin verliert das klassische gedruckte Lehrbuch an Bedeutung. An seine Stelle treten zunehmend eBooks und Online-Lehrbücher wie z.B. UpToDate (<http://www.uptodate>). Der Vorteil der Online-Lehrbücher liegt in der Möglichkeit zur fortlaufenden Aktualisierung, den Querverweisen und der Verlinkung mit den Studien, die den Aussagen und Empfehlungen zugrunde liegen.

Strukturierte Zusammenfassungen von Originalstudien, die aus zahlreichen Fachzeitschriften herausgefiltert werden, bieten Publikationen bzw. Datenbanken wie Journal Watch (<http://www.jwatch.org>) und ACP Journal Club (<http://acpjournals.org>) und für die Pflege Evidence Based Nursing (<http://ebn.bmj.com>). Aufgearbeitete Evidenz zur Unterstützung klinischer Entscheidungen bietet Clinical Evidence (<http://clinicalevidence.bmj.com>).

### Leitlinien und HTA-Berichte

Leitlinien werden definiert als »systematisch entwickelte Aussagen, die den gegenwärtigen Erkenntnisstand wiedergeben, um die Entscheidungsfindung von Ärzten und Patienten für eine angemessene Versorgung bei spezifischen Gesundheitsproblemen zu unterstützen.« (AWMF 2012, S. 5)

*Health Technology Assessment (HTA) bezeichnet eine umfassende und systematische Bewertung neuer oder bereits auf dem Markt befindlicher Technologien der medizinischen Versorgung (Medikamente, Medizinprodukte, Prozeduren, Organisationssysteme etc.) hinsichtlich ihrer vor allem medizinischen, sozialen und finanziellen Auswirkungen.*

### Fünf wichtige Quellen für Leitlinien/HTA-Berichte

- Nationale Versorgungsleitlinien <http://www.versorgungsleitlinien.de> *ausgewählte häufige Krankheitsbilder*
- Leitlinien der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften <http://tinyurl.com/q5btqpfz> (*vorzugsweise Entwicklungsstufe S3*)
- Leitlinien des National Institute for Health and Clinical Excellence <http://www.nice.org.uk/guidance> *Leitlinien für den National Health Service, Medizin, Public Health*
- Leitlinien des Scottish Intercollegiate Guidelines Network (SIGN) <http://tinyurl.com/q5zotqy>
- HTA-Berichte Deutsches Institut für Medizinische Dokumentation und Information (DIMDI), <http://tinyurl.com/p67hrwf> *Krankheitsfrüherkennung, Therapie, Rehabilitation, Pflege*

Methodisch hochwertige Leitlinien und HTA-Berichte sind eine wichtige und meist leicht

zugängliche Form aufgearbeiteter Evidenz. Hervorzuheben sind die mit hohem Aufwand und dokumentierter Transparenz erstellten S3-Leitlinien und die Nationalen Versorgungsleitlinien. Die unterschiedlichen Entwicklungsstufen der AWMF-Leitlinien sind in Abb. 22 dargestellt.

Hier einige Schlaglichter auf die Beispielfragen (Recherche Juni 2015):

#### Frage 1

Die Frage, ob Bettruhe oder Aktivierung bei akutem Kreuzschmerz schneller zur Beschwerdefreiheit führt, lässt sich mit der Nationalen Versorgungsleitlinie Kreuzschmerz (<http://www.leitlinien.de/nvl/kreuzschmerz>) beantworten.

#### Frage 2

Eine Suche in Pubmed mit den Begriffen «nursing homes physical restraints», eingegrenzt auf »Randomized Controlled Trial«, ergibt 49 Studien. Besonders relevant erscheint die Studie »Effect of a Guideline-Based Multicomponent Intervention on Use of Physical Restraints in Nursing Homes. A Randomized Controlled Trial«.

### Schritt 3 Kritische Bewertung der Suchergebnisse

Die kritische Bewertung (critical appraisal) soll klären, inwieweit das Ergebnis einer Studie oder Übersichtsarbeit valide und relevant ist. Dafür wird die methodische Qualität geprüft, um die Irrtumswahrscheinlichkeit bzw. die Verlässlichkeit (interne Validität) einschätzen zu können. Die interne Validität ist eine notwendige Voraussetzung für die Prüfung der externen Validität (Schritt 4).

#### Definition

*Interne Validität bezeichnet das Ausmaß, mit dem die Ergebnisse einer Studie den »wahren« Effekt einer Intervention (oder Exposition) wiedergeben, d.h. frei von systematischen Fehlern (Bias) sind.*

Dazu wird die Planung, Durchführung und Erstellung der jeweiligen Studie, Übersichtsarbeit und Leitlinie anhand vorgegebener Qualitätskriterien beurteilt.

Zur Prüfung verschiedener Qualitätsaspekte stehen eine Reihe von Prüfinstrumenten zur Verfügung so z.B.:

- CONSORT (Consolidated Statement of Reporting Trials) zur Beurteilung der methodischen Qualität von RCTs (Moher 2014, S. 80 ff.)
- STROBE (STrengthening the Reporting of OBServational studies in Epidemiology) für Beobachtungsstudien (Moher et al. 2014, S. 169 ff.)
- AGREE (Appraisal of Guidelines Research and Evaluation [www.agreertrust.org](http://www.agreertrust.org)) bzw. DELBI (Deutsches-Leitlinien-Bewertungsinstrument <http://tinyurl.com/ptx6oax>) für die Erstellung und Beurteilung von Leitlinien
- GRADE (Grading of Recommendations Assessment, Development, and Evaluation <http://www.gradeworkinggroup.org>) für Empfehlungen, die in systematischen Übersichtsarbeiten, HTA-Berichten und Leitlinien ausgesprochen werden.

### Schritt 4 Anwendung der ausgewählten und bewerteten Evidenz beim individuellen Fall

Die Gültigkeit der Studienergebnissen für den individuellen Patienten wird als externe Validität bezeichnet.

#### Definition

*Die externe Validität beschreibt die Übertragbarkeit (Generalisierbarkeit) der Studienergebnisse auf Patientengruppen oder Bedingungen der Routineversorgung. Die Generalisierbarkeit hängt zum Beispiel von der Auswahl der Studienteilnehmer und von der Qualifikation der beteiligten Ärzte ab (IQWiG, Glossar, Stichwort Validität online <http://tinyurl.com/of4d8hd>).*

Die Übertragbarkeit der Studienergebnisse auf einen bestimmten Patienten kann angenommen werden, wenn dieser – bezogen auf Merkmalen wie Alter, Geschlecht, Krankheitsstadium, Risikokonstellationen, Multimorbidität und Kontraindikationen – den Patienten der Studienpopulation ähnlich ist. Ist er viel älter oder seine Krankheit weiter fortgeschritten, ist zu überlegen, wie sich dies auf den Behandlungserfolg auswirken kann: So kann vernünftigerweise angenommen werden, dass ein schon geschwächter Krebspatient eine Chemotherapie weniger gut verträgt als die Studienpatienten im Frühstadium der Krankheit. Auch ist zu bedenken, dass Patienten, die an einer Studie teilnehmen, häufig mehr Zeit und Aufmerksamkeit erhalten, besser informiert und aufgeklärt werden und möglicherweise eine stärkere positive Erwartungshaltung entwickeln. Auch kann sich die Durchführung der Therapie unter Studienbedingungen von der Durchführung im Alltag unterscheiden (Dosierung, Begleittherapien, Applikationsart, Training der Behandler (Windeler et al. 2008)).

### Schritt 5 Selbst-Evaluation

Die kritische Reflexion und Überprüfung der Fertigkeiten und Fähigkeiten zum Meistern der vorangegangenen 4 Schritte sind das Anliegen von Schritt 5: Wie gut bin ich im Formulieren beantwortbarer Fragen, in der Literaturrecherche, in der kritischen Bewertung von Studien, im Anwenden der externen Evidenz auf den individuellen, einmaligen Patienten und was tue ich, um meine EBM-Kompetenz weiterzuentwickeln?

Detaillierte Fragen zur Selbst-Evaluation finden sich auf der Website des Centre for Evidence-Based Medicine <http://ktclearinghouse.ca/cebmlpractise/evaluation>.

### Wirksamkeit von Therapieverfahren

Zur Untersuchung der Effekte von Therapieverfahren und zur vergleichenden Untersuchung

zweier oder mehrerer Therapieverfahren ist der faire Vergleich durch eine randomisierte kontrollierte Studie wegen der hohen internen Validität (geringe Irrtumswahrscheinlichkeit) am besten geeignet. Liegen bereits mehrere Studien zu einer Fragestellung vor, können diese in einer systematischen Übersichtsarbeit oder einer Meta-Analyse zusammengefasst werden. Je tiefer die Methoden in der Methodenhierarchie für Therapieverfahren liegen, desto höher die Irrtumswahrscheinlichkeit. Liegt keine Evidenz höherer Stufe vor, kann die Expertenmeinung die beste verfügbare Evidenz darstellen. Viele Fragen in der Medizin sind (noch) nicht in Studien untersucht – daher spielt die auf Logik, Plausibilität und Erfahrung beruhende Expertenmeinung im klinischen Alltag eine wichtige Rolle.

### Evidenzstufen für therapeutischen Methoden

- I a Systematische Übersichtsarbeiten von Studien der Evidenzstufe I b
- I b Randomisierte klinische Studien
- II a Systematische Übersichtsarbeiten von Studien der Evidenzstufe II b
- II b Prospektive vergleichende Kohortenstudien
- III Retrospektive vergleichende Studien
- IV Fallserien und andere nicht vergleichende Studien
- V Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte, u. ä.; nicht mit Studien belegte Meinungen anerkannter Expertinnen und Experten, Berichte von Expertenkomitees und Konsensuskonferenzen (G-BA 2015a, Kapitel 2 § II).

### Evidenzstufen für diagnostische Methoden

- I a Systematische Übersichtsarbeiten von Studien der Evidenzstufe I b
- I b Randomisierte kontrollierte Studien
- I c Andere Interventionsstudien

- II a Systematische Übersichtsarbeiten von Studien zur diagnostischen Testgenauigkeit der Evidenzstufe II b
- II b Querschnitt- und Kohortenstudien, aus denen sich alle diagnostischen Kenngrößen zur Testgenauigkeit (Sensitivität und Spezifität, Wahrscheinlichkeitsverhältnisse, positiver und negativer prädiktiver Wert) berechnen lassen
- III Andere Studien, aus denen sich die diagnostischen Kenngrößen zur Testgenauigkeit (Sensitivität und Spezifität, Wahrscheinlichkeitsverhältnisse) berechnen lassen
- IV Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte, u.ä.; nicht mit Studien belegte Meinungen anerkannter Expertinnen und Experten, Berichte von Expertenkomitees und Konsensuskonferenzen (G-BA 2015b, Kapitel 2 § II).

### Diagnostik

Eine diagnostische Untersuchung hat zum Ziel, das Behandlungsergebnis für einen Patienten zu verbessern, indem das Vorliegen einer Krankheit bestätigt oder ausgeschlossen wird. Für die Fragen nach der Genauigkeit und dem Nutzen diagnostischer Tests stehen unterschiedliche Methoden zur Verfügung. Genauigkeit bedeutet, dass der Test sicher zwischen Gesunden und Kranken unterscheidet, der Test also eine hohe Sensitivität und Spezifität aufweist (S. 218 f.). Dazu kann der Vergleich der Testmethode mit einem »Goldstandard« dienen – wie z.B. der Vergleich von Belastungs-EKG und Koronarangiographie in der Diagnose der koronaren Herzkrankheit, bei der die direkte Betrachtung der Koronargefäße in der Angiographie den Goldstandard darstellt. Die Frage, ob ein (neuer) diagnostischer Test zu besseren Behandlungsergebnissen führt, beantwortet die randomisierte kontrollierte Studie mit der geringsten Irrtumswahrscheinlichkeit.

### Warum?

Lautet die Frage »warum?«, sind befragende Methoden einzusetzen. Um zu verstehen, warum die einen Patienten das vom Arzt verschriebene Medikament einnehmen und andere es in den Müll werfen, müssen die Patienten befragt werden. Dafür stehen verschiedene Formen des Interviews zur Verfügung, wie z.B. das offene und leitfadengestützte Interview, das Tiefeninterview und die Fokusgruppe (eine Form der Gruppendiskussion). Die Ergebnisse qualitativer Studien dienen der Entwicklung und Verbesserung von Theorien sowie der Begriffs- und Konstruktbildung.

Lautet das Ergebnis der qualitativen Studie, dass die Patienten ihre Medikamente einnehmen, die sich über erwünschte und unerwünschte Arzneiwirkungen gut informiert fühlen, und jene nicht, die sich schlecht informiert fühlen, liegt es nahe, eine Intervention zur Verbesserung der Information von Patienten zu entwickeln. Ob diese Intervention funktioniert, sollte dann im nächsten Schritt durch eine randomisierte kontrollierte Studie geprüft werden. Erbringt die Befragung von Patienten beispielsweise, dass der Nicht-Einnahme von Medikamenten Defizite in der Arzt-Patient-Kommunikation und damit einhergehend mangelndes Vertrauen in das Medikament zugrunde liegen, können Items gebildet werden, welche den Informationsstatus und das Vertrauen abbilden. Veränderungen durch Interventionen können dann anhand dieser Items gemessen, also quantitativ untersucht werden.

### Effekte von Umweltexpositionen oder anderen toxischen Expositionen auf die Gesundheit

Die Effekte von Tabakrauch, Lärm oder elektromagnetischer Strahlung sind am besten durch Beobachtungsstudien wie Kohortenstudien und Fall-Kontroll-Studien zu erfassen. Wegen der immer nur näherungsweise kontrollierbaren

Störfaktoren wird es auf einige Fragen wohl nie eine abschließende Antwort geben. Probanden gezielt einer schädlichen Einwirkung auszusetzen und die daraus folgenden Schäden zu messen, ist natürlich nicht zulässig. Keine Ethikkommission würde einer Studie zustimmen, in welcher der Zusammenhang von Feinstaub und Lungenerkrankungen durch den Vergleich einer experimentell exponierten Gruppe im Vergleich zu einer nicht-exponierten Kontrollgruppe untersucht werden soll. Daher stammt die Evidenz zu Fragen wie dieser aus Beobachtungsstudien und natürlichen Experimenten, wie z.B. dem Vergleich der Gesundheit in unterschiedlich exponierten Regionen.

### **Unerwünschte Arzneimittelwirkungen/ Komplikationen von Interventionen**

Unerwünschte Wirkungen und Wechselwirkungen von Arzneimitteln nach der Marktzulassung werden über ein Pharmakovigilanzsystem erfasst, das in erster Linie auf den Meldungen von Ärzten beruht und in Deutschland beim Bundesinstitut für Arzneimittel und Medizinprodukte angesiedelt ist. Die Häufigkeit von Komplikationen, wie die Durchstoßung der Darmwand bei der Darmspiegelung oder die Auslösung eines Herzinfarkts bei Ballondehnung eines Herzkranzgefäßes, kann in Fallserien erfasst werden. Die Risiken einer Strahlenbelastung durch Röntgenaufnahmen können z.B. in Fall-Kontroll-Studien untersucht werden.

### **Krankheitsmechanismen**

Die Erforschung der Mechanismen, die zu Krankheit oder zur Heilung führen, erfordert ein breites Spektrum von Methoden von der Grundlagenforschung bis hin zur klinischen Epidemiologie. Die Erforschung von Therapieeffekten durch RCTs geben zwar Einblicke in den – durch die Therapie unterbrochenen oder zumindest modifizierten – Kausalmechanismus, können ihn jedoch nicht vollständig offenlegen.

### **3.8.3. Exkurs: Pioniere einer evidenzbasierten Praxis**

Das Prinzip der klinischen Studie (clinical trial) ist die Prüfung einer Therapiemethode durch Vergleich. Verglichen werden zwei Gruppen von Kranken. Die Therapiegruppe erhält die zu prüfende Therapie, die Vergleichsgruppe, auch Kontrollgruppe genannt, erhält keine Behandlung, Plazebo oder eine andere Behandlung. Im Laufe der Zeit hat man erkannt, dass die Methode, mit der die Gruppen gebildet werden, von großer Bedeutung für die Validität der Ergebnisse sein kann (S. 172 ff.). Auch das Wissen der Probanden wie auch der beteiligten Forscher über die Zugehörigkeit zur Therapiegruppe oder zur Kontrollgruppe kann das Ergebnis verzerren. Zum Standard der Wirksamkeitsmessung in der Medizin ist die klinische Studie erst in den 1950er und 1960er Jahren geworden. Bis dahin galt die klinische Erfahrung von Experten als wegweisend für die Wirksamkeits- und Nutzenbeurteilung (S. 170 ff.).

Das Verhältnis von klinischen Studien und Erfahrungswissen hat Pickering schon 1949 schlüssig dargelegt. Letztlich ist jede Therapie ein Experiment, bei dem eine oder mehrere Variablen verändert werden, um einen Zustand zu verändern. Der Ausgang des Experiments wird beobachtet und gemessen (Pickering 1949). Im medizinischen Alltag führt der Arzt eine Behandlung bei einem Patienten durch. Als »Kontrolle« für den aktuellen Patienten dient das Ergebnis der Beobachtung ähnlicher Fälle in der Vergangenheit. Dabei vergleicht der Arzt den Verlauf der Krankheit bei seinem aktuellen Patienten mit dem erinnerten Verlauf eines oder mehrerer Patienten in der Vergangenheit – sozusagen eine »erinnerte historische Kontrolle«. Er führte also eine Messung durch, ob er es will oder nicht, ob es ihm bewusst ist oder nicht. Liegt der Messung kein Plan zugrunde, fehlen präzise definierte Endpunkte. Bei der so erhobenen »subjektiven Statistik« sind die

Ergebnisse wenig präzise und zumeist ungeeignet, Unterschiede zwischen Behandlung und Nicht-Behandlung zu erfassen. Die klinische Studie ist hingegen ein geplantes und dokumentiertes Experiment, bei dem viele Patienten gleichzeitig beobachtet und Unterschiede in der Behandlungsgruppe im Vergleich zur Kontrollgruppe anhand vorab definierter Parameter gemessen werden. Da auch in der klinischen Studie jeder Proband individuell und einzigartig ist, streuen die Ergebnisse in beiden Gruppen.

### Frühe Studien mit fairem Vergleich

Vergleichende Experimente mit Plan scheinen eine naheliegende Vorgehensweise zu sein. Beispiele lassen sich bis zu biblischen Zeiten zurückverfolgen. Besonders bemerkenswert erscheinen die homöopathischen Kochsalzversuche in Nürnberg im Jahr 1834, weil hier Struktur-, Behandlungs- und Beobachtungsgleichheit explizit berücksichtigt wurden. Man führte eine doppelt verblindete randomisierte kontrollierte Studie nach allen Regeln der Kunst durch – mehr als 100 Jahre, bevor der Begriff geprägt wurde.

### Studie zu Ernährung und Gesundheit – das Buch Daniel 1

Teil der Heranbildung des biblischen Helden Daniel und seiner drei Freunde am babylonischen Hof war die Ernährung. Daniels Hypothese lautete, dass eine Ernährung aus Gemüse und Wasser gesünder sei als die Ernährung des Königs mit Fleisch und Wein. Die Studie wurde hinter dem Rücken des Königs durchgeführt. Als Vergleichsgruppe (Kontrolle) dienten die Knechte des Aufsehers, die Dauer des Experimentes wurde auf 10 Tage angesetzt, der Endpunkt etwas unbestimmt mit »Aussehen« definiert. Nach den 10 Tagen, so heißt es in Satz 15, »sahen sie schöner und kräftiger aus als alle jungen Leute, die von des Königs Speise aßen« (Neuhauser et al. 2004).

Das Ergebnis wurde umgehend in die Tat umgesetzt – ein Schicksal, dass nicht allen neuen Erkenntnissen vergönnt ist.

### Behandlung von Schusswunden, Ambroise Paré 1537

Schießpulver galt im 16. Jahrhundert als toxisch. Dieser Vorstellung folgend wurden Schusswunden gereinigt, und zwar mit einer Paste, die Holunderöl enthielt und siedend heiß aufgetragen wurde. Im Jahr 1537 behandelte der Chirurg Ambroise Paré einen Soldaten, der infolge der Explosion seines Schießpulverbehälters eine ausgedehnte Verbrennung im Gesichtsbereich erlitten hatte. Eine Gesichtshälfte behandelte er in üblicher Weise (»usual care«), auf die andere Gesichtshälfte trug er – dem Rat einer alten Frau folgend – eine Zwiebelsalbe auf. Das Ergebnis dieses direkt vergleichenden Experiments sprach für die Zwiebelsalbe. Ein zweiter Versuch folgte, als ihm während der Behandlung die Holunderpaste ausging. Aus der Not heraus erhielt ein Teil der Verletzten lediglich die als zweite Schicht dienende Salbe aus Eigelb, Terpentin und Rosenöl. Die Ergebnisse für die neue Therapie waren deutlich besser (Wootton 2006, Kindle Position 2308; Donaldson 2004.). Parés Erkenntnisse wurden interessanterweise von seinen chirurgischen Kollegen angenommen, vermutlich weil die Chirurgie als ein Handwerk ohne wissenschaftlichen Hintergrund galt, die Chirurgen also keine »scientific community« (S. 122 ff.) bildeten und somit kein Dogma den Weg der Evidenz in die Praxis behinderte.

### James Lind und die Skorbut-Studie 1747

Skorbut war eine in der Schifffahrt gefürchtete, unbehandelt häufig zum Tode führende Krankheit, von der wir heute wissen, dass sie durch Vitamin C-Mangel verursacht wird. Ende des 16. Jahrhunderts wiesen einige Erfahrungen bereits auf die Wirksamkeit von Orangen und Zitronen in der Behandlung von Skorbut hin.

Im Jahr 1601 führte der Kapitän James Lancaster im Rahmen einer Ostindien-Fahrt eine – nicht geplante – kontrollierte Studie durch. Auf 3 der 4 Schiffe grassierte Skorbut, nicht jedoch auf seinem eigenen, auf dem die Besatzung jeden Tag 3 Teelöffel Zitronensaft erhalten hatte. Die Admiralität erkannte die Erfolge aber nicht an, weil Skorbut ja wie andere Krankheiten auch nach den damaligen fest gefügten Vorstellungen durch schlechte Luft oder ein Ungleichgewicht der Säfte entstand. Der Einsatz von Zitrusfrüchten wurde daher den englischen Kapitänen nicht empfohlen und Skorbut blieb ein gravierendes Problem. Im Jahr 1747 führte der Schiffsarzt James Lind eine geplante kontrollierte Studie durch. Auf dem Schiff HMS Salisbury waren 80 der 800 Seeleute an Skorbut erkrankt. Lind zog eine Stichprobe von 12 Erkrankten mit möglichst ähnlicher Symptomatik. Zur üblichen Ernährung erhielten je 2 von ihnen die damaligen Standardtherapien: Cider, Vitriol, Essig, Meerwasser, eine Mischung aus Muskat, Senf und Knoblauch. Deutliche und schnelle Effekte stellte er nur bei der Orangen-Zitronen-Therapie fest. Im Jahr 1753 veröffentlichte er sein berühmtes Buch »A Treatise of Scurvy« (Lind 1753), das als frühes Beispiel einer systematischen Übersichtsarbeit gilt (Milne und Chalmers 2004). Auch Lind war im damaligen Miasma-Dogma gefangen und erklärte sich den Effekt durch die Ent-Blockierung der Hautporen, die durch schlechte Luft verschlossen seien. Die Anerkennung seines Erfolgs hielt sich in Grenzen. Erschwerend kam hinzu, dass er als Leiter des größten Marinekrankenhauses die zahlreichen Fälle von Skorbut mit Zitronensaft behandelte, der zuvor erhitzt und dadurch seines wirksamen Bestandteils, dem Vitamin C, beraubt worden war. In seiner medizinischen Praxis blieb er konventionell und behandelte seine Patienten – auch die an Skorbut erkrankten – mit Aderlass. Lind, wenn er denn einer der Wegbereiter der klinischen Studie sei, habe

dann aber nicht viel daraus gemacht, merkt der Medizinhistoriker David Wootton kritisch an (Wootton 2006).

### Mesmerismus um 1790

Der Arzt Franz Anton Mesmer war zutiefst von dem Einfluss der Planeten mittels einer »sehr feinen Materie, welche alle Körper durchdringt« überzeugt. Durch Anwendung dieser Kraft, die er auch als »tierischen Magnetismus« bezeichnete, versprach er Heilung von bisher nicht behandelbaren Krankheiten (Mesmer 1781, S. 2 ff.).

Damit erregte er Ende der 1780er Jahre in Paris große Aufmerksamkeit und rekrutierte – auch durch die Patronage von Königin Marie Antoinette – zahlreiche Patientinnen und Patienten aus der Oberschicht. Das ärztliche Establishment erwirkte daraufhin die Einsetzung einer Kommission zur Prüfung der Heilmethode. Diese hochrangig besetzte Kommission führte in Zusammenarbeit mit einem Schüler Mesmers einige Experimente durch. So »magnetisierten« sie einen von mehreren Bäumen und ließen sie von Probanden umarmen, die nicht wussten, welcher Baum »magnetisiert« war. In ähnlicher Weise wurde Probanden fälschlich mitgeteilt, hinter einer geschlossenen Tür stünde ein Arzt, der seine magnetischen Heilkräfte sende. Im Zusammenhang mit der Anwendung der »magnetischen Heilkräfte« traten bei Patienten häufig »Krisen« (Erregungszustände) mit anschließender Mattigkeit und Benommenheit auf. So auch bei den geschilderten Experimenten, allerdings bei den nicht »magnetisierten« genauso wie bei den »magnetisierten« Bäumen und auch dann, wenn hinter der Tür niemand stand. Bei einem weiteren Experiment fiel auf, dass »Krisen« bei Frauen viel häufiger als bei Männern auftraten. Die Kommission vermutete, dass hier die sexuelle Anziehung eine Rolle spiele – schließlich »heilte« der Magnetiseur ja durch Auflegen der Hand auf die von Krankheit oder Unwohlsein betroffene Körperstelle. Hier wurde also mit ei-

nem geplanten Experiment, das die Technik der Verblindung der Probandinnen einschloss, mit höhergradiger Wahrscheinlichkeit das Fehlen eines spezifischen Effekts belegt.

### **Alexander Hamiltons Versuch zum Aderlass 1809**

Der Aderlass galt für mehr als 2500 Jahre als ein wirksames Heilmittel. Der englische Militärarzt Alexander Hamilton stellte diese Annahme in Frage und führte im Jahr 1809 einen Versuch durch. 366 kranke Soldaten wurden durch alternierende Einteilung auf 3 gleiche Gruppen verteilt. In der Behandlungsgruppe wurden alle Soldaten mit Aderlass behandelt. Die beiden anderen Gruppen erhielten als Kontrollgruppe keinen Aderlass. In der Behandlungsgruppe starben 35 Soldaten, in den Vergleichsgruppen 2 bzw. 4 Soldaten. Der Effekt des Aderlasses erwies sich in diesem Experiment also als sehr stark – die Behandlung brachte vielen Probanden den Tod (Milne und Chalmers 2002). Ob Hamilton die Versuche tatsächlich durchgeführt oder die Ergebnisse für seine Dissertation einfach erfunden hat, ist nicht eindeutig geklärt. Auch wenn die Ergebnisse erfunden sein sollten – die Idee, eine weithin anerkannte und ausgeübte »Heil«-Methode in einer vergleichenden Untersuchung zu testen, erscheint bis heute bemerkenswert.

### **Die homöopathischen Kochsalzversuche von Nürnberg 1835**

»Ich wette zehn gegen eins, dass Sie etwas Ungewöhnliches fühlen werden auch ohne Glauben« – diese Aussage eines Arztes und Homöopathen wurde im Jahr 1834 in einem öffentlich durchgeführten Experiment in Nürnberg geprüft. Es ging um die Effekte einer C30-Verdünnung (von Kochsalz. Eine »Vereinigung wahrheitsliebender Männer« entwarf dafür ein Studiendesign aus Sicht der Homöopathie: Potenzierung), das in vollem Umfang einer doppelt verblindeten randomisierten kontrollierten Studie (Medical Re-

search Council 1948), entsprach – 114 Jahre vor der Streptomycin-Studie (Mesmer 1781, S. 2 ff.), die als erste randomisierte kontrollierte Studie gilt (Klemperer 2011a). Als Studienendpunkte waren definiert: »Gar nichts Ungewöhnliches im Befinden« bzw. »Etwas Ungewöhnliches« im Befinden. Das Ergebnis war negativ, es zeigten sich keine Unterschiede zwischen Therapie- und Kontrollgruppe. Ein Blick in die Schrift von Löhner (1835) lohnt als frühes Beispiel für ein sehr sorgfältig und präzise verfasstes Studienprotokoll.

### **Ignaz Semmelweis und das Kindbettfieber**

Semmelweis (1818–1865), ein in Ungarn geborener und in Wien von 1846 bis 1849 in der Geburtshilfe tätiger Arzt, war dort mit der hohen Sterblichkeit der Mütter am Kindbettfieber konfrontiert. Aufgrund seiner scharfsinnigen Analyse stellte er die Hypothese auf, dass die Ärzte selbst mit ihren kontaminierten Händen die Krankheit übertrugen. Die eindrucksvolle Senkung der Sterblichkeit durch Händedesinfektion mit Chlorkalk bzw. Chlorkalk führte zwar unmittelbar zu einer Verminderung der Mortalität, jedoch nicht zur Anerkennung durch seine Kollegen. Auch hier griff wieder das Muster: Der Erfolg konnte keiner sein, weil er mit der vorherrschenden Theorie nicht übereinstimmte. Das Ignorieren von Studienergebnissen, die nicht in vorgefasste Konzepte passen, wird als »Semmelweis-Reflex« bezeichnet (Website Semmelweis Society <http://tinyurl.com/n86szfx>)

### **Vertiefung:**

- Semmelweis IP (1861). Die Aetiologie, der Begriff und die Prophylaxis des Kindbettfiebers. C.A. Hartleben's Verlags-Expedition. Volltext online Harvard University Library: <http://pds.lib.harvard.edu/pds/view/6591473>
- Semmelweis in der James Lind Library: <http://tinyurl.com/buh8nnt>



### Paul Martini und die Methodenlehre

Martini (1889–1964) veröffentlichte im Jahr 1932 die »Methodenlehre der therapeutischen Untersuchung«. Die darin dargelegten und in weiteren Auflagen vertieften und erweiterten Methoden zur Untersuchung der Wirksamkeit von Therapien unter Einschluss statistischer Methoden gelten als ein Frühwerk evidenzbasierter Medizin.

#### Vertiefung:

Martini P (1932). Methodenlehre (Auszüge). James Lind Library: <http://tinyurl.com/kkld8bh>

### Florence Nightingale

Nightingale (1820–1910) gilt allgemein als Begründerin der modernen Pflege. Ihre Arbeit zeigt deutliche Zeichen einer evidenzbasierten Praxis. Ihre Tätigkeit im Militärkrankenhaus in Scutari (in der Nähe des heutigen Istanbul) während des Krimkrieges (1854–1856) wurde als größtes jemals durchgeführtes Qualitätsentwicklungsprojekt bezeichnet (Neuhauser 2003). Sie reformierte das Krankenhaus und dokumentierte dabei die Ausgangssituation quantitativ (Mortalität 47,2% im Februar 1855), die Veränderungsbereiche (Überbelegung, schlechte Lüftung, Abwässer, Abfall, Komfort), die Prozesse (z.B. Ausheben von Entwässerungsgräben, Abfallbeseitigung, Reinigen der Kleidung, Lüftung) und das Ergebnis (Mortalität 2,2% im Juni 1855). Die militärische Führung wollte ihre Erfolge nicht anerkennen und begründete die Senkung der Sterblichkeit mit geringerer Verletzungsschwere. Nightingale führte geschickte Krankenhausvergleiche durch (z.B.: Mortalität der Soldaten auf der Krim, die zu krank für einen Transport waren: 27/1000, Krankenhaus Scutari vor Reform: 427/1000), drohte der Admiralität mit der Veröffentlichung ihres statistischen Berichts (»Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army Founded Chiefly on the Ex-

periences of the Late War«), verbündete sich mit Queen Victoria und erreichte so die Reform des gesamten britischen Lazarettwesens. Eine statistische Abteilung bezeichnete sie als »absolute Notwendigkeit« (Neuhauser 2003).

### Mary Ellen Richmond, Sozialarbeiterin

In ihrem 1917 erschienenen Buch »Social Diagnosis« befasst sich Richmond ausführlich mit Fragen des Beweises der Wirksamkeit sozialarbeiterischer Interventionen. Die Idee des fairen Vergleichs unter kontrollierten Bedingungen ist ihr genauso nah wie der Gedanke, dass die Anwendung der Methode auf Fragen sozialarbeiterischer Interventionen wünschenswert, aber nicht ganz einfach ist.

»The scientist uses controlled experimentation because he works with material which may be brought under complete control. He may, for instance, till half of an orchard whose physical conditions, soil, grade, exposure, etc., are the same throughout. If the tilled half bears much better than the untilled, he concludes that tilling increases the product of fruit trees. When, however, the farmer in the fable digs in his orchard for buried treasure, and in place of gold finds his promised fortune in an unprecedented yield of fruit, he probably draws no causal inferences whatever« (Richmond 1917, S. 40) »Social work has its own approach to evidence, but as regards the testing of its evidential material it has much to learn from law, medicine, history, logic, and psychology.« (Richmond 1917, S. 49) Der Wissenschaftler führt kontrollierte Experimente durch, weil er mit Material arbeitet, das vollständig kontrolliert werden kann. Er kann beispielsweise die Hälfte eines Obstgartens bearbeiten, dessen Erde, Bodengüte, Sonnenexposition usw. gleich ist. Wenn die bearbeitete Hälfte eine viel bessere Ernte erbringt als die nicht bearbeitete, schließt er daraus, dass die Bearbeitung die Leistung der Obstbäume erhöht. Wenn jedoch der Farmer – wie im Märchen – in seinem Obstgarten nach einem verborgenen Schatz gräbt und es an Stelle des Goldes zu einer unerwartet guten Obsternte kommt, sieht er

*vermutlich keine kausalen Zusammenhänge. (...) Soziale Arbeit hat ihre eigene Herangehensweise an Evidenz. Was aber die Untersuchung von Evidenz in der Sozialen Arbeit angeht, kann sie viel von Recht, Medizin, Geschichte, Logik und Psychologie lernen.*

### Ronald Fisher und das Feld-Experiment

Fisher (1890–1962), Biologe und Statistiker, beschrieb in seiner wegweisenden Schrift »The Arrangement of Field Experiments« die Randomisierung als notwendige Voraussetzung für die Abschätzung der Irrtumswahrscheinlichkeit bzw. die Durchführung des Testes auf statistische Signifikanz. Anhand der Erträge von gedüngten und nicht gedüngten Ackerparzellen führte er aus, ab wann ein Vorteil zugunsten der gedüngten Parzellen auf den Dünger und nicht auf den Zufall zurückzuführen ist.

*»What reason is there to think that, even if no manure had been applied, the acre which actually received it would not still have given the higher yield? (...) This level, which we may call the 5 per cent point, would be indicated, though very roughly, by the greatest chance deviation observed in twenty successive trials.« (Fisher 1926, S. 85) Welche Gründe gab es anzunehmen, dass der Acker, der Dünger erhalten hatte, auch ohne Dünger den höheren Ertrag gebracht hätte? (...) Dieses Maß, das wir den 5-Prozent-Punkt nennen, wäre gegeben, wenn auch sehr grob, als die höchste zufällige Abweichung, die in 20 aufeinander folgenden Versuchen beobachtet wird.*

### Weiterführende Literatur

Moser G, Stöckel S, Kuhn J (Hrsg.): Die statistische Transformation der Erfahrung. Centaurus-Verlag, Heidelberg 2012 (Beiträge zur Geschichte des Evidenzdenkens in der Medizin).

### 3.8.4. Vorbehalte – die Top 4

#### Die vier größten Missverständnisse zur Evidenzbasierten Praxis und Evidenzbasierten Medizin

1. *EBM/EBP erkennen nur randomisierte kontrollierte Studien an*

Falsch! Richtig ist: Evidenz umfasst systematisch generiertes Wissen, das auf seine Validität geprüft wurde. EBM und EBP klassifizieren Wissensquellen entsprechend ihrer Irrtumswahrscheinlichkeit (Evidenzhierarchie). Richtig ist: Der Gemeinsame Bundesausschuss beschließt die Aufnahme neuer Therapien in den Leistungskatalog nur dann, wenn ihr Nutzen belegt ist, und zwar wenn möglich durch Unterlagen der Evidenzstufe I mit patientenbezogenen Endpunkten. Dieser Anspruch an die Evidenz ist eine normative Entscheidung des Gemeinsamen Bundesausschusses und kein der EBM innewohnendes Merkmal. EBM umfasst das volle Spektrum wissenschaftlicher Methoden einschließlich qualitativer Methoden – die Frage bestimmt die Methode!

2. *EBM/EBP ist Teil des naturwissenschaftlichen Paradigmas und daher reduktionistisch.*

Vorab: Forschung kann nicht die ganze Welt auf einmal untersuchen, muss also immer die Komplexität reduzieren. Diese Reduzierung ist nur dann zu kritisieren, wenn sie dem Forschungsgegenstand nicht angemessen ist. Ansonsten: Falsch! Richtig ist: Evidenzbasierung zielt auf die wissenschaftliche Begründung und systematische Überprüfung der Wirksamkeit beruflichen Handelns. Richtig ist auch: Eine trennscharfe Kategorisierung des medizinischen und gesundheitlichen Wissensfundus in naturwissenschaftlich, sozialwissenschaftlich und geisteswissenschaftlich funktioniert nicht. Gesundheit und Krankheit haben stets biologische, psychologische und soziale Ursachen. Daher sind Fragen häufig nur durch das Einbeziehen un-

terschiedlicher Disziplinen formulierbar und beantwortbar. Auch hier bestimmt die Frage die Methode – es gilt also der Grundsatz, dass Methoden gegenstandsadäquat sein müssen. Die Testung von Hypothesen zur Wirksamkeit von Interventionen im Kontrollgruppendesign (z.B. RCT) kann auch in anderen Wissenschaftsbereichen angemessen sein. So wurde im Kontrollgruppendesign untersucht, ob der Erwerb bestimmter Wissensinhalte aus dem Bereich der Physik mit aktivierenden Methoden, angewandt von einem wenig erfahrenen Assistenten (postdoctoral fellow), besser gelingt als mit einem erfahrenen, von Studenten hoch eingeschätzten Dozenten (faculty member) (Deslauriers et al. 2011).

3. *Jeder Patient ist einmalig. Statistiken hingegen normieren den Menschen und helfen daher nicht bei Entscheidungen.*

Richtig ist: Jeder Mensch ist einmalig, aber nicht jeder Patienten reagiert völlig anders als die anderen. Die systematische Erfassung und Auswertung der Behandlungsverläufe und -ergebnisse vieler Patienten ergibt aggregiertes Wissen in Form von Statistiken, das es ermöglicht, unserem einmaligen Patienten die Wahrscheinlichkeit von Nutzen und Schaden einer Behandlung näherungsweise mitzuteilen. Richtig ist auch: Eine genauere Methode der Vorhersage von Behandlungsergebnissen gibt es bislang nicht.

4. *Methodenpluralismus ist besser als Methodenimperialismus.*

Richtig! Aber das gilt unter Berücksichtigung der übergeordneten Regel, dass die Methoden zu den untersuchten Gegenständen passen müssen. Ungeeignete, mit hoher Irrtumswahrscheinlichkeit einhergehende methodische Herangehensweisen als vorteilhaften »Methodenpluralismus« darzustellen, ist der Wissenschaftlichkeit nicht förderlich.

### 3.8.5. Wissenstransfer – die Kluft zwischen Wissen und Handeln überbrücken

In der Medizin und in den weiteren Gesundheitsberufen wird fortlaufend neues Wissen hervorgebracht. Die Veröffentlichung neuer Erkenntnisse führt jedoch häufig nicht zu den erforderlichen Veränderungen in der beruflichen Praxis.

Der Transfer von Wissen in die Praxis als Element der Qualitätsentwicklung gesundheitlicher Versorgung hat sich unter Begriffen wie »knowledge transfer«, »knowledge translation« oder »evidence into practice« zu einem eigenen Wissenschaftszweig mit zahlreichen, aber empirisch erst in Ansätzen überprüften Theorien entwickelt (Straus et al. 2013). Die Evidenz für die Frage des Transfers von Evidenz in die Praxis ist zwar lückenhaft. Deutlich ist jedoch, dass Wissen nur einer unter vielen Faktoren ist, die das Verhalten von Individuen wie auch von Organisationen determinieren.

Eine einfache Theorie der Veränderung geht davon aus, dass wissenschaftliche Erkenntnisse z.B. in Fachzeitschriften veröffentlicht werden, auf diesem Wege zu den Praktikern gelangen und diese daraufhin ihre Vorgehensweisen verändern. Diese Theorie funktioniert in der Praxis eher nicht (NHS Centre for Reviews and Dissemination 1999). Das Konzept von Wissen-Haltung-Verhalten (Knowledge-Attitude-Behavior) beschreibt die Sequenz einer Verhaltensänderung. Folgende Barrieren sind hier in einer systematischen Übersichtsarbeit für die Umsetzung von Leitlinienempfehlungen benannt worden: (Cabana et al. 1999).

#### *Im Bereich des Wissens:*

- ein Mangel an Wissen um das Vorhandensein und die Inhalte von Leitlinien
- unzureichende Verfügbarkeit
- zu großer Umfang

#### *Im Bereich der Haltung:*

- mangelndes Vertrauen in Leitlinien generell

- mangelndes Vertrauen in die Autoren
- fehlende Praxisnähe
- das Gefühl, dass Leitlinien zu rigide sind und »Kochbuchmedizin« darstellen
- Einschränkung der ärztlichen Autonomie
- fehlende Übereinstimmung mit den Inhalten
- Mangel an Motivation, Festhalten an der bisherigen Praxis

#### *Im Bereich des Verhaltens:*

- Nicht-Vereinbarkeit mit Patientenwünschen
- Leitlinien mit gegensätzlichen Empfehlungen
- Mangel an Zeit und Ressourcen
- fehlende oder falsche finanzielle Anreize

Die Lerntheorie untersucht Faktoren, die erklären, warum Personen, Gruppen oder Organisationen bestimmte Verhaltensweisen herausbilden, verändern, beibehalten oder aufgeben. Die Motivation zur Verhaltensänderung wird danach vor allem durch positive Konsequenzen, Verstärkung, Bestätigung und Belohnung gefördert. Als wenig wirksam gelten Bestrafung und Kritik. Der Effekt der positiven Verstärker hängt davon ab, wie erstrebenswert sie denjenigen erscheinen, die sie erhalten. Gelingt es nicht, Motivation zu wecken bzw. zu verstärken, ist die Wahrscheinlichkeit für Verhaltensänderungen gering. Theorien des »Modelllernens« gehen davon aus, dass Problemlösungen aus kulturellen oder sozialen Kontexten übernommen werden, z.B. aus der Familie, der peer group oder von Vorbildern.

Veränderungskonzepte wie Audit, Feedback, Fee-for-service und Beitragsermäßigungen stützen sich auf Elemente von Lerntheorien. Das noch immer gültige Fazit einer systematischen Übersichtsarbeit aus dem Jahr 1999 lautet, dass zwar manche Methoden in manchen Situationen wirksam sind, aber keine Methode in allen Situationen (NHS Centre for Reviews and Dissemination 1999).

Modelle der sozialen Kognition befassen sich mit den Annahmen, Wahrnehmungen und Abwägungen von Vorteilen (benefit) und

mit Barrieren für Verhaltensänderungen, mit Haltungen anderer wichtiger Personen und der Selbstwirksamkeit, also dem Glauben an die eigene Fähigkeit, ein bestimmtes Verhalten zu entwickeln. Hierzu zählt das Modell der Stadien der Veränderungsbereitschaft von Prochaska (Prochaska und DiClemente 1982). In dem Modell der Diffusion von Innovationen von (Rogers 1995) werden Personen nach ihrer Veränderungsbereitschaft in »*innovators, early adopters, early majority, late majority* und *laggards*« unterteilt und danach differenzierte Vorgehensweisen für Veränderungen entwickelt.

Modelle der Organisationsentwicklung fokussieren nicht auf Individuen, sondern auf Organisationen, von denen ebenfalls angenommen wird, dass diese Veränderungsprozesse in verschiedenen Phasen durchlaufen werden. So unterscheidet das Modell von Lewin drei Phasen (Lewin 1963). Unfreezing geht von der Einsicht aus, dass es so wie bisher nicht weitergehen kann. Changing bedeutet die Realisierung der notwendigen Veränderungen und Refreezing deren Stabilisierung durch neue Haltungen, Unterstützung und Verstärkung.

Theorien zu Veränderungsprozessen können hier nur angerissen werden. Deutlich werden soll, dass der Transfer von Evidenz in die Praxis systematische, theoretisch begründete und empirische Erkenntnisse einbeziehende Strategien erfordert.

### **Kritische Würdigung Möglichkeiten von EBM/EBP**

Das EBM/EBP-Konzept

- kann Entscheidungen und Behandlungsergebnisse verbessern
- unterstützt bei der Aufgabe, die Interventionen – medizinische Behandlung, Psychotherapie, psychosoziale Beratung, Krisenintervention usw. – zu identifizieren, die dem Patienten bzw. Klienten den größten Nutzen bringen

- unterstützt eine kritische Haltung und die fortlaufende Beschäftigung mit der Verbesserung beruflichen Handelns
- trägt zur wissenschaftlichen Begründung beruflichen Handelns bei
- korrigiert falsche Annahmen – z.B.: Bettruhe bei akutem unkompliziertem Kreuzschmerz ist schädlich, nicht nützlich
- schärft das Bewusstsein für Bias, der als Folge von Interessenkonflikten, kommerziellen Interessen und Eigeninteressen in allen Phasen des Forschungsprozesses entstehen kann, insbesondere für die Einflussnahme der Industrie auf alle Phasen der Evidenzproduktion und -verbreitung

### Probleme von EBM/EBP

Das EBM/EBP-Konzept

- fokussiert auf die Frage der Wirksamkeit von Interventionen und vernachlässigt tendenziell die Fragen nach den Wirkmechanismen und den zugrunde liegenden Theorien
- verführt dazu, die Frage der Übertragbarkeit von Studienergebnissen auf den konkreten Patienten zu vernachlässigen
- verführt dazu, die vorhandene Evidenz mit der zu treffenden Entscheidung gleichzusetzen – tatsächlich soll Evidenz lediglich einen Teil der Informationen darstellen und Entscheidungen unterstützen, in die stets auch Werte und subjektive Bewertungen eingehen
- verführt zu vorschnellen Bewertungen und Empfehlungen – Lücken in der Evidenz (z.B. über unerwünschte Wirkungen), Bias (z.B. Schlussfolgerungen, die mit Daten nicht übereinstimmen) und die Vorläufigkeit von Evidenz werden übersehen, erwünschte Ergebnisse werden unangemessen positiv bewertet
- kann bei unsicherer Evidenzlage bzw. bei falscher Bewertung vorhandener Evidenz zu schlechten Entscheidungen führen
- verführt zu der falschen Annahme, dass die

Ergebnisse randomisierter kontrollierter Studien verlässlich sind: Mangels Zugang zu Studienprotokoll und Rohdaten ist die interne Validität der meisten Studien nicht ausreichend überprüfbar; dabei gibt es starke Hinweise dafür, dass die Mehrzahl der veröffentlichten Studien schlecht sind (*»Why Most Published Research Findings Are False«*) (Ioannidis 2005)

- verführt ihre Enthusiasten zu Arroganz und Snobismus (Bastian 2004).

## 3.9 Shared Decision Making

### 3.9.1. Definition

*Der englische Begriff Shared Decision Making (SDM) bedeutet gemeinsame Entscheidungsfindung. Die deutsche Übersetzung »partizipative Entscheidungsfindung« hat sich eher nicht durchgesetzt. Der letztendliche Zweck von SDM besteht darin, dem Patienten durch evidenzbasierte Informationen zu ermöglichen, Entscheidungen zu treffen, die seinen geklärten Präferenzen entsprechen.*

### Definition

*Partizipative Entscheidungsfindung/Shared Decision Making ist ein Interaktionsprozess mit dem Ziel, unter gleichberechtigter aktiver Beteiligung von Patient und Arzt auf Basis geteilter Information zu einer gemeinsam verantworteten Übereinkunft zu kommen (Härter 2004).*

Eine Reihe von Besonderheiten prägt die Arzt-Patient-Beziehung. Das Verhältnis ist hierarchisch angelegt. Der Arzt befindet sich aufgrund seiner Expertise in einer Machtposition gegenüber dem Patienten. Der Patient entscheidet häufig nicht aus freien Stücken, dass er die medizinische Dienstleistung in Anspruch nimmt, sondern wird durch seine Beschwerden dazu gezwungen. Im ambulanten System entscheidet sich der Patient für den Arzt, aber nicht umgekehrt – jeder Arzt ist dazu verpflichtet, jeden Patienten zu behandeln. Im Krankenhaus