



WO IST DER BEWEIS?

Plädoyer für eine
evidenzbasierte Medizin



IMOGEN EVANS, HAZEL THORNTON
IAIN CHALMERS, PAUL GLASZIOU

Deutsche Ausgabe
herausgegeben von
Gerd Antes

HUBER



Evans et al.
Wo ist der Beweis?

Verlag Hans Huber
Programmbereich Medizin



© 2013 by Verlag Hans Huber, Hogrefe AG, Bern

Dieses Dokument ist nur für den persönlichen Gebrauch bestimmt und darf in keiner Form vervielfältigt und an Dritte weitergegeben werden.
Aus: Imogen Evans, Hazel Thornton, Iain Chalmers, Paul Glasziou; Wo ist der Beweis? – Plädoyer für eine evidenzbasierte Medizin. 1. Auflage.

Imogen Evans
Hazel Thornton
Iain Chalmers
Paul Glasziou

Wo ist der Beweis?

Plädoyer für eine evidenzbasierte Medizin

Aus dem Englischen von Karin Beifuss

Deutsche Ausgabe herausgegeben von Gerd Antes

Verlag Hans Huber

© 2013 by Verlag Hans Huber, Hogrefe AG, Bern

Dieses Dokument ist nur für den persönlichen Gebrauch bestimmt und darf in keiner Form vervielfältigt und an Dritte weitergegeben werden.
Aus: Imogen Evans, Hazel Thornton, Iain Chalmers, Paul Glasziou; Wo ist der Beweis? – Plädoyer für eine evidenzbasierte Medizin. 1. Auflage.

Lektorat: Dr. Klaus Reinhardt
Herstellung: Jörg Kleine Büning
Umschlaggestaltung: Anzinger Wüschner Rasp, München
Druckvorstufe: punktgenau gmbh, Bühl
Druck und buchbinderische Verarbeitung:
AALEXX Buchproduktion GmbH, Großburgwedel
Printed in Germany

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.



Die Verfasser haben größte Mühe darauf verwandt, dass die therapeutischen Angaben insbesondere von Medikamenten, ihre Dosierungen und Applikationen dem jeweiligen Wissensstand bei der Fertigstellung des Werkes entsprechen. Da jedoch die Medizin als Wissenschaft ständig im Fluss ist und menschliche Irrtümer und Druckfehler nie völlig auszuschließen sind, übernimmt der Verlag für derartige Angaben keine Gewähr. Jeder Anwender ist daher dringend aufgefordert, alle Angaben in eigener Verantwortung auf ihre Richtigkeit zu überprüfen.

Anregungen und Zuschriften an:

Verlag Hans Huber
Lektorat Medizin
Länggass-Strasse 76
CH-3000 Bern 9
Tel: 0041 (0)31 300 4500
verlag@hanshuber.com
www.verlag-hanshuber.com

1. Auflage 2013

© 2013 by Verlag Hans Huber, Hogrefe AG, Bern

ISBN 978-3-456-85245-8

Inhalt

Geleitwort zur deutschsprachigen Ausgabe (Gerd Antes)	9
Geleitwort (Ben Goldacre)	13
Vorwort	19
Einleitung	23
1 Neu – aber auch besser?	31
Warum wir faire Tests von medizinischen Therapien brauchen. Unerwartete negative Wirkungen: Thalidomid, Vioxx, Avandia, mechanische Herzklappen. Zu schön, um wahr zu sein: Herceptin	
2 Erhoffte, aber nicht eingetretene Wirkungen	45
Empfehlungen zur richtigen Schlafposition von Säuglingen – Medikamente gegen Herzrhythmusstörungen – Diethylstilbestrol – Hormonersatztherapie – Nachtkerzenöl bei Ekzemen	
3 Mehr heißt nicht unbedingt besser	55
Intensivtherapien bei Brustkrebs: verstümmelnde Operationen – Knochenmarktransplantation. Den Gedanken wagen, weniger zu tun	

4 Früher ist nicht zwangsläufig besser 67

Lehren aus dem Neuroblastom-Screening. Nutzen und Schaden gegeneinander abwägen: Screenings auf Phenylketonurie, Bauchortenaneurysma, Brustkrebs, Prostatakarzinom und Lungenkrebs. Genetische Tests. Was man mit Screening erreichen will und warum Evidenz so wichtig ist. Ist überhaupt jemand normal?

5 Wie man mit Unsicherheit im Hinblick auf Therapieeffekte umgeht 89

Dramatische Behandlungseffekte: selten und leicht erkennbar. Moderate Behandlungseffekte: normal und nicht ganz so offensichtlich. Wenn Ärzte sich nicht einig sind: Koffein bei Frühgeborenen – Antibiotika bei vorzeitigen Wehen – Brustkrebs. Unsicherheiten in Bezug auf Therapieeffekte thematisieren. Medizinische Versorgung im Rahmen fairer Tests anbieten

6 Faire Tests von Therapien 107

Warum wir faire Therapietests brauchen. Gleiches mit Gleichem vergleichen. Vergleiche von Therapien mit dramatischen Effekten. Vergleiche von Therapien mit mäßigen, aber wichtigen Effekten. Faire Erfassung des Behandlungsergebnisses. Untersuchungen zu unerwünschten Therapiewirkungen

7 Den Zufallsfaktor berücksichtigen 133

Der Zufallsfaktor und das Gesetz der großen Zahl. Wie man beurteilt, welche Rolle der Zufall in fairen Tests gespielt haben könnte. Was bedeutet «signifikanter Unterschied»? Wie man ausreichend große Teilnehmerzahlen gewinnt

8 Bewertung der relevanten verlässlichen Evidenz 141

Kann eine einzige Studie genügen? Systematische Reviews der relevanten zuverlässigen Evidenz. Was passieren kann, wenn nicht alle relevante verlässliche Evidenz ausgewertet wird. Systematische Reviews am Anfang und am Ende neuer Forschung

9 Reglementierung von Therapietests: hilfreich oder hinderlich? 155

Sind Reglementierungssysteme das Richtige für Therapiestudien? Information und Einverständnis. Was Reglementierungssysteme nicht tun

10 Gute, schlechte und überflüssige klinische Forschung . . . 167

Gute Forschung: Schlaganfall, Präeklampsie, HIV-Infektion bei Kindern. Schlechte Forschung: psychiatrische Störungen, Epiduralanalgesie bei Wehenschmerzen. Überflüssige Forschung: Atemnotsyndrom, Schlaganfall, Aprotinin. Verkehrte Forschungsprioritäten

11 Richtige Forschung geht uns alle an 185

Welchen Beitrag können Patienten und Öffentlichkeit zur Verbesserung der Forschung leisten? Mitwirkung von Patienten an der Forschung. Wie Patienten faire Therapiestudien gefährden können. Sind Patientenverbände unabhängig? Zusammenarbeit von Patienten und Wissenschaftlern

12 Wodurch zeichnet sich eine bessere Gesundheitsversorgung aus? 199

Wie könnten die hier gemachten Vorschläge bei Ihnen selbst aussehen? Partizipative Entscheidungsfindung. Fragen zur Umsetzung von wissenschaftlicher Evidenz in die Praxis. Wie geht es weiter?

**13 Aus den richtigen Gründen forschen:
ein Entwurf für eine bessere Zukunft** 221

Leitsätze für bessere Forschung: Die richtigen Forschungsfragen stellen – Forschungsvorhaben richtig planen und durchführen – sämtliche Ergebnisse veröffentlichen und zugänglich machen – unverzerrte und brauchbare Forschungsberichte erstellen. Entwurf für eine bessere Zukunft. Es ist Zeit zu handeln

Literatur 233

Zusätzliche Quellen 247

Über die Autoren 253

Sachregister 255

Geleitwort zur deutschsprachigen Ausgabe

Die Medizin und unsere Gesundheitsversorgung gehören zu den beeindruckendsten Errungenschaften der Menschheit und haben unser Leben gerade im letzten Jahrhundert tiefgehend beeinflusst. Alle Segnungen aus dieser Entwicklung können jedoch nicht verdecken, dass die Medizin uns gleichzeitig auch regelmäßig in negativer Form berührt. Medienberichte über Skandale in der medizinischen Forschung oder über ärztliche Kunstfehler zeigen das ebenso wie unsere Ohnmacht, mit der wir miterleben müssen, wie nahe Angehörige oder Freunde an einer Krebserkrankung sterben, ohne dass die Medizin dies verhindern kann.

Selbst einfache Routinebehandlungen und -eingriffe können unerwartet zu Schäden führen und schlimmstenfalls mit dem Tod enden. Der enorme Nutzen von medizinischen Maßnahmen ist untrennbar mit Risiken verbunden. Die Medizin war, ist und bleibt ein Risikogebiet, dessen müssen wir uns bewusst sein. Entscheidungen und Handlungen sind mit einem unvermeidlichen Maß an Unsicherheit behaftet, das deswegen immer berücksichtigt werden muss.

Die Bewertung von Verfahren geschah traditionell aufgrund der Erfahrung, die ein Arzt in seiner Praxis im Laufe der Jahre sammelte. Während dieses so erworbene Wissen zwangsläufig subjektiv und unsystematisch ist, haben die letzten Jahrzehnte eine beeindruckende Verbesserung gezeigt: Das Wissen über Vor- und Nachteile einzelner medizinischer Verfahren – von Arzneimitteln bis hin zu Operationstechniken – wird immer systematischer durch wissenschaftliche Studien belegt. Heute ist Normalität, dass weltweit jedes Jahr Hunderttausende von Patienten ihre Diagnose oder Behandlung innerhalb von klinischen Studien erhalten, damit die Ergebnis-

se anderen Patienten zu einem späteren Zeitpunkt zur Verfügung stehen. Jedes Jahr werden deswegen zehntausende Studien durchgeführt, in denen Verfahren A bei einer Hälfte und Verfahren B bei der anderen Hälfte der Patienten angewendet wird, um damit die Überlegenheit von A oder B nachzuweisen, gleichzeitig aber auch, um Hinweise auf mögliche Schäden zu bekommen.

Die Information aus den klinischen Studien ist in den letzten Jahren eine mächtige Ergänzung der ärztlichen Erfahrung geworden und hat sogar einen eigenen Namen bekommen: Evidenz. Die evidenzbasierte Medizin oder besser Gesundheitsversorgung ist ein seit 1990 vereinheitlichtes Konzept, unter dem das Wissen aus klinischen Studien systematisch zur Entscheidungsunterstützung herangezogen wird. Nutzer hiervon ist in erster Linie der Arzt. Heute werden Entscheidungen jedoch zunehmend nicht mehr vom Arzt allein gefällt, sondern «partizipativ», d. h. gemeinsam mit dem Patienten und oft auch den Angehörigen. Diese partizipative Entscheidungsfindung – mit dem Arzt auf Augenhöhe – erfordert ein grundlegend anderes Verständnis der Rollen auf beiden Seiten mit wesentlich mehr Verantwortung aufseiten der Patienten, als das traditionell der Fall war. Das erfordert nicht nur die Bereitschaft, diese Verantwortung zu tragen, sondern auch zumindest grundlegende Fähigkeiten, die Qualität von Behandlungen so weit zu verstehen und einschätzen zu können, dass die notwendigen Entscheidungen mit dem Arzt geteilt werden können.

Damit kommen auf den Patienten völlig neue Anforderungen zu, die nicht leicht zu bewältigen sind. Mehr Kompetenz für die eigene Gesundheit zu entwickeln, ist für viele Menschen reizvoll, wenn sie intensiver darüber nachgedacht haben. Daraus entwickelt sich zunehmend mehr Interesse an diesen Themen. Die alles überlagernde Schlüsselfrage ist dabei, ob die möglichen Therapien ihre Wirksamkeit und Unbedenklichkeit in geeigneten klinischen Studien bewiesen haben. Das bedeutet keine Abwertung der oft jahrelangen Vorarbeiten in Theorie, im Labor und im Tierversuch, die die Voraussetzungen für die meisten Verfahren schaffen. Es ist nur die Einsicht, dass für den Erfolg der eigenen Behandlung nur der letzte Schritt entscheidend ist, und das ist der Nachweis in einer klinischen Studie.

Ben Goldacre beschreibt in seinem folgenden Geleitwort eindrücklich die vielfältigen Gefahren, die sowohl bei der Durchführung wie auch in der nachträglichen Verarbeitung und Verbreitung der Ergebnisse lauern. Eine der Hauptfehlerquellen liegt in Erwartungen, Wünschen und Glauben aller Beteiligten, die dem objektiven Wissen aus Studien im Wege stehen und oft zu stark verfälschten Ergebnissen führen. Diese in den letzten Jahrzehnten intensiv untersuchten Fehler haben zu einer Reihe strikter Regeln geführt, die die Studiendurchführung und Berichterstattung aufwendig und anstrengend machen. Sie sind jedoch unbedingt notwendig und unverzichtbar, um verlässliche Studienergebnisse zu erhalten, von denen in der Folge Gesundheit und Leben anderer Patienten abhängen.

Dieses Buch ist der konsequente Versuch, die strengen Regeln für verlässliche Information für Patienten und Betroffene verständlich und nutzbar zu machen. Dass offensichtlich haltlosen Versprechungen von Eiferern und Scharlatanen nicht getraut werden kann, haben die meisten Menschen verstanden. Dass man jedoch auch seriösen Wissenschaftlern, Ärzten, Krankenkassen, Behörden und gesundheitspolitischen Organisationen nicht blind vertrauen sollte, ist weit weniger leicht einsichtig. Täglich verbreitete Empfehlungen und Versprechungen zeigen das jedoch sofort: Neue Behandlungen sind nicht automatisch besser; mehr ist nicht notwendigerweise besser; das frühe Erkennen einer Krankheit ist keinesfalls immer besser; und etabliert bedeutet nicht immer bewährt!

Behauptungen dieser Art führen uns immer wieder in die Irre, ihre häufige Wiederholung macht sie nicht richtiger. Leider beherrschen sie auch das Denken vieler Patienten. In den folgenden Kapiteln wird erklärt, warum vieles, was scheinbar plausibel ist, nicht stimmt und uns zu falschen Entscheidungen führt. Neben den systematischen Fehlern, die dazu führen, werden auch die Zufallsfehler erläutert, die in Studien an Menschen unvermeidlich sind und das Geschehen noch weiter komplizieren. Gerade weil Studien an uns Menschen so fehlerträchtig sind, muss größter Wert auf strenge wissenschaftliche Regeln bei der Durchführung gelegt werden. Diese nicht nur grob zu verstehen, sondern auch die Notwendigkeit und den Sinn einzusehen, sollte nach dem Lesen dieses Buches sehr viel leichter möglich sein.

Wiederholt wurde der Vorläufer dieses Buchs als «zu englisch» kritisiert. Die Kritik ist einerseits berechtigt, andererseits jedoch nicht abzustellen. Die angelsächsischen Länder haben, vor allem durch ihre um Größenordnungen umfangreicheren finanziellen Mittel für dieses Themenfeld, einen jahrelangen Vorsprung gegenüber Deutschland. Das bezieht sich sowohl auf die Anzahl der mit öffentlichen Mitteln durchgeführten Studien als auch auf die systematische Berücksichtigung von Studienergebnissen in der Gesundheitsversorgung, also der Praxis der evidenzbasierten Medizin. Insofern müssen wir mit dem Kompromiss leben, das Verständnis von klinischen Studien an Studien aus anderen Ländern zu erläutern.

Neben den Studien aus englischsprachigen Ländern ist auch die Übertragung der Terminologie in die deutsche Sprache eine stete Herausforderung, die viel Aufwand erfordert. Für ihre tatkräftige Mithilfe bedanke ich mich bei Caroline Mavergames und Britta Lang, die wesentlich zur Qualität der deutschen Ausgabe beigetragen haben.

Februar 2013

Gerd Antes

Geleitwort

In der Medizin sollte es nicht um Autorität gehen, und die wichtigste Frage, die man bezüglich jeder Behauptung stellen kann, ist ganz einfach. Sie lautet: «Woher wissen Sie das?» Diese Frage soll im vorliegenden Buch beantwortet werden.

Die Beziehung zwischen den verschiedenen medizinischen Berufsgruppen und Patienten hat sich enorm gewandelt. In ferner Vergangenheit erschöpfte sich das Training der «Kommunikationsfertigkeiten» von Ärzten, soweit überhaupt vorhanden, darin zu lernen, wie man einem Patienten verschweigt, dass er an Krebs versterben würde. Heute bringen wir den Studierenden bei – und dies ist ein direktes Zitat aus den entsprechenden Handouts –, «wie man gemeinsam mit dem Patienten auf ein optimales Gesundheitsergebnis hinarbeitet». Wenn sich die Medizin heutzutage von ihrer besten Seite zeigt, dann werden die Patienten, wenn sie wollen, in die Besprechung und Auswahl ihrer Therapien eingebunden.

Damit es aber dazu kommen kann, muss jedem klar sein, woher wir wissen, ob eine Therapie wirksam ist, woher wir wissen, ob sie schädliche Wirkungen hat und wie wir ihre Vor- und Nachteile gegeneinander abwägen können, um das jeweilige Risiko zu bestimmen. Leider fallen Ärzte, genauso wie alle anderen auch, diesbezüglich mitunter hinter die Erwartungen zurück. Und was noch bedauerlicher ist: Da draußen gibt es ganze Heerscharen, die nur darauf warten, uns in die Irre zu führen.

Doch in allererster Linie sind wir selbst es, die sich täuschen können. Die meisten Krankheiten haben einen natürlichen Verlauf: Sie bessern oder verschlechtern sich periodisch oder auch rein zufällig. Wenn die Symptome sehr schlimm sind, dann kann deshalb auch alles, was man ggf. dagegen unternimmt, den Anschein erwecken,

als habe diese Therapie Wirkung gezeigt, weil sich der Zustand des Patienten ohnehin gebessert hätte.

Ebenso kann uns auch der Placebo-Effekt in die Irre führen: In manchen Fällen kann es tatsächlich vorkommen, dass es dem Patienten einfach deswegen besser geht, weil er ein Scheinmedikament ohne aktive Wirkstoffe genommen hat und glaubt, dass es ihm geholfen hat. Robert M. Pirsig drückte dies in *Zen und die Kunst ein Motorrad zu warten* folgendermaßen aus: «Der eigentliche Zweck der wissenschaftlichen Methode ist es, sich zu vergewissern, ob die Natur einen nicht zu der falschen Annahme verleitet hat, man wüsste etwas, was man in Wirklichkeit nicht weiß.»

Und dann gibt es auch noch diejenigen, die gegen wissenschaftliche Studien zu Felde ziehen. Wenn dieses Buch eine wichtige Botschaft vermittelt, dann ist es das Konzept des «fairen Testens» – einen Begriff, den ich selber von anderen übernommen und bereits endlose Male benutzt habe. Nicht alle Studien sind aus dem gleichen Holz geschnitzt, denn es gibt so viele Möglichkeiten, wie eine wissenschaftliche Forschungsarbeit verzerrt werden kann und irrtümlicherweise eine Antwort gibt, die irgendwer irgendwo für die «richtige» hält.

Manchmal gehen solche Verzerrungen nur auf Unaufmerksamkeit zurück oder entspringen sogar den lautersten Motiven. Ärzte, Patienten, Professoren, Pflegekräfte, Beschäftigungstherapeuten und Gesundheitsmanager – sie alle können der Idee aufsitzen, dass die eine Therapie, in die sie persönlich so viel Energie gesteckt haben, die einzig wahre ist.

Zuweilen kommt es auch aus anderen Gründen zu einer Verzerrung der Evidenz. Es wäre falsch, sich den oberflächlichen Verschwörungstheorien über die pharmazeutische Industrie anzuschließen: Sie hat gewaltige, lebensrettende Fortschritte erzielt. Aber in manchen Bereichen der Forschung geht es um viel Geld, und aus Gründen, die Sie in diesem Buch nachlesen können, werden 90 % der Studien von der Industrie durchgeführt. Dies kann zu einem Problem werden, wenn die von der Industrie gesponserten Studien für das Medikament des Sponsors viermal häufiger zu einem positiven Ergebnis gelangen als unabhängig finanzierte Studien. Ein neues Medikament in den Handel zu bringen, kostet bis zu 800 Millio-

nen Dollar: Der Löwenanteil davon muss aufgebracht werden, bevor das Medikament überhaupt auf den Markt kommt, und wenn sich dann herausstellt, dass es nichts taugt, ist dieses Geld weg. Wo so viel auf dem Spiel steht, können die Idealvorstellungen von einem fairen Test schon mal baden gehen.¹

Ebenso kann auch die Art und Weise, wie Evidenz kommuniziert wird, verzerrend und irreführend sein. Manchmal liegt das daran, wie die Fakten und Zahlen dargestellt werden: Mitunter geben sie nur einen Teil der Geschichte wieder, beschönigen Schwachstellen und picken sich die wissenschaftlichen Belege heraus, die eine Therapie in einem besonders günstigen Licht erscheinen lassen.

Doch in der Alltagskultur können auch noch sehr viel interessantere Prozesse am Werk sein. Verständlicherweise hegen wir den Wunsch nach Wunderheilung, selbst wenn es in der Forschung häufig nur um bescheidene Verbesserungen, um Risikoverringering und knappe Ermessensentscheidungen geht. In den Medien geht dieser Aspekt leider nur allzu oft in einem Schwall von Wörtern wie «Heilung», «Wunder», «Hoffnung», «Durchbruch» und «Opfer» unter.²

In einer Zeit, in der es so vielen Menschen wichtig ist, die Kontrolle über ihr eigenes Leben zu haben und an Entscheidungen über ihre Gesundheitsversorgung beteiligt zu sein, ist es besonders schmerzlich, mit ansehen zu müssen, wie viele verzerrte Informationen es gibt, denn sie machen ohnmächtig. Manchmal geht es bei diesen verzerrten Darstellungen um ein bestimmtes Medikament: Das vielleicht überzeugendste Beispiel aus der jüngeren Vergangenheit ist die Darstellung von Herceptin in den britischen Medien als Wunderdroge gegen Brustkrebs.³

Gelegentlich können diese Eiferer und ihre Freunde in den Medien, wenn sie ihre eigenen Produkte (Therapien) bewerben und die gegen sie sprechenden Belege in Zweifel ziehen, aber auch noch größeres Unheil anrichten, indem sie nämlich unsere Auffassung darüber, woher wir wissen, dass etwas gut oder schlecht für uns ist, aktiv untergraben.

Homöopathische Zuckerpillen schneiden nicht besser ab als Schein-Zuckerpillen, wenn sie in wirklich fairen Tests miteinander verglichen werden. Doch wenn man Homöopathen mit diesen Er-

gebnissen konfrontiert, führen sie ins Feld, dass mit dem ganzen Konzept der Durchführung von Studien etwas nicht stimmen könne, dass es irgendeinen komplizierten Grund gäbe, warum ausgerechnet ihre Pillen als einzige nicht getestet werden können. Auch Politiker warten, wenn man sie mit den Belegen für das Versagen ihres bevorzugten Schulungsprogramms zur Verhütung von Teenagerschwangerschaften konfrontiert, mitunter mit dieser besonderen Art von Verteidigung auf. In Wahrheit kann aber, wie dieses Buch zeigen wird, jede Behauptung, die im Hinblick auf die Wirkung einer Maßnahme (Intervention) aufgestellt wird, einem transparenten, fairen Test unterzogen werden.⁴

Gelegentlich können diese Verzerrungen das öffentliche Meinungsbild sogar noch tiefer untergraben. So ergab eine neuere «systematische Übersichtsarbeit» zu allen fairen und unverzerrten Tests von antioxidativ wirkenden Vitamintabletten, dass es keine Hinweise darauf gibt, dass ihre Einnahme das Leben verlängern kann (tatsächlich können sie es sogar verkürzen). Bei dieser Art von systematischer Zusammenfassung werden – wie dieses Buch wunderbar erklärt – klare Regeln befolgt: So wird dargelegt, wo nach der Evidenz gesucht, welche Evidenz eingeschlossen und wie ihre Qualität bewertet werden sollte. Aber wenn systematische Übersichtsarbeiten zu einem Ergebnis kommen, das die Behauptungen über antioxidativ wirkende Nahrungsergänzungsmittel infrage stellt, sind Unternehmen, Zeitungen und Illustrierte voll der falschen Kritik: Sie behaupten, dass für diese systematische Übersichtsarbeit aufgrund bestimmter politischer Überzeugungen oder offener Korruption Studien bewusst herausgepickt und positive Belege absichtlich ignoriert worden seien usw.⁵

Das ist bedauerlich. Denn das Konzept der systematischen Übersichtsarbeit gehört – wenn man die Gesamtheit der Evidenz im Blick hat – in der Medizin eigentlich zu den wichtigsten Errungenschaften der letzten 30 Jahre. Während sie ihre kleine Geschäftsnische verteidigen, können Journalisten und Arzneimittelunternehmen dadurch, dass sie den öffentlichen Zugang zu diesen Ideen unterlaufen, uns allen einen Bärenienst erweisen.

Und genau da liegt das Problem. Es gibt viele Gründe, warum man dieses Buch lesen sollte. Im einfachsten Fall wird es Ihnen hel-

fen, hinsichtlich Ihrer eigenen Gesundheit fundiertere Entscheidungen zu treffen. Wenn Sie im Gesundheitswesen arbeiten, wird das, was Sie in den folgenden Kapitel lesen werden, allem, was sie über evidenzbasierte Medizin bislang gelernt haben, haushoch überlegen sein. Auf die Bevölkerung bezogen bedeutet das: Wenn mehr Menschen verstehen, was einen fairen Vergleich ausmacht, und erkennen können, ob eine Intervention besser ist als eine andere, dann könnte sich die Öffentlichkeit – so die Argumentation der Autoren – aktiv für eine stärkere Mitwirkung beim Abbau von Unsicherheiten bei für sie wichtigen medizinischen Behandlungen einsetzen, anstatt die Forschung mitunter zu fürchten.

Aber es gibt noch einen weiteren Grund, dieses Buch zu lesen, und zwar um die Tricks der Branche kennenzulernen, und dieser Grund hat nichts mit praktischen Überlegungen zu tun: Tatsache ist, dass dieses Thema interessant, herrlich und vernünftig ist. Und in diesem Buch wird der Stoff besser erklärt, als es mir jemals untergekommen ist, weil die Menschen, die es geschrieben haben, über Erfahrung, Wissen und Empathie verfügen.

Wo ist der Beweis? rückt den menschlichen Aspekt in den Fokus realer Fragestellungen. In der Medizin geht es um das Leiden und den Tod von Menschen, aber auch um die menschlichen Schwächen bei Entscheidungsträgern und Forschern: Und das wird in diesem Buch in den persönlichen Geschichten und Zweifeln von Wissenschaftlern, ihren Motiven, Anliegen und Meinungsänderungen eingefangen. Nur selten wird diese Seite der Wissenschaft der Öffentlichkeit offenbart. Dabei bedienen sich die Autoren sowohl des ernsthaften wissenschaftlichen Schrifttums als auch der kurzweiligeren Sparten der medizinischen Literatur. So ist es ihnen mit diesem Buch gelungen, aus den Diskussionssträngen, die wissenschaftlichen Beiträgen, Kommentaren, Autobiographien und Randnotizen zugrunde liegen, bislang ungehobene Schätze ans Tageslicht zu fördern.

Dieses Buch gehört in jede Schule und alle ärztlichen Wartezimmer. Bis dahin liegt es in Ihrer Hand. Lesen Sie es.

August 2011

Ben Goldacre

Vorwort

Die erste englische Ausgabe von *Testing Treatments*, die 2006 erschien, ging von folgender Frage aus: «Wie kann sichergestellt werden, dass Therapieforschung den Bedürfnissen der Patienten Rechnung trägt?» Unsere kollektiven Erfahrungen – kollektiv zu diesem Zeitpunkt bedeutete, dass Imogen Evans Ärztin (davor Wissenschaftlerin und Journalistin) war, Hazel Thornton eine Patientin und unabhängige Laienvertreterin für Qualität in der Forschung und im Gesundheitswesen und Iain Chalmers aus der Gesundheitsversorgungsforschung stammte – hatten gezeigt, dass die Forschung in diesem wichtigen Punkt oftmals versagte. Darüber hinaus war uns nur allzu klar, dass viele (sowohl neue als auch alte) Therapien nicht durch solide Evidenz abgesichert waren. Deshalb haben wir ein Buch geschrieben, um durch die Förderung des Dialogs zwischen Patienten und Ärzten sowie anderen Gesundheitsberufen zu einer kritischeren öffentlichen Bewertung von Therapieeffekten beizutragen.

Ermutigt wurden wir zum einen durch das Interesse, das *Testing Treatments* bei den Lesern geweckt hat (sowohl an der *British-Library*-Originalausgabe als auch an der Online-Version, die unter www.jameslindlibrary.org kostenfrei zugänglich gemacht wurde), zum anderen aber auch durch die Tatsache, dass es sowohl Laien als auch Fachleser ansprach. In vielen Ländern fand die erste Ausgabe von *Testing Treatments* als Unterrichtsmaterial Verwendung, und inzwischen stehen unter www.testingtreatments.org vollständige Übersetzungen zum kostenlosen Herunterladen zur Verfügung.

Es war von Anfang an klar, dass *Testing Treatments* ein *work in progress* sein würde; denn es würde fast immer irgendeine Unsicherheit über die Wirkungen von Therapien bestehen bleiben – ganz

gleich, ob es sich dabei nun um neue oder bekannte Therapien handelt – und deshalb auch weiterhin nötig sein, alle Therapien ausreichend zu testen. Zu diesem Zweck ist es unverzichtbar, sich wieder und immer wieder mit den wissenschaftlichen Erkenntnissen (Evidenz) auseinanderzusetzen, die vorhandene Evidenz kritisch und systematisch auszuwerten, bevor man sich neuen Forschungsvorhaben zuwendet, und auch neue Ergebnisse vor dem Hintergrund aktualisierter systematischer Übersichtsarbeiten zu interpretieren.

In der zweiten Auflage von *Testing Treatments* wurden aus drei vier Autoren. Neu hinzugekommen ist der Allgemeinmediziner und Wissenschaftler Paul Glasziou, der es sich zur Aufgabe gemacht hat, hochwertigen wissenschaftlichen Erkenntnissen auch im Praxisalltag Rechnung zu tragen. Mit Pinter & Martin haben wir auch einen neuen Verlag gefunden, der die erste Auflage 2010 nachdruckte. Die neue Auflage wird, wie auch schon vorher, kostenlos unter www.testingtreatments.org zugänglich gemacht.

In dieser Neuauflage bleiben wir unseren Grundgedanken treu, doch haben wir den ursprünglichen Text umfassend überarbeitet und aktualisiert. So wurden beispielsweise die Abschnitte über Vor- und Nachteile des Screenings in einem eigenen Kapitel (Kapitel 4) unter der Überschrift «Früher ist nicht zwangsläufig besser» erheblich erweitert. Und in Kapitel 9 («Reglementierung von Therapietests: hilfreich oder hinderlich?») beschreiben wir, wie es dazu kommen kann, dass Forschung zum Schaden der Patienten überkontrolliert wird. Im vorletzten Kapitel (Kapitel 12) stellen wir die Frage «Wodurch zeichnet sich eine bessere Gesundheitsversorgung aus?» und zeigen, wie man die verschiedenen Erkenntnisstränge so zusammenführen kann, dass sich für uns alle entscheidende Verbesserungen erzielen lassen. Den Abschluss bilden unser Entwurf für eine bessere Zukunft und ein Aktionsplan (Kapitel 13).

Wir hoffen, dass unser Buch zu einem besseren Verständnis darüber beiträgt, wie Therapien fair getestet werden können und auch fair getestet sollten und welchen Beitrag wir selbst dazu leisten können, damit dies auch passiert. Dieses Buch ist kein «Führer zu den besten Behandlungen» und kein Ratgeber zu den Wirkungen einzelner Therapien. Vielmehr thematisieren wir gezielt die Aspekte, die eine grundlegende Voraussetzung dafür sind, dass Forschung ein

solides Fundament hat, richtig durchgeführt wird, zwischen schädlichen und nützlichen Therapien unterscheiden kann und dazu ausgelegt ist, die Fragen zu beantworten, auf die es den Patienten, der Öffentlichkeit sowie Ärzten und anderen Gesundheitsberufen ankommt.

Danksagungen

Für ihre wertvollen Kommentare und anderen Beiträge, die uns bei der Bearbeitung der zweiten Auflage von *Testing Treatments* geholfen haben, danken wir folgenden Personen:

Claire Allen, Doug Altman, Patricia Atkinson, Alexandra Barratt, Paul Barrow, Ben Bauer, Michael Baum, Sarah Boseley, Joan Box, Anne Brice, Rebecca Brice, Amanda Burls, Hamish Chalmers, Jan Chalmers, Yao-long Chen, Olivia Clarke, Catrin Comeau, Rhiannon Comeau, Katherine Cowan, John Critchlow, Sally Crowe, Philipp Dahm, Chris Del Mar, Jenny Doust, Mary Dixon-Woods, Ben Djulbegovic, Iain Donaldson, George Ebers, Diana Elbourne, Murray Enkin, Chrissy Erueti, Curt Furberg, Mark Fenton, Lester Firkins, Peter Götzsche, Muir Gray, Sally Green, Susan Green, Ben Goldacre, Metin Gülmezoğlu, Andrew Herxheimer, Jini Hetherington, Julian Higgins, Jenny Hirst, Jeremy Howick, Les Irwig, Ray Jobling, Bethan Jones, Karsten Juhl Jørgensen, Bridget Kenner, Debbie Kennett, Gill Lever, Alessandro Liberati, Howard Mann, Tom Marshall, Robert Matthews, Margaret McCartney, Dominic McDonald, Scott Metcalfe, Iain Milne, Martin McKee, Sarah Moore, Daniel Nicolae, Andy Oxman, Kay Pattison, Angela Raffle, June Raine, Jake Ranson, James Read, Kiley Richmond, Ian Roberts, Nick Ross, Peter Rothwell, Karen Sandler, Emily Savage-Smith, Marion Savage-Smith, John Scadding, Lisa Schwartz, Haleema Shakur, Ruth Silverman, Ann Southwell, Pete Spain, Mark Starr, Melissa Sweet, Tilli Tansey, Tom Treasure, Ulrich Tröhler, Liz Trotman, Liz Wager, Renee Watson, James Watt, Hywel Williams, Norman Williams, Steven Woloshin, Eleanor Woods, and Ke-hu Yang.

Iain Chalmers und Paul Glasziou danken dem *National Institute for Health Research* (Großbritannien) für Unterstützung. Darüber

hinaus ist Paul Glasziou auch dem *National Health and Medical Research Council* (Australien) zu Dank verpflichtet.

Ein besonderer Dank geht an unseren Verleger Martin Wagner von Pinter & Martin, der uns stets Nachsicht und aufmunternde Unterstützung entgegengebracht und jederzeit einen kühlen Kopf bewahrt hat.

August 2011

Imogen Evans, Hazel Thornton, Iain Chalmers und Paul Glasziou

Einleitung

Es gibt keine Möglichkeit herauszufinden, wann unsere Beobachtungen über komplexe Ereignisse in der Natur vollständig sind. Unser Wissen ist begrenzt, wie Karl Popper betont hat, während unsere Unwissenheit grenzenlos ist. In der Medizin können wir uns niemals sicher sein, welche Folgen unser Handeln haben wird. Wir können lediglich das Ausmaß der Unsicherheit eingrenzen. Dieses Eingeständnis ist nicht so pessimistisch, wie es klingt: Behauptungen, die wiederholten hartnäckigen Infragestellungen standhalten, erweisen sich oftmals als recht zuverlässig. Solche ‚sich ändernden Wahrheiten‘ sind die Bausteine für die einigermaßen festen Strukturen, auf die sich unser ärztliches Handeln am Krankenbett tagtäglich stützen kann.

*(William A. Silverman. Where's the evidence?
Oxford: Oxford University Press, 1998, S. 165)*

Die moderne Medizin kann ungeheure Erfolge verzeichnen. Es ist nur schwer vorstellbar, wie das Leben ohne Antibiotika ausgesehen haben muss. Die Entwicklung weiterer wirksamer Arzneimittel hat die Behandlung von Herzinfarkten und Bluthochdruck revolutioniert und das Leben an Schizophrenie erkrankter Menschen verändert. Die Impfung von Kindern hat Polio und Diphtherie in den meisten Ländern zu einer fernen Erinnerung verblassen lassen, und dank künstlicher Gelenke leiden zahllose Menschen weniger stark an Schmerzen und Beeinträchtigungen. Die modernen Bildgebungsverfahren wie Ultraschall, Computertomografie (CT) und Magnetresonanztomografie (MRT, «Kernspin») haben dazu beigetragen, dass genaue Diagnosen gestellt werden können und die Patienten die richtige Behandlung erhalten. Früher kam die Diagnose zahlreicher Krebserkrankungen einem Todesurteil gleich, während

Patienten heutzutage regelrecht mit ihrer Krebserkrankung leben anstatt daran zu versterben. Und weitgehend hat sich auch HIV/AIDS von einer rasch tödlichen zu einer chronischen (lang dauernden) Krankheit gewandelt.

Natürlich sind viele Verbesserungen im Gesundheitswesen aufgrund von sozialen Fortschritten und Fortschritten im Bereich der öffentlichen Gesundheit wie etwa sauberes Leitungswasser, Kanalisation und besseren Wohnverhältnissen zustande gekommen. Doch selbst Skeptikern würde es schwerfallen, die beeindruckenden Einflüsse der modernen medizinischen Versorgung zu leugnen. Während des letzten halben Jahrhunderts hat eine bessere Gesundheitsversorgung wesentlich zu einer höheren Lebenserwartung beigetragen und die Lebensqualität vor allem chronisch kranker Menschen verbessert.^{1,2}

Doch die Triumphe der modernen Medizin können uns leicht dazu verführen, viele der aktuellen Probleme zu übersehen. Selbst heute kann sich ein viel zu großer Teil der medizinischen Entscheidungsfindung nicht auf ausreichende Belege stützen. Es gibt immer noch zu viele Therapien, durch die Patienten Schaden nehmen können, etliche, die keinen Nutzen haben oder deren Nutzen nicht nachgewiesen ist, und wieder andere, welche die Mühe lohnen, die aber zu selten angewendet werden. Wie kann das sein, wenn doch jedes Jahr Studien über die Wirkungen von Therapien einen ganzen Berg von Ergebnissen erzeugen? Leider sind die Erkenntnisse oftmals unzuverlässig, und außerdem befasst sich ein Großteil der Forschung nicht mit den Fragen, die aus Sicht der Patienten beantwortet werden müssten.

Ein Teil des Problems ist, dass Therapieeffekte nur in ganz seltenen Fällen klar auf der Hand liegen oder dramatisch ausfallen. Stattdessen herrscht im Normalfall Unsicherheit darüber, wie gut neue Therapien wirken oder ob sie vielleicht sogar mehr schaden als nützen. Um Therapieeffekte zuverlässig zu identifizieren, sind deshalb sorgfältig geplante faire Tests vonnöten – Tests also, die darauf ausgelegt sind, systematische Fehler (Bias) zu verringern und dem Zufallsfaktor Rechnung zu tragen (s. Kap. 6).

Die Unmöglichkeit vorherzusagen, was genau passiert, wenn ein Mensch eine Krankheit entwickelt oder eine Therapie erhält, be-

zeichnet man gelegentlich als Franklins Gesetz, nach dem amerikanischen Staatsmann Benjamin Franklin (18. Jahrhundert), der einmal die berühmten Worte gesprochen hat: «Nichts in dieser Welt ist sicher außer dem Tod und den Steuern.»³ Und doch ist Franklins Gesetz unserer Gesellschaft wohl kaum zur zweiten Natur geworden. Auf die Unvermeidlichkeit von Unsicherheit wird in unseren Schulen nicht stark genug hingewiesen, und auch nicht auf die grundlegenden Konzepte, wie man Erkenntnisse (Evidenz) gewinnt und interpretiert oder wie Angaben zu Wahrscheinlichkeiten und Risiken zu verstehen sind. Wie ein Kommentator es einmal formuliert hat:

*In der Schule wurde uns etwas über Chemikalien in Reagenzgläsern beigebracht, über Gleichungen, mit denen man Bewegung beschreibt, und vielleicht haben wir noch etwas über Photosynthese gelernt. Aber aller Wahrscheinlichkeit nach haben Sie nichts über Todesfälle, Risiken, Statistik und die Wissenschaft erfahren, die Sie entweder kuriert oder aber umbringt.*⁴

Und während die auf soliden wissenschaftlichen Erkenntnissen beruhende angewandte Medizin schon zahllose Leben gerettet hat, so würden Sie arg in Verlegenheit geraten, wenn Sie in einem naturwissenschaftlichen Museum ein einziges Exponat ausfindig machen sollten, das die wichtigsten Prinzipien der wissenschaftlichen Untersuchung erklärt.

Seien Sie sich nicht zu sicher!

«Aber im Laufe der Zeit finden wir, suchend, das Bess're. Sichere Wahrheit erkennt kein Mensch (...) Es ist alles durchwebt von Vermutung.»

(Xenophanes, 6. Jahrhundert v. Chr.)

«Ich bin mir immer sicher in Sachen, die man nicht wissen kann.»

(Charlie Brown, «Peanuts»)

«Unsere zahlreichen Irrtümer zeigen, dass der Umgang mit Kausalschlüssen ... eine Kunst bleibt. Auch wenn wir zu unserer Unterstützung analytische Verfahren, statistische Methoden und Konventionen sowie logische Kriterien entwickelt haben, so bleiben die Schlussfolgerungen, zu denen wir gelangen, letztlich doch eine Frage des Ermessens.»

(Susser M. Causal thinking in the health sciences. Oxford: Oxford University Press, 1983)

Die Begriffe Unsicherheit und Risiko spielen dabei wirklich eine große Rolle. Nehmen wir zum Beispiel die logische Unmöglichkeit, etwas «Negatives zu beweisen» – d.h. zu zeigen, dass etwas nicht existiert oder dass eine Therapie wirkungslos ist. Dies ist nicht bloß ein philosophischer Streit; es hat auch wichtige praktische Konsequenzen, wie die Erfahrungen mit einem Kombinationspräparat namens Bendectin (aktive Wirkstoffe Doxylamin und Pyridoxin bzw. Vitamin B₆) zeigen. Bendectin (auch unter den Namen Debendox und Diclectin im Handel) wurde Frauen häufig zur Linderung von Übelkeit in der Frühschwangerschaft verordnet. Allerdings kamen bald Behauptungen auf, dass Bendectin Geburtsdefekte verursacht. Diese Behauptung wurde bald in einer ganzen Lawine von Gerichtsprozessen untersucht. Unter dem Druck all dieser anhängigen Gerichtsverfahren nahmen die Hersteller Bendectin 1983 aus dem Handel. In mehreren nachfolgenden Übersichtsarbeiten der verfügbaren Evidenz konnte aber keine Unterstützung für einen Zusammenhang mit dem Auftreten von Geburtsdefekten erbracht werden – es war nicht möglich, den schlüssigen Beweis dafür zu erbringen, dass es nicht schadet, aber es gab auch keine Anhaltspunkte dafür, dass Bendectin schädlich war. Ironischerweise war Bendectin da schon vom Markt genommen worden, und die einzigen für die Behandlung der morgendlichen Übelkeit bei Schwangeren nun noch verfügbaren Medikamente sind solche, über deren Potenzial zur Verursachung von Geburtsdefekten wir noch viel weniger wissen.⁵

In der Regel kann Forschung solche Unsicherheiten bestenfalls verringern helfen. Therapien können sowohl schaden als auch nützen. Solide, gut durchgeführte Forschung kann die *Wahrscheinlichkeit* angeben, dass eine Therapie für ein gesundheitliches Problem Vor- oder Nachteile bewirkt, indem sie diese Therapie mit einer anderen Therapie oder überhaupt keiner Behandlung vergleicht. Da immer Unsicherheit herrschen wird, ist es sinnvoll, der Versuchung zu widerstehen, alle Dinge nur schwarz und weiß zu sehen. Denn das Denken in Wahrscheinlichkeiten macht uns stärker.⁶ Die Menschen müssen wissen, mit welcher Wahrscheinlichkeit eine Krankheit einen bestimmten Ausgang nimmt – sagen wir die Wahrscheinlichkeit, mit der bei einem Bluthochdruck-Patienten ein Schlaganfall eintritt – also die Faktoren, die Einfluss nehmen auf das Risiko eines

Schlaganfallereignisses sowie die Wahrscheinlichkeit, dass eine Therapie das Risiko eines Schlaganfallereignisses verändert. Wenn ausreichend verlässliche Informationen vorliegen, können Patienten und Ärzte gemeinsam zu beurteilen versuchen, in welchem Verhältnis zueinander die Vor- und Nachteile einer Therapie stehen. Anschließend können sie die Option wählen, die sich für den betreffenden Patienten – gemessen an seinen persönlichen Umständen und Präferenzen – am besten eignet.⁷

Dieses Buch hat sich zum Ziel gesetzt, die Kommunikation zu verbessern und das Vertrauen zu stärken; es ist nicht unser Ziel, das Vertrauen der Patienten in die Ärzteschaft zu untergraben. Dazu kann es aber nur kommen, wenn die Patienten sich selbst und ihren Ärzten helfen können, die verschiedenen Behandlungsoptionen kritisch zu bewerten.

In Kapitel 1 wird kurz erläutert, warum wir faire Therapietests brauchen und inwieweit manche neuen Therapien schädliche Wirkungen gezeigt haben, mit denen man nicht gerechnet hatte. In Kapitel 2 beschreiben wir Fälle, in denen sich erhoffte Therapieeffekte nicht eingestellt haben, und beleuchten die Tatsache, dass zahlreiche verbreitet eingesetzte Therapien nicht ausreichend evaluiert sind. Kapitel 3 veranschaulicht, warum eine intensivere Therapie nicht unbedingt besser ist. Kapitel 4 erklärt, warum das Screenen gesunder Menschen auf Krankheitsfrühzeichen sowohl nachteilig als auch hilfreich sein kann. In Kapitel 5 befassen wir uns mit den vielen Unsicherheiten, die nahezu alle Aspekte der Gesundheitsversorgung durchdringen, und zeigen auf, wie man mit diesen Unsicherheiten umgehen kann.

Kapitel 6, 7 und 8 enthalten verschiedene «fachliche Informationen», die aber in allgemein verständlicher Sprache präsentiert werden. In Kapitel 6 stellen wir die Grundlagen für faire Therapietests dar, wobei besonders darauf abgehoben wird, wie wichtig es ist, Gleiches mit Gleichem zu vergleichen. Kapitel 7 beleuchtet die Frage, warum die Berücksichtigung des Zufallsfaktors unerlässlich ist, und Kapitel 8 erläutert, warum es so wichtig ist, alle relevanten zuverlässigen Erkenntnisse systematisch auszuwerten.

Kapitel 9 legt dar, warum die Systeme zur Reglementierung von Forschung über Therapieeffekte (in Gestalt von Ethikkommissio-

nen und anderen Gremien) für die Durchführung guter Forschung hinderlich sein können, und erläutert, warum diese Art von Reglementierung den Patienteninteressen unter Umständen zuwider laufen kann. In Kapitel 10 werden die wichtigsten Unterschiede zwischen guter, schlechter und überflüssiger Forschung über Therapieeffekte herausgearbeitet; es legt dar, inwieweit Forschung häufig durch wirtschaftliche und akademische Prioritäten verzerrt wird und es ihr deshalb nicht gelingt, die Fragen zu thematisieren, die für das Wohlergehen der Patienten wirklich wichtig wären.

Kapitel 11 zeigt auf, was Patienten und die Öffentlichkeit tun können, um sicherzustellen, dass Therapien besser getestet werden. In Kapitel 12 wenden wir uns der Frage zu, wie robuste Evidenz aus der Therapieforschung wirklich zu einer besseren Gesundheitsversorgung des einzelnen Patienten beitragen kann. Und in Kapitel 13 stellen wir unseren Entwurf für eine bessere Zukunft vor; den Abschluss dieses Kapitels bildet ein Aktionsplan.

In jedem Kapitel wird auf eine Auswahl an wichtigen Quellen Bezug genommen. Am Ende des Buchs (S. 247) findet sich ein eigener Abschnitt mit zusätzlichen Quellen. Für alle Leser, die sich mit dem einen oder anderen Aspekt noch näher befassen möchten, bietet die *James Lind Library* unter www.jameslindlibrary.org einen guten Einstieg. Die kostenlose elektronische Version der zweiten Auflage von *Testing Treatments* ist auf einer neuen Webseite (*Testing Treatments Interactive*) unter www.testingtreatments.org zu finden; diese Webseite soll im Laufe der kommenden Jahre um andere Materialien ergänzt werden.

Die Autoren dieses Buchs haben sich dem Prinzip eines gerechten Zugangs zu einer auf die Bedürfnisse der Menschen eingehenden wirkungsvollen Gesundheitsversorgung verpflichtet. Diese soziale Verantwortung hängt wiederum von verlässlichen und zugänglichen Informationen über die Wirkungen von Tests und Therapien ab, die aus solider Forschung stammen. Da die Gesundheitsressourcen überall begrenzt sind, müssen sich Therapien auf robuste Evidenz stützen können und effizient und gerecht eingesetzt werden, wenn die ganze Bevölkerung eine Chance haben soll, vom medizinischen Fortschritt zu profitieren. Es ist unverantwortlich, wertvolle Ressourcen auf Therapien zu verschwenden, die nur wenig nützen,

oder ohne guten Grund auf Gelegenheiten zur Überprüfung von Therapien zu verzichten, über die wir bislang nur wenig wissen. Faire Therapietests sind deshalb von entscheidender Bedeutung, weil es darum geht, allen die gleichen Behandlungsoptionen zu ermöglichen.

Wir hoffen, dass Sie, liebe Leserinnen und Leser, nach der Lektüre von *Testing Treatments* die Begeisterung für dieses Thema mit uns teilen und sich daran machen werden, unbequeme Fragen über Therapien zu stellen, Lücken im medizinischen Wissen aufzudecken und sich an der Forschung zu beteiligen, um so Antworten zu finden, von denen Sie selbst und auch alle anderen profitieren werden.

1

Neu – aber auch besser?

Warum wir faire Tests von medizinischen Therapien brauchen

Wenn die Wirksamkeit medizinischer Behandlungen nicht in fairen – unverzerrten, unvoreingenommenen – Tests geprüft wird, kann es passieren, dass Ärzte ihren Patienten nutzlose oder sogar schädliche Therapien verordnen, in der Annahme, sie würden helfen. Umgekehrt kann es vorkommen, dass nützliche Therapien als nutzlos verworfen werden. Alle Therapien sollten einer solchen fairen Überprüfung unterzogen werden – egal, woher sie stammen, und ganz gleich, ob sie der Schulmedizin oder eher der Komplementär-/Alternativmedizin zuzurechnen sind. Ungeprüfte Theorien über Behandlungseffekte, wie überzeugend sie auch klingen mögen, sind schlicht nicht ausreichend. Denn es ist schon vorgekommen, dass die Wirksamkeit von Behandlungen rein theoretisch vorhergesagt, diese Annahme durch «faire Tests» aber widerlegt wurde. In wieder anderen Fällen gab es zuversichtliche Vorhersagen über die Unwirksamkeit einer Therapie, bis Therapiestudien den Gegenbeweis erbrachten.

Naturgemäß neigen wir zu der Auffassung, dass «neu» – wie in der Waschmittelwerbung – immer mit «besser» gleichgesetzt wird. Doch wenn man neue Therapien fair bewertet, dann ist die Wahrscheinlichkeit, dass sie schlechter abschneiden als bereits bestehende Behandlungen, ebenso hoch wie die Wahrscheinlichkeit, dass sie sich als besser erweisen. Ebenso natürlich ist unsere Neigung, etwas für sicher und wirksam zu halten, nur weil wir es schon lange kennen. Doch in der Medizin gibt es unzählige Beispiele dafür, dass Therapien aus Gewohnheit oder fester Überzeugung angewendet

werden und nicht, weil ihre Wirksamkeit durch verlässliche Hinweise aus klinischen Studien (Evidenz) belegt ist: Therapien, die oftmals gar nicht helfen und manchmal sogar erheblichen Schaden anrichten.

Dass wir «faire Tests» für die Wirksamkeit medizinischer Behandlungen brauchen, ist nichts Neues: Schon im 18. Jahrhundert wendete James Lind (**Abb. 1**) einen solchen fairen Test an. Er wollte wie in einem Experiment sechs Heilmittel vergleichen, die seinerzeit zur Behandlung von Skorbut angewendet wurden – einer Krankheit, der auf langen Seereisen unzählige Seeleute zum Opfer fielen. Lind erbrachte den Beweis, dass Orangen und Zitronen, von denen wir heute wissen, dass sie Vitamin C enthalten, ein sehr wirksames Heilmittel waren.

Als James Lind 1747 als Schiffsarzt an Bord der *HMS Salisbury* diente, brachte er zwölf seiner Skorbut-Patienten mit vergleichbarem Krankheitsstadium im selben Teil des Schiffs unter und sorgte dafür, dass sie dieselbe Nahrung erhielten. Dies war entscheidend, denn er schaffte damit «gleiche Ausgangsbedingungen» (s. Kap. 6

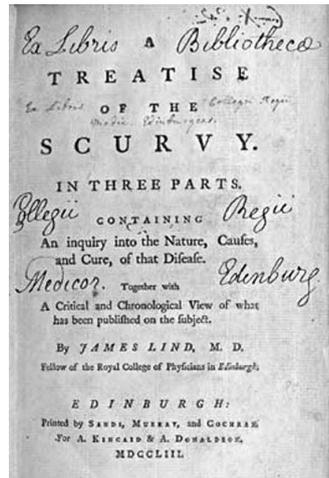
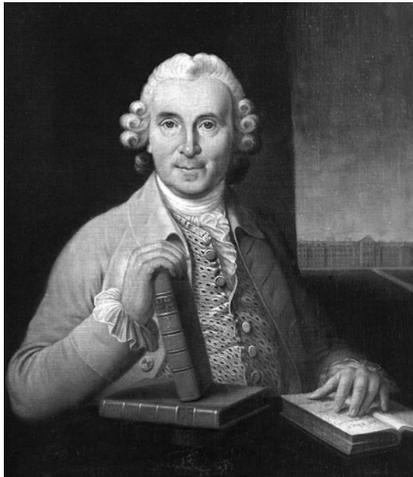


Abbildung 1: Der schottische Schiffsarzt James Lind (1716–1794) mit den von ihm verfassten Büchern und der Titelseite seiner berühmtesten Schrift. In dieser Schrift hatte er eine 1747 durchgeführte kontrollierte Studie dokumentiert, mit der er nachweisen konnte, dass Orangen und Zitronen zur Behandlung von Skorbut wirksamer waren als fünf andere damals übliche Behandlungen (s. www.jameslindlibrary.org).

und Kasten in Kap. 3, S. 61). Je zwei Seeleute teilte Lind darauf jeweils einer der damals bei Skorbut üblichen Behandlungen (Apfelwein, Schwefelsäure, Essig, Meerwasser, Muskatnuss) oder der «Behandlung» mit zwei Orangen und einer Zitrone zu. Die Südfrüchte siegten mühelos. Die Admiralität ordnete später auf allen Schiffen die Ausgabe von Zitronensaft an – mit dem Ergebnis, dass die tödliche Krankheit in der britischen Marine gegen Ende des 18. Jahrhunderts nur noch Geschichte war.

Die damalige britische «Ärztelkammer», das *Royal College of Physicians*, favorisierte von den Behandlungen, die Lind verglich, die Schwefelsäure, während die Admiralität dem Essig den Vorzug gab. Linds faire Untersuchung erbrachte den Nachweis, dass beide Autoritäten damit falsch lagen. Erstaunlicherweise kommt es gar nicht so selten vor, dass einflussreiche Experten sich irren. Sich allzu sehr auf Meinungen, Gewohnheiten oder Präzedenzfälle anstatt auf die Ergebnisse fairer Tests zu verlassen, stellt auch heute noch ein schwerwiegendes Problem im Gesundheitswesen dar (s. u. und Kap. 2).

Anekdoten sind Anekdoten

«Unser Gehirn scheint auf Anekdoten programmiert zu sein, und am leichtesten lernen wir ja auch mithilfe spannender Geschichten. Doch macht es mich immer wieder fassungslos, wie viele Menschen, darunter auch nicht wenige meiner Freunde, die darin liegenden Tücken nicht erkennen. In den Naturwissenschaften weiß man, dass Anekdoten und persönliche Erfahrungen uns auf verhängnisvolle Weise in die Irre führen können. Was man hier braucht, sind überprüfbare und wiederholbare Ergebnisse. In der Medizin dagegen kommt man damit nicht allzu weit. Zwischen den einzelnen Menschen bestehen zu viele Unterschiede, als dass man sich seiner Sache völlig sicher sein könnte, wenn es um den einzelnen Patienten geht. Deshalb ist der Spielraum für Mutmaßungen häufig recht groß. Aber es gibt Grenzen, die nicht überschritten werden dürfen, denn sonst wären die Wissenschaften verraten und verkauft: Es würde gefpuscht werden, und Fakten und Meinungen würden sich vermischen, bis wir kaum mehr das eine vom anderen unterscheiden könnten.»

Ross N. Foreword. In: Ernst E, ed. *Healing, hype, or harm? A critical analysis of complementary or alternative medicine*. Exeter: Societas, 2008:vi-vii.

Heutzutage werden Unsicherheiten hinsichtlich der Wirkungen von Therapien oftmals dann deutlich, wenn Ärzte und andere Kliniker sich über das jeweils beste Vorgehen bei einer Krankheit uneinig sind (s. Kap. 5). Bei der Ausräumung dieser Unsicherheiten kommt

sowohl den Patienten und der Öffentlichkeit, aber auch den Ärzten selbst eine wichtige Rolle zu. Patienten sowie Gesundheitsexperten haben ein vitales Interesse daran, dass in der Erforschung medizinischer Behandlungen strikte Regeln eingehalten werden. Zum einen muss Ärzten die Gewissheit gegeben werden, dass ihre Behandlungsempfehlungen auf solider Evidenz beruhen, zum anderen müssen die Patienten darauf bestehen, dass genau dies der Fall ist. Nur durch den Aufbau dieser kritischen Partnerschaft kann die Öffentlichkeit Vertrauen in das haben, was die moderne Medizin zu bieten hat (s. Kap. 11, 12 und 13).

Unerwartete negative Wirkungen

Thalidomid

Ein besonders abschreckendes Beispiel für ein neues Medikament, das mehr Schaden als Nutzen¹ gebracht hat, ist Thalidomid, das in Deutschland unter dem Namen «Contergan» verkauft wurde. Dieses Schlafmittel wurde gegen Ende der 1950er-Jahre als eine angeblich sichere Alternative zu den damals regelmäßig verordneten Barbituraten auf den Markt gebracht; denn anders als bei den Barbituraten führte eine Überdosierung von Thalidomid nicht zum Koma. Thalidomid wurde speziell für Schwangere empfohlen, bei denen es auch zur Linderung der morgendlichen Übelkeit eingesetzt wurde.

Zu Beginn der 1960er-Jahre beobachteten Geburtshelfer bei Neugeborenen dann aber eine starke Zunahme von Fällen schwerer Arm- und Beifehlbildungen. Diese bis dahin seltene Form von Fehlbildungen führte zu derart extrem verkürzten Gliedmaßen, dass Hände und Füße direkt aus dem Körper zu wachsen schienen. Deutsche und australische Ärzte führten diese Fehlbildungen bei Säuglingen darauf zurück, dass die Mütter in der Frühschwangerschaft Thalidomid eingenommen hatten.²

Ende 1961 nahm der Hersteller Thalidomid vom Markt. Erst viele Jahre später – nach öffentlichen Kampagnen und gerichtlichen Auseinandersetzungen – begann man den Opfern eine Entschädi-

gung zu zahlen. Der durch diese verheerenden Fehlbildungen ange- richtete Schaden war immens hoch. In den rund 46 Ländern, in de- nen Ärzte Thalidomid verordnet hatten (in einigen Ländern war es sogar frei verkäuflich), waren Tausende von Säuglingen betroffen. Ärzte, pharmazeutische Industrie und Patienten waren angesichts der Thalidomid-Tragödie gleichermaßen fassungslos. In der Folge kam es weltweit zu einer Neugestaltung der Prozesse, die die Ent- wicklung und Zulassung von Arzneimitteln steuern.³

Eine tragische Epidemie von Blindheit bei Babys

«Unmittelbar nach dem Zweiten Weltkrieg wurden zahlreiche neue Therapien eingeführt, welche die Überlebenschancen von Frühgeborenen verbessern sol- len. Im Laufe der nächsten paar Jahre zeigte sich jedoch in schmerzlicher Weise, dass verschiedene Änderungen bei den pflegerischen Maßnahmen völlig uner- wartete schädliche Wirkungen zeitigten. Die bedeutendste dieser tragischen klini- schen Erfahrungen war die in den Jahren 1942 bis 1954 auftretende ‹Epidemie› von Blindheit, der sogenannten Frühgeborenen-Retinopathie oder retrolentalen Fibroplasie. Wie sich herausstellte, stand diese Erkrankung mit der Art der Sauer- stoffzufuhr im Rahmen der Behandlung von unvollständig entwickelten Neu- geborenen in Zusammenhang. Der zwölf Jahre währende Kampf gegen den Ausbruch der Krankheit machte auf ernüchternde Weise deutlich, wie wichtig es ist, alle medizinischen Neuerungen einer gründlichen Bewertung zu unterziehen, bevor sie zur allgemeinen Anwendung zugelassen werden.»

Silverman WA. *Human experimentation: a guided step into the unknown*.
Oxford: Oxford University Press, 1985: vii-viii.

Vioxx

Doch obwohl die Arzneimittelprüfrichtlinien erheblich verschärft wurden, können selbst die besten Prüfmethode keine absolute Sicherheit garantieren. Nichtsteroidale Antiphlogistika (NSAID) sind ein gutes Beispiel dafür, warum wir so etwas wie Arzneimittel- überwachung brauchen. Gewöhnlich werden NSAID außer zur Schmerz- lindering und Entzündungshemmung bei verschiedenen Erkrankungen (z. B. Arthritis) auch zur Fiebersenkung eingesetzt. Zu den ‹traditionellen› NSAID gehören viele frei verkäufliche Me- dikamente wie Aspirin und Ibuprofen. Zu ihren Nebenwirkungen zählen bekanntlich Magen- und Darmreizungen, die zu Dyspepsie (‹Verdauungsstörungen›) und gelegentlich auch zu Blutungen und sogar gastrischen Ulzera (Magengeschwüren) führen können.

Die Arzneimittelhersteller hatten folglich gute Gründe, nach NSAID zu suchen, die keine dieser Komplikationen nach sich ziehen.

Rofecoxib (besser bekannt unter seinem Handelsnamen Vioxx, auch als Ceoxx bzw. Ceeoxx vermarktet) wurde 1999 als eine vermeintlich sicherere Alternative zu den älteren Substanzen eingeführt und schon bald darauf häufig verordnet. Kaum fünf Jahre später nahm der Hersteller Vioxx wegen eines erhöhten Risikos für kardiovaskuläre Komplikationen wie Herzinfarkt und Schlaganfall vom Markt. Was war passiert?

1999 hatte Vioxx von der US-amerikanischen Zulassungsbehörde (*Food and Drug Administration*, FDA) die Zulassung zur «Linderung der Zeichen und Symptome von Arthrose, zur Behandlung akuter Schmerzen bei Erwachsenen und zur Behandlung von Menstruationssymptomen» (d.h. Regelschmerzen) erhalten. Später wurde die Zulassung auf die Linderung der Zeichen und Symptome rheumatoider Arthritiden bei Erwachsenen und Kindern erweitert. Im Laufe der Entwicklung von Vioxx waren die Wissenschaftler des Pharmaunternehmens auf die potenziell schädlichen Wirkungen aufmerksam geworden, die das Medikament auf das Blutgerinnungssystem des Körpers hat und die das Risiko der Blutgerinnelbildung erhöhten. Die hauptsächlich kleinen Studien, die im Rahmen des Zulassungsverfahrens bei der FDA eingereicht wurden, hatten sich jedoch auf den Nachweis der entzündungshemmenden Effekte von Vioxx konzentriert und waren nicht für die Untersuchung der möglichen Komplikationen ausgelegt gewesen.⁴

Bereits vor der FDA-Zulassung hatte das Unternehmen eine große Studie begonnen, in der bei Patienten mit rheumatoider Arthritis hauptsächlich die gastrointestinalen Nebenwirkungen mit denen von Naproxen, einem anderen NSAID, verglichen werden sollten. Auch diese Studie war nicht speziell für den Nachweis kardiovaskulärer Komplikationen ausgelegt. Darüber hinaus kamen später noch Fragen nach Interessenkonflikten bei den Mitgliedern des unabhängigen Datenüberwachungskomitees (*Data and Safety Monitoring Board*, DSMB) auf (ein Gremium, das mit der Überwachung der Studienergebnisse betraut ist, um beim Vorliegen ausreichender Gründe eine Studie ggf. vorzeitig abzubrechen).

Dennoch ließen die Ergebnisse – die zeigten, dass Vioxx weniger Episoden von Magenulzera und gastrointestinalen Blutungen verursachte als Naproxen – eine höhere Anzahl von Herzinfarkten in der Vioxx-Gruppe erkennen. Immerhin wurde der in einer bekannten medizinischen Fachzeitschrift veröffentlichte Studienbericht heftig kritisiert. Zu den Vorwürfen gehörte, dass die Auswertung und Darstellung der Ergebnisse so gewählt worden war, dass sie die Schwere der kardiovaskulären Risiken herunterspielten. Der Herausgeber der Zeitschrift klagte später darüber, dass die Forscher kritische Daten zu diesen Nebenwirkungen zurückgehalten hätten. Allerdings waren die Ergebnisse, die der FDA im Jahre 2000 vorgelegt und von ihrem Arthritis-Beirat 2001 diskutiert worden waren, schließlich der Anlass dafür, dass die FDA die Sicherheitsinformation auf dem Vioxx-Etikett 2002 um den Hinweis auf ein erhöhtes Herzinfarkt- und Schlaganfallrisiko ergänzen ließ.

Der Arzneimittelhersteller setzte seine Untersuchungen zu weiteren Anwendungsmöglichkeiten von Vioxx fort und begann im Jahr 2000 eine Studie, in der geprüft werden sollte, ob das Medikament kolorektale Polypen (kleine gutartige Tumoren im unteren Darm, aus denen sich ein kolorektales Karzinom entwickeln kann) verhindern konnte. Diese Studie wurde vorzeitig abgebrochen, als Zwischenergebnisse erkennen ließen, dass das Medikament mit einem erhöhten Risiko für kardiovaskuläre Komplikationen vergesellschaftet war. In der Folge nahm der Hersteller Vioxx 2004 aus dem Handel. Im veröffentlichten Studienbericht behaupteten die Studienautoren, die weder beim Hersteller angestellt waren noch Beratungshonorare vom Unternehmen erhalten hatten, dass die kardiovaskulären Komplikationen erst nach 18-monatiger Einnahme von Vioxx auftraten. Diese Behauptung beruhte auf einer fehlerhaften Datenauswertung und wurde später von der Zeitschrift, in welcher der Studienbericht erschienen war, offiziell richtiggestellt.⁴ Trotz der zahlreichen juristischen Schritte, die Patienten im Anschluss unternahmen, behauptet der Hersteller weiterhin, dass das Unternehmen jederzeit – angefangen bei den Zulassungsstudien bis hin zur Sicherheitsüberwachung nach der Markteinführung von Vioxx – verantwortlich gehandelt habe. Das Unternehmen bekräftigte zudem seine Überzeugung, dass die Belege zeigen würden, dass

vorbestehende kardiovaskuläre Risikofaktoren und nicht Vioxx verantwortlich zu machen seien.⁵

Der Vioxx-Skandal zeigt, dass auch ein halbes Jahrhundert nach Thalidomid noch immer viel zu tun bleibt, um die faire Überprüfung der Wirksamkeit von Behandlungen, die Transparenz des Prozesses und die Robustheit der Belege gewährleisten zu können. Wie einige Kommentatoren es ausgedrückt haben:

Unser System beruht darauf, dass die Interessen der Patienten an erster Stelle stehen. Die Zusammenarbeit zwischen in der Forschung tätigen Ärzten, praktizierenden Ärzten, der Industrie und den Fachzeitschriften ist für die Erweiterung unserer Kenntnisse und die Verbesserung der Patientenversorgung essenziell. Vertrauen ist ein notwendiger Bestandteil dieser Partnerschaft. Allerdings machen die jüngsten Ereignisse die Einrichtung geeigneter Systeme zum Schutz der Patienteninteressen erforderlich. Ein erneutes Bekenntnis aller Beteiligten und die Einrichtung besagter Systeme stellen die einzige Möglichkeit dar, wie man dieser unglücklichen Geschichte noch etwas Positives abgewinnen kann.⁴

Avandia

2010 schaffte es ein weiteres Medikament wegen unerwünschter Nebenwirkungen auf das Herz-Kreislauf-System in die Schlagzeilen: und zwar Rosiglitazon, besser bekannt unter seinem Handelsnamen Avandia. Zehn Jahre zuvor war Avandia von europäischen und US-amerikanischen Zulassungsbehörden als ein neuer Ansatz in der Behandlung des Typ-2-Diabetes zugelassen worden. Diese Form des Diabetes tritt auf, wenn der Körper nicht genug Insulin produziert oder wenn die Körperzellen nicht auf Insulin ansprechen. Sie kommt sehr viel häufiger vor als der Typ-1-Diabetes, bei dem der Körper überhaupt kein Insulin herstellen kann. Der Typ-2-Diabetes, der oftmals mit Adipositas (Fettleibigkeit) einhergeht, ist durch Ernährungsumstellung, körperliche Bewegung und orale Medikamenteneinnahme (statt der Injektion von Insulin) normalerweise gut behandelbar. Zu den Langzeitkomplikationen des Typ-2-Diabetes gehört ein erhöhtes Herzinfarkt- und Schlaganfallrisiko. Das

Hauptziel der Behandlung ist es daher, das Risiko für diese Komplikationen zu verringern.

Beworben wurde Avandia damit, dass es die Wirkung des körpereigenen Insulins auf eine neue Weise verbessere und dass es den älteren Medikamenten bei der Kontrolle des Blutzuckerspiegels überlegen sei. Dabei stand der Blutzucker im Mittelpunkt und nicht die schwerwiegenden, Leiden verursachenden Komplikationen, welche die Patienten letztendlich umbringen.

Als Avandia zugelassen wurde, gab es wenige aussagekräftige Belege für seine Wirksamkeit und keine Nachweise für seine Auswirkungen auf das Herzinfarkt- und Schlaganfallrisiko. Die Zulassungsbehörden forderten den Hersteller deshalb zur Durchführung zusätzlicher Studien auf, denn inzwischen wurde Avandia weltweit häufig und begeistert verschrieben. Dann tauchten aber allmählich erste Berichte über unerwünschte kardiovaskuläre Effekte auf, deren Anzahl stetig zunahm. 2004 zeigte sich die Weltgesundheitsorganisation (WHO) derart beunruhigt, dass sie den Hersteller aufforderte, die Evidenz bezüglich dieser Komplikationen noch einmal zu überprüfen. Der Hersteller kam dieser Aufforderung nach und bestätigte das erhöhte Risiko.⁶

Allerdings dauerte es weitere sechs Jahre, bevor die Zulassungsbehörden die Evidenz genauer unter die Lupe nahmen und handelten. Im September 2010 verkündete die FDA, dass die Anwendung von Avandia rigoros auf Patienten beschränkt werden müsse, deren Typ-2-Diabetes mit anderen Medikamenten nicht beherrschbar sei, und im selben Monat empfahl die europäische Arzneimittelzulassungsbehörde (*European Medicines Agency*, EMA), Avandia im Laufe der nächsten zwei Monate vom Markt zu nehmen. Beide Behörden gaben als Grund für ihre Entscheidung das erhöhte Herzinfarkt- und Schlaganfallrisiko an. Mittlerweile wurde in unabhängigen Untersuchungen eine lange Liste verpasster Handlungsmöglichkeiten aufgezeigt, und aus den Reihen der Ärzteschaft wurde auf die grundsätzliche Notwendigkeit hingewiesen, dass Zulassungsbehörden und Ärzte «bessere Nachweise verlangen müssten, bevor wir uns auf die Massenmedikation einer großen Gruppe von Patienten einlassen, die sich zur Beratung und Behandlung an uns wenden».⁷

Mechanische Herzklappen

Medikamente sind aber nicht die einzigen Therapieformen, die unerwartete negative Auswirkungen haben können: Auch nicht-medikamentöse Behandlungen können schwerwiegende Risiken bergen. Bei Patienten mit einer schweren Herzklappenerkrankung zählen mechanische Herzklappen heutzutage zu den Standardtherapien. Das Design der Herzklappen wurde im Laufe der Jahre mehrfach verbessert. Wie die Erfahrungen mit einem bestimmten Typ von mechanischer Klappe allerdings gezeigt haben, hatte einer dieser Verbesserungsversuche desaströse Folgen. In den frühen 1970er-Jahren kam die sogenannte Björk-Shiley-Herzklappe in den Handel. Die frühen Modelle waren jedoch anfällig für Thrombosen (Blutgerinnselbildung), was ihre Funktion beeinträchtigte. Zur Behebung dieses Nachteils wurde das Design Ende der 1970er-Jahre modifiziert, um die Gefahr der Gerinnselbildung zu verringern.

Die neue Bauform umfasste eine von zwei Metallbügeln (dem Gerüst) gehaltene Scheibe, von denen weltweit viele tausend Exemplare implantiert wurden. Leider hatte die Klappenstruktur einen gravierenden Nachteil: Einer der Bügel neigte dazu abzureißen (sogenannte Bügelfraktur), mit der Folge einer katastrophalen und oftmals tödlichen Klappenfehlfunktion.

Wie es der Zufall wollte, waren die Bügelbrüche bereits während der Prüfungen vor dem Inverkehrbringen der Prothesen als Problem erkannt worden. Dies hatte man jedoch der fehlerhaften Verschweißung der Bügel zugeschrieben, deren Ursache nicht vollständig untersucht wurde. Nichtsdestotrotz wurde diese Erklärung von der FDA ebenso akzeptiert wie die Versicherung des Herstellers, dass das Bügelfrakturrisiko durch die Senkung des Risikos für Klappenthrombosen mehr als ausgeglichen wurde. Als die Beweise für das desaströse Klappenversagen nicht mehr ignoriert werden konnten, handelte die FDA schließlich und erzwang 1986 den Rückruf der Klappen, aber erst nachdem Hunderte von Patienten unnötigerweise verstorben waren. Auch wenn die Systeme zur Produktreglementierung inzwischen verbessert worden sind und eine bessere Beobachtung der Patienten nach der Markteinführung sowie

flächendeckende Patientenregister umfassen, besteht bei der Einführung neuer Medizinprodukte nach wie vor ein dringender Bedarf für größere Transparenz.⁸

Zu schön, um wahr zu sein

Herceptin

Kommerzielle Unternehmen sind nicht die einzigen, welche die Vorzüge neuer Therapien lauthals verkünden, während sie ihre Nachteile herunterspielen. Auch der Medizinbetrieb kann so etwas aufbauschen, ebenso wie die enthusiastische Berichterstattung in den Medien – und beide tragen dazu bei, die Vorteile herauszustreichen und gleichzeitig ihre potenziellen Nachteile zu missachten. Bei diesen Nachteilen kann es sich nicht nur um schädliche Nebenwirkungen handeln, sondern auch um diagnostische Probleme, wie die Ereignisse im Zusammenhang mit dem Brustkrebsmedikament Trastuzumab, besser bekannt unter seinem Handelsnamen Herceptin, gezeigt haben (s. auch Kap. 3).

Anfang 2006 veranlassten lautstarke Forderungen von Patienten- und Ärzteverbänden, angeheizt von der pharmazeutischen Industrie und den Massenmedien, den staatlichen britischen Gesundheitsdienst (*National Health Service*, NHS) dazu, Patientinnen mit Brustkrebs im Frühstadium Herceptin zugänglich zu machen. Die Patientinnen mit ihren hartnäckigen Forderungen (und ihrer sogenannten *Pester Power*) triumphierten, und Herceptin wurde als Jahrhundertmedikament vorgestellt (s. Kap. 11).

Damals war Herceptin aber nur für die Behandlung von metastasiertem (gestreutem) Brustkrebs zugelassen und für Brustkrebs im Frühstadium nicht ausreichend getestet worden. Tatsächlich hatte der Hersteller die Zulassung für die Behandlung von Frühstadien der Erkrankung bei einer sehr kleinen Untergruppe von Frauen – die für ein als HER2 bezeichnetes Protein positiv getestet worden waren – gerade erst beantragt. Dieses genetische Profil findet sich aber nur bei einer von fünf Frauen. Über die Schwierigkeiten und Kosten, die mit der genauen Klärung der Frage, ob eine Patientin HER2-positiv

ist, verbunden sind, und über die Möglichkeiten einer falschen Diagnose – und damit auch einer «falsch positiven» Behandlung wurde in der begeisterten, aber unkritischen Presse nur selten berichtet. Und es wurde auch nicht darauf hingewiesen, dass mindestens vier von fünf Brustkrebspatientinnen *nicht* HER2-positiv sind.^{9, 10, 11, 12}

Das britische *National Institute for Health and Clinical Excellence* (NICE) – das mit dem Deutschen Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) vergleichbar und wie dieses mit der unparteiischen Prüfung medizinischer Evidenz und der Erarbeitung von Empfehlungen beauftragt ist – konnte Herceptin erst im späteren Verlauf des Jahres 2006 als Behandlungsoption für Frauen mit HER2-positivem Mammakarzinom im Frühstadium empfehlen. Aber selbst zu diesem Zeitpunkt wurde eine deutliche Warnung ausgesprochen. Weil sich die Hinweise verdichteten, dass Herceptin unerwünschte Wirkungen auf die Herzfunktion haben könnte, riet NICE allen Ärzten, vor der Verordnung des Medikaments die Herzfunktion der Patientinnen zu prüfen. Frauen mit Herzproblemen (von Angina pectoris bis hin zu Herzrhythmusstörungen) sollte das Medikament nicht angeboten werden. NICE war zu der Auffassung gelangt, dass angesichts der vorliegenden Kurzzeitdaten über die Nebenwirkungen, von denen einige als schwer eingestuft wurden, Vorsicht geboten sei. Bis Langzeitwirkungen – nützliche wie schädliche – erkennbar werden, kann geraume Zeit vergehen.¹³

Ähnlicher Druck wurde, was die Anwendung von Herceptin betraf, auch in anderen Ländern ausgeübt. In Neuseeland beispielsweise forderten Patientenorganisationen, Presse und Medien, Arzneimittelhersteller wie auch Politiker einmütig, dass Brustkrebspatientinnen Herceptin erhalten sollten. Zu einer vergleichbaren Einschätzung der Evidenz über die Anwendung von Herceptin beim Mammakarzinom im Frühstadium kam auch die neuseeländische *Pharmaceutical Management Agency* (PHARMAC), die eine ähnliche Funktion erfüllt wie das IQWiG in Deutschland und NICE in Großbritannien. Im Juni 2007 entschied die PHARMAC aufgrund ihrer Prüfungsergebnisse, dass bei Patientinnen mit Brustkrebs im Frühstadium die neunwöchige Gabe von Herceptin angebracht sei. Zudem sollten Herceptin und andere Antikrebsmedikamente nicht

nacheinander, sondern gleichzeitig verabreicht werden. Dieser Neun-Wochen-Zyklus gehörte zu den drei Therapieschemata, die damals auf der ganzen Welt ausprobiert wurden. Zudem beschloss PHARMAC die Mitfinanzierung einer internationalen Studie, mit der die ideale Dauer der Herceptin-Therapie bestimmt werden sollte. Im November 2008 ignorierte die neu gewählte Regierung PHARMACs evidenzbasierte Entscheidung und sagte die Finanzierung eines 12-monatigen Behandlungszyklus mit Herceptin zu.¹⁴

Vom Strudel mitgerissen

Im Jahr 2006 sah sich eine Patientin in Großbritannien, die selber Ärztin war, unversehens von der Herceptin-Welle erfasst. Im Jahr zuvor war bei ihr die Diagnose eines HER2-positiven Mammakarzinoms gestellt worden.

«Vor meiner Diagnose wusste ich nur wenig über die moderne Behandlung von Brustkrebs und zog, wie viele Patientinnen, deshalb Online-Quellen zu Rate. Die britische Brustkrebsstiftung *Breast Cancer Care* führte auf ihrer Internetseite eine Kampagne mit dem Ziel durch, allen Frauen mit HER2-positivem Brustkrebs Herceptin zugänglich zu machen; ich unterzeichnete die Petition, weil ich angesichts der auf der Internetseite und in den Medien veröffentlichten Daten einfach nicht verstehen konnte, warum man Frauen ein so wirksames Medikament vorenthalten sollte, die es bei einem Rezidiv [Krankheitsrückfall] ohnehin erhalten würden. ... Mir war, als hätte ich kaum eine Chance, meine Krebserkrankung zu überleben, wenn ich dieses Medikament nicht bekäme! Zudem wurde ich auch noch von der Zeitschrift *The Sun* kontaktiert, welche die Herceptin-Kampagne unterstützte und sich an meiner Geschichte – aus der Sicht einer Ärztin und eines «Krebsopfers» – interessiert zeigte.

Nach Beendigung der Chemotherapie sprach ich mit meinem Onkologen über eine Behandlung mit Herceptin. Er äußerte Bedenken, denn bei der Behandlung über einen längeren Zeitraum waren in Studien kardiale Effekte [Auswirkungen auf das Herz] beobachtet worden, denen man auf besagter Internetseite und seitens der Medien aber nur sehr wenig Beachtung geschenkt hatte, zumal das Medikament ja ansonsten gesunden Frauen verabreicht würde. Darüber hinaus übersetzte sich der «50%-Vorteil», der vielfach zitiert worden war und sich in meinem Kopf festgesetzt hatte, für mich persönlich tatsächlich in einen 4- bis 5-prozentigen Vorteil, dem ein gleich hohes kardiales Risiko gegenüberstand! Also entschied ich mich gegen die Einnahme des Medikaments und werde zu dieser Entscheidung wohl auch dann noch stehen, wenn der Tumor wieder auftreten sollte.

Diese Geschichte zeigt, dass sogar eine medizinisch ausgebildete und normalerweise rational denkende Frau anfällig wird, wenn bei ihr eine potenziell lebensbedrohliche Krankheit diagnostiziert wird. ... Ein Großteil der Informationen zur Anwendung von Herceptin bei Brustkrebs im Frühstadium entstammte dem Riesenrummel, der von den Medien und der Industrie künstlich erzeugt und durch Einzelfälle wie den meinen angeheizt worden war.»

Cooper J. Herceptin (rapid response). *BMJ*. Eingestellt am 29. November 2006 unter www.bmj.org.

Im Zusammenhang mit Herceptin bleiben viele Fragen offen, z. B. wann und wie lange das Medikament verordnet werden sollte, ob die Vorteile durch die bei einigen Frauen auftretenden Langzeitschäden aufgewogen werden und ob das Medikament das Wiederauftreten der Krebserkrankung verzögert bzw. verhindert. Und es wurden weitere Bedenken geäußert, u. a. auch dass Herceptin, wenn es in Kombination mit anderen Brustkrebsmitteln wie Anthrazyklinen und Cyclophosphamid verabreicht wird, das Risiko der Patientinnen für unerwünschte Wirkungen auf das Herz (von rund 4 auf etwa 27 Patientinnen pro 100) erhöhen kann.¹⁵

- Die Überprüfung neuer Therapien ist notwendig, weil die Wahrscheinlichkeit, dass neue Therapien schlechter sind als bestehende Behandlungen, ebenso hoch ist wie die Wahrscheinlichkeit, dass sie besser sind.
- Verzerrte (unfaire) Tests von medizinischen Behandlungen können aufseiten des Patienten Leiden und Tod verursachen.
- Die Tatsache, dass eine Therapie zugelassen ist, ist keine Gewähr dafür, dass sie auch sicher ist.
- Oftmals dauert es gewisse Zeit, bis die Nebenwirkungen einer Therapie erkennbar werden.
- Die nützlichen Effekte einer Therapie werden häufig übertrieben, während ihre schädliche Wirkungen heruntergespielt werden.

2

Erhoffte, aber nicht eingetretene Wirkungen

Manche Behandlungen werden bereits längere Zeit angewandt, bevor erkannt wird, dass sie mehr schaden als nutzen und die erhofften Effekte möglicherweise nicht eingetreten sind. In diesem Kapitel erklären wir, wie es dazu kommen kann.

Empfehlungen zur richtigen Schlafposition von Säuglingen

Denken Sie bloß nicht, dass nur Medikamente Schaden anrichten können – auch Ratschläge können tödliche Folgen haben. Viele Menschen haben schon einmal von dem amerikanischen Kinder spezialisten Dr. Benjamin Spock gehört. Sein Bestseller *Baby and Child Care* (dt. *Säuglings- und Kinderpflege*) wurde vor allem in den USA und Großbritannien im Laufe mehrerer Jahrzehnte zu einer Art «Bibel». Bei einem sehr gut gemeinten Ratschlag hatte sich Dr. Spock allerdings furchtbar geirrt. Mit scheinbar unbestreitbarer Logik – und sicher auch mit einem gewissen Maß an Autorität – stellte er von 1956 bis in die späten 1970er-Jahre folgende Behauptung auf:

Es hat zwei Nachteile, wenn ein Baby auf dem Rücken schläft. Zum einen kann es, wenn es erbricht, eher an dem Erbrochenen ersticken. Zum anderen neigt ein Säugling dazu, seinen Kopf auf dieselbe Seite zu drehen ... dies könnte die Kopfseite abflachen ... Meiner Meinung nach sollte man ein Baby besser daran gewöhnen, von Anfang an auf dem Bauch zu schlafen.

Die Empfehlung, Babys zum Schlafen auf den Bauch zu legen, wurde in Krankenhäusern zum Standard und von Millionen Eltern auch zu Hause pflichtbewusst befolgt. Heute wissen wir jedoch, dass diese Maßnahme – die niemals gründlich evaluiert wurde – zu Tausenden vermeidbaren Fällen von plötzlichem Kindstod führte.¹ Als in den 1980er-Jahren dann eindeutige Belege für die schädlichen Wirkungen des Schlafens in Bauchlage verfügbar wurden, begannen Ärzte und die Medien vor den Gefahren dieser Praxis zu warnen, woraufhin die Anzahl der Fälle von plötzlichem Kindstod drastisch abzunehmen begann. Verstärkt wurde diese Botschaft später durch konzertierte «Wieder ruhig schlafen!»-Kampagnen, mit denen der negative Einfluss von Dr. Spocks bedauerlicher Empfehlung ein für alle Mal aus der Welt geschafft werden sollte (**Abb. 2**).

Medikamente zur Korrektur von Herzrhythmusstörungen bei Herzinfarkt-Patienten

Dr. Spocks Empfehlung mag logisch angemutet haben, doch beruhte sie auf einer ungeprüften Theorie. Andere Beispiele für die Gefährlichkeit eines solchen Vorgehens sind nicht besonders schwer zu finden. Nach einem Herzinfarkt kommt es bei manchen Menschen zu Herzrhythmusstörungen – sogenannten Arrhythmien. Patienten mit solchen Arrhythmien haben ein höheres Risiko, vorzeitig zu versterben, als Patienten ohne Arrhythmien. Da es Medikamente gibt, mit denen sich diese Arrhythmien unterdrücken lassen, schien die Annahme, dass diese Medikamente damit auch das Risiko, nach einem Herzinfarkt zu versterben, senken würden, nur folgerichtig zu sein. Tatsächlich hatten diese Medikamente aber genau den gegenteiligen Effekt. Sie waren zwar in klinischen Studien getestet worden, aber nur, um herauszufinden, ob sie Herzrhythmusstörungen verringerten. Als die aus klinischen Studien angesammelten Daten 1983 erstmals systematisch ausgewertet wurden, ergaben sich aber keine Hinweise darauf, dass diese Medikamente auch die Sterberate der betroffenen Patienten senkten.²

Dennoch wurden die Medikamente fast ein Jahrzehnt lang weiter verordnet – und waren auch weiterhin für den Tod von Menschen



Abbildung 2: Veränderung der Empfehlungen zur richtigen Schlafposition von Säuglingen im zeitlichen Verlauf.

verantwortlich. Auf dem Höhepunkt ihrer Anwendung in den späten 1980er-Jahren sollen einer Schätzung zufolge allein in den USA jedes Jahr Zehntausende von vorzeitigen Todesfällen durch diese Medikamente verursacht worden sein. Sie töteten jährlich mehr Amerikaner als die Kampfhandlungen im gesamten Vietnamkrieg.³ Später stellte sich heraus, dass die Ergebnisse einiger früher Studien, die erste Hinweise auf die tödliche Wirkung der Medikamente gegeben hatten, aus kommerziellen Gründen nie veröffentlicht worden waren (s. Kap. 8, S. 146).⁴

Diethylstilbestrol

Es gab eine Zeit, da bestand in der Ärzteschaft Unsicherheit darüber, ob man Schwangeren mit vorausgegangenem Fehl- und Totgeburten nicht mit einem synthetischen (nichtnatürlichen) Östrogen namens Diethylstilbestrol (DES) helfen könne. Manche Ärzte ver-

schrieben es, andere nicht. DES war in den frühen 1950er-Jahren recht populär geworden, denn man glaubte, es könne eine Funktionsstörung der Plazenta beheben, die für die Schwangerschaftsprobleme verantwortlich gemacht wurde. Ermutigt worden waren die Frauen, die DES anwendeten, durch einzelne (anekdotische) Fallberichte über Frauen mit früheren Fehl- und Totgeburten, die nach der Behandlung mit DES ein lebendes Kind geboren hatten.

So verordnete z.B. ein britischer Geburtshelfer einer Frau mit zwei vorausgegangenem Totgeburten das Medikament von der Frühschwangerschaft an. Die Schwangerschaft endete mit der Geburt eines lebenden Babys. Mit der Begründung, dass die «natürliche» Fähigkeit der Frau zum Austragen einer erfolgreichen Schwangerschaft sich im Laufe der Zeit möglicherweise verbessert habe, wurde während ihrer vierten Schwangerschaft auf die Gabe von DES verzichtet: Das Baby verstarb in der Gebärmutter aufgrund einer «Plazenta-Insuffizienz». Deshalb bestanden bei Mutter und Geburtshelfer während der fünften und sechsten Schwangerschaft der Frau auch keinerlei Zweifel, dass erneut DES gegeben werden sollte. Beide Schwangerschaften endeten mit lebendgeborenen Kindern. Sowohl der Geburtshelfer als auch seine Patientin schlossen daraus auf die Wirksamkeit von DES. Leider war die Richtigkeit dieser auf einzelnen Fallberichten beruhenden Schlussfolgerung nie in fairen Studien nachgewiesen worden. Tatsächlich waren nämlich im selben Zeitraum, in dem diese Frau behandelt wurde, unverzerrte Studien durchgeführt und veröffentlicht worden, die keine Beweise für den Nutzen von DES liefern konnten.⁵

Obwohl also keine Evidenz aus fairen Wirksamkeitsstudien vorlag, die belegte, dass DES zur Verhütung von Totgeburten wirksam war, endete die DES-Geschichte hier noch nicht. Hinweise auf die schädlichen Nebenwirkungen tauchten erst 20 Jahre später auf, als die Mutter einer sehr jungen, an einem seltenen Vaginalkarzinom erkrankten Frau eine sehr wichtige Beobachtung äußerte. Der Mutter war während der Schwangerschaft DES verordnet worden, und sie hatte den Verdacht, dass die Krebserkrankung ihrer Tochter durch dieses Medikament verursacht worden sein könnte.⁶ In diesem Fall war die Beobachtung richtig. Noch wichtiger aber war, dass auch der *Nachweis* ihrer Richtigkeit erbracht wurde. In zahlreichen

Studien sowohl bei Männern als auch Frauen, die im Mutterleib DES ausgesetzt gewesen waren, wurde seitdem eine Vielzahl an schwerwiegenden Nebenwirkungen von DES nachgewiesen. Dazu gehörten neben dem vermehrten Vorkommen seltener Krebsarten auch andere Anomalien des Fortpflanzungsapparats.

Als endlich offiziell verlautbart wurde, dass DES nicht in der Schwangerschaft angewendet werden sollte, waren bereits mehrere Millionen Menschen dem Medikament ausgesetzt worden. Nach heutigem Kenntnisstand kann man sagen, dass viel weniger Ärzte DES verschrieben hätten, wenn sie die sehr zuverlässige wissenschaftliche Evidenz zu DES genutzt hätten, die in den 1950er-Jahren verfügbar war. Denn der Beweis, dass DES bezüglich der Beschwerden, für die es in erster Linie verordnet wurde, auch wirkte, war im Grunde nie erbracht worden. Tragischerweise war das Fehlen der Wirksamkeitsbelege auf breiter Front übersehen worden.⁷

Hormonersatztherapie

Bei Frauen in der Menopause ist eine Hormonersatztherapie (engl. *hormone replacement therapy*, HRT) zur Linderung der häufig auftretenden und belastenden Hitzewallungen sehr wirksam, und es gibt etliche Anhaltspunkte dafür, dass sie auch vorbeugend gegen Osteoporose (Knochenschwund) wirkt. Nach und nach wurden der HRT immer mehr positive Effekte zugeschrieben, u. a. auch die Verhütung von Herzinfarkten und Schlaganfällen, sodass Millionen von Frauen auf ärztlichen Rat hin die HRT wegen dieser und weiterer angeblicher Vorteile längerfristig anzuwenden begannen. Die Grundlage für diese Behauptungen war allerdings mehr als schwach.

Sehen wir uns nur einmal die Herzinfarkte an. Mehr als 20 Jahre lang hat man Frauen erzählt, dass die HRT ihr Risiko für diese schwere Krankheit reduziert – tatsächlich beruhte diese Aussage auf den Ergebnissen verzerrter (unfairer) Studien (s. Kap. 1 und Kap. 6). 1997 wurde dann eine Warnung herausgegeben, dass diese Aussage falsch sein könnte: Forscher aus Finnland und Großbritannien hatten die Ergebnisse gut durchgeführter Studien systematisch ausgewertet⁸ und dabei festgestellt, dass die HRT, anstatt das Auftreten

Kein Wunder, dass sie verwirrt war

Im Januar 2004 schrieb eine Hysterektomie-Patientin an die britische medizinische Fachzeitschrift *The Lancet* folgenden Brief:

«1986 wurde bei mir wegen Gebärmutterfibromen eine Hysterektomie durchgeführt. Außer der Gebärmutter entfernte der Operateur dabei auch die Eierstöcke und stellte zudem fest, dass ich unter Endometriose litt. Weil ich damals erst 45 Jahre alt war und die sofortige Menopause eingesetzt hätte, erhielt ich eine Hormonersatztherapie (HRT). Im ersten Jahr nahm ich konjugierte Östrogene (Premarin) ein; zwischen 1988 und 2001 bekam ich als Privatpatientin dann aber von dem Chirurgen, der mich operiert hatte, alle sechs Monate Östrogenimplantate. Ich war bezüglich der Behandlung immer etwas unsicher, weil ich zum einen das Gefühl hatte, keine Kontrolle mehr zu haben, wenn das Implantat erst einmal eingesetzt war, und zum anderen, weil ich nach mehreren Jahren dann oft auch Kopfschmerzen bekam. Abgesehen davon fühlte ich mich aber sehr fit.

Mein Operateur versicherte mir, dass die HRT viele Vorteile hätte und mir gut täte, was ich nur bestätigen konnte. Mit der Zeit war von immer mehr Vorteilen der HRT zu lesen. Anders als in den Anfangsjahren galt die HRT inzwischen auch nicht mehr nur als Modetherapie. Jetzt war sie auch gut für das Herz, gegen Osteoporose, und teilweise sollte sie auch Schlaganfällen vorbeugen. Immer wenn ich meinen Operateur aufsuchte, schien er noch mehr Belege für die Vorteile der HRT parat zu haben.

Als er 2001 dann in den Ruhestand ging, wendete ich mich an meinen Arzt beim *National Health Service*. Was für ein Schock! Von ihm erfuhr ich genau das Gegenteil: dass es gut wäre, von der HRT abzukommen: Sie könne das Risiko für Herzkrankheiten, Schlaganfälle und Brustkrebs erhöhen und auch Kopfschmerzen auslösen. Ich erhielt dann noch ein weiteres Implantat und stellte für kurze Zeit dann wieder auf Premarin um, aber seitdem habe ich acht Monate lang keine HRT mehr gehabt. Mein Arzt meinte, es sei meine Entscheidung, ob ich die Therapie beibehalten wolle oder nicht. Ich war so was von verwirrt ...

Ich kann nicht verstehen, wie die HRT und all ihre wundervollen Vorzüge sich in so kurzer Zeit in ihr Gegenteil verkehren können. Wie kann ein Laie wie ich da noch die richtige Entscheidung treffen? Ich habe viele Stunden damit zugebracht, darüber zu reden und nachzudenken, ob ich besser bei der HRT geblieben wäre, obwohl ich seitdem kaum negative Wirkungen verspürt habe. Ich bin wegen dieser ganzen Sache sehr irritiert und bin mir sicher, dass es anderen Frauen ähnlich ergeht.»

Huntingford CA. Confusion over benefits of hormone replacement therapy.
Lancet 2004; 363: 332.

von Herzkrankheiten zu reduzieren, diese tatsächlich sogar begünstigen könne. Einige prominente Kommentatoren verwarfen diese Schlussfolgerung zwar, aber dieses vorläufige Ergebnis wurde mittlerweile durch zwei große, gut durchgeführte Studien bestätigt. Hätte man die Wirkungen der HRT bei ihrer Einführung bereits richtig bewertet, wären die betroffenen Frauen nicht falsch informiert wor-

den, und viele von ihnen wären nicht vorzeitig verstorben. Zu allem Übel kommt noch hinzu, dass wir heutzutage wissen, dass die HRT das Risiko für Schlaganfall und die Entwicklung von Brustkrebs erhöht.⁹

Insgesamt ist die HRT nach wie vor eine wertvolle Therapie für Frauen mit menopausaler Symptomatik.¹⁰ Tragisch ist allerdings, dass sie vor allem als Möglichkeit zur Verringerung des Herzinfarkt- und Schlaganfallrisikos so stark beworben wurde. Das Risiko für diese schweren Krankheiten mag zwar nur geringfügig erhöht sein, es betrifft aber, da die HRT so häufig verschrieben wurde, insgesamt doch eine sehr große Anzahl von Frauen.

Nachtkerzenöl bei Ekzemen

Auch wenn unzureichend bewertete Therapien nicht immer zum Tod führen oder Schaden verursachen, so können sie immerhin eine beträchtliche Geldverschwendung bedeuten. Ein Ekzem ist ein belastendes Hautleiden, von dem Kinder und Erwachsene gleichermaßen betroffen sind. Die Hautveränderungen sind nicht nur unansehnlich, sie verursachen auch starken Juckreiz. Die Anwendung von steroidhaltigen Salben hilft zwar, doch wurden wegen ihrer Nebenwirkungen (wie z. B. dem Dünnwerden der Haut) Bedenken geäußert. In den frühen 1980er-Jahren tauchte als mögliche Alternative mit wenigen Nebenwirkungen ein natürlicher Pflanzenöl-Extrakt, das Nachtkerzenöl, auf.¹¹ Nachtkerzenöl enthält eine essenzielle Fettsäure, die sogenannte Gamma-Linolensäure (engl. *gamma-linolenic acid*, GLA), deren Anwendung auf plausiblen Annahmen beruhte. Eine solche Annahme lautete beispielsweise, dass die Verstoffwechslung (Metabolisierung) von GLA im Körper bei Patienten mit einem Ekzem beeinträchtigt ist. Theoretisch sollte also die Gabe von GLA-Ergänzungsmitteln helfen. Ebenso empfohlen wurde bei Ekzemen Borretschöl (auch als Gurkenkrautöl bekannt), das sogar noch höhere Anteile an GLA enthält.

Man hielt GLA für sicher, aber war sie auch wirksam? Um das herauszufinden, wurden zahlreiche Studien durchgeführt, die jedoch widersprüchliche Ergebnisse lieferten. Die veröffentlichte Evi-

denz war außerdem stark durch Studien beeinflusst, die von den Herstellern der Ergänzungsstoffe gesponsert waren. 1995 beauftragte das britische Gesundheitsministerium Wissenschaftler, die keine Beziehungen zu den Herstellern von Nachtkerzenöl unterhielten, mit der Bewertung von 20 veröffentlichten und unveröffentlichten Studien: Es ließen sich keine Belege für die Wirksamkeit von Nachtkerzenöl finden (**Abb. 3**). Das Ministerium hat seinen Bericht nie veröffentlicht, weil die Hersteller des Präparats dagegen protestierten. Doch fünf Jahre später zeigte eine weitere, von denselben Wissenschaftlern erstellte systematische Übersichtsarbeit (systematischer Review) zu Nachtkerzenöl und Borretschöl – der dieses Mal publiziert wurde –, dass selbst die größten und weitgehend vollständigen Studien keine überzeugenden Belege für die Wirksamkeit dieser Behandlungen erbracht hatten.¹²

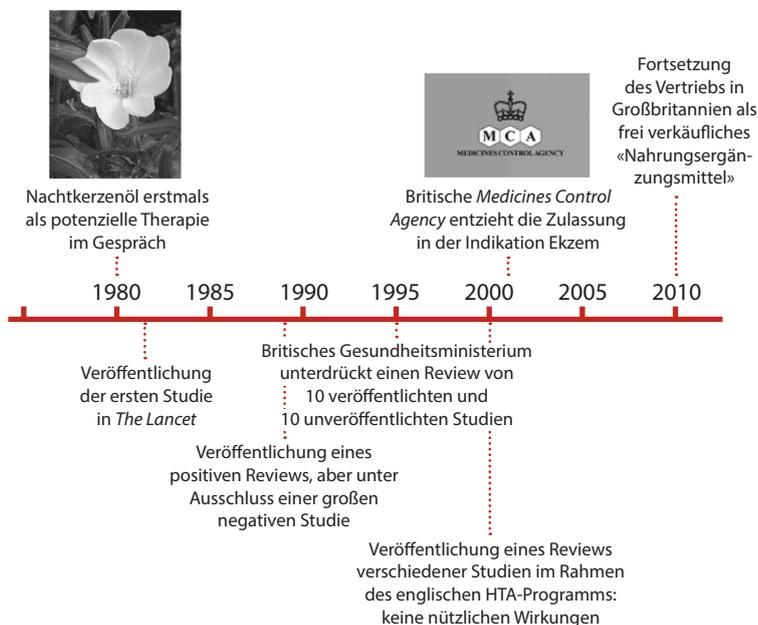


Abbildung 3: Chronologie der Evidenz zum Nachtkerzenöl und seiner Anwendung bei Ekzemen.

Eine Möglichkeit gab es allerdings noch: Vielleicht wirkte GLA ja nur in sehr hoher Dosierung. Aber 2003 wurde auch diese Behauptung durch eine sorgfältig durchgeführte faire Studie widerlegt.¹³ Wie es die Ironie der Geschichte wollte, hatte die britische Zulassungsbehörde, die *Medicines Control Agency*, aber bereits im Oktober 2002 – lange bevor diese neuen Ergebnisse veröffentlicht wurden – den Entzug der Produktzulassungen für zwei wichtige Nachtkerzenözübereitungen verfügt, weil ihre Wirksamkeit nicht belegt war.

Da jedoch keine Bedenken hinsichtlich der Sicherheit von Nachtkerzenöl verlautbart wurden, ist es noch immer als frei verkäufliches «Nahrungsergänzungsmittel» für verschiedenste Beschwerden erhältlich. Was seinen Nutzen bei Ekzemen anbelangt, so finden sich die Wirksamkeitsbehauptungen in vagen Formulierungen wieder wie «kann bei Ekzem-Patienten lindernd wirken», «kann helfen» und «verfügt über bestimmte heilende Kräfte, die beispielsweise bei Ekzemleiden entzündungshemmend wirken können».

- Weder Theorien noch Fachmeinungen sind für sich allein ein verlässlicher Ratgeber für sichere, wirksame Behandlungen.
- Nur weil eine Behandlung «etabliert» ist, bedeutet dies nicht, dass sie mehr nützt als schadet.
- Selbst wenn Patienten durch unzureichend getestete Behandlungen keinen Schaden nehmen, so kann ihre Anwendung doch eine Verschwendung individueller und öffentlicher Ressourcen darstellen.

3

Mehr heißt nicht unbedingt besser

Ein weit verbreiteter Irrglaube besagt: Wenn eine Behandlung gut ist, dann muss mehr davon noch besser sein. Das ist aber schlicht falsch. Denn es hat sich herausgestellt, dass «mehr Behandlung» auch schlechter sein kann. Die «richtige» Dosis zu finden – bei der der Nutzen hoch ist und nur geringfügige unerwünschte Nebenwirkungen auftreten – stellt bei allen medizinischen Behandlungen eine Herausforderung dar. Wird die Dosis erhöht, dann erreichen die nützlichen Effekte irgendwann ein Plateau; die unerwünschten Wirkungen nehmen dagegen normalerweise noch zu. Das bedeutet: «Mehr Therapie» kann den eigentlichen Nutzen mindern oder insgesamt sogar Schaden verursachen.

Ein gutes Beispiel sind Diuretika (Entwässerungstabletten): Niedrig dosiert senken sie den Blutdruck bei nur wenigen unerwünschten Wirkungen. Eine Erhöhung der Dosis bewirkt keine weitere Blutdrucksenkung, kann aber unerwünschte Wirkungen wie übermäßige Harnausscheidung, Impotenz und erhöhte Blutzuckerspiegel hervorrufen. Ein ähnliches Beispiel ist Aspirin: In niedrigen Dosen trägt es bei nur sehr wenigen unerwünschten Wirkungen zur Vorbeugung gegen Schlaganfälle bei. Mehrere Aspirin-tabletten pro Tag können zwar Kopfschmerzen lindern, bewirken aber keine bessere Schlaganfallprävention und erhöhen gleichzeitig das Risiko für Magengeschwüre.

Dieses Prinzip der «richtigen Dosis» hat über die medikamentöse Therapie hinaus auch für viele andere Behandlungen Gültigkeit, so auch für Operationen.

Intensivtherapien bei Brustkrebs

Über die mit der Annahme, dass intensivere Therapien zwangsläufig vorteilhaft sind, verbundenen Gefahren können wir einige besonders wertvolle Lektionen aus den bei Brustkrebs empfohlenen Therapien ziehen, von denen in den Nachrichten so oft zu hören ist.

Wir tun das, weil ...

«Wir [Ärzte] tun manchmal Dinge, weil andere Ärzte es auch so machen und wir nicht anders sein wollen als sie; aus diesem Grund machen wir es genauso, oder weil es uns [von Dozenten, Ober- und Assistenzärzten] so beigebracht wurde oder weil wir [von Dozenten, Verwaltungsdirektoren, Aufsichtsbehörden, Leitlinienentwicklern] dazu angehalten wurden und denken, dass wir das so machen müssen, oder weil die Patienten es so wollen und wir der Meinung sind, dass wir ihren Wünschen nachkommen sollten, oder weil wir das Gefühl haben, dass wir das wegen größerer Anreize [unnötige Untersuchungen, vor allem bei apparateorientierten Ärzten, und Arztbesuche] oder aus Angst [vor dem Rechtssystem, vor Audits] so machen sollten, oder wir machen es so, weil wir Zeit gewinnen wollen [um der Natur ihren Lauf zu lassen]; letzten Endes und gar nicht so selten machen wir es aber, damit wir etwas zu tun haben und weil wir unseren gesunden Menschenverstand nicht einschalten.»

Parmer MS. We do things because (rapid response). BMJ.
Eingestellt am 1. März 2004 unter www.bmj.com.

Während des gesamten 20. Jh. und bis in das 21. Jh. hinein haben Frauen mit Brustkrebs etliche ausgesprochen brutale und qualvolle Behandlungen sowohl verlangt als auch über sich ergehen lassen. Einige dieser operativen und medikamentösen Behandlungen gingen aber weit über das hinaus, was zur Bekämpfung der Krankheit eigentlich nötig gewesen wäre. Bei etlichen Patientinnen wie auch bei ihren Ärzten waren diese Therapien zweifellos aber auch durchaus populär. Je radikaler oder toxischer die Therapie – so die Überzeugung der Patientinnen –, umso größer auch ihre Chance, die Krankheit besiegen zu können. Ärzte und Patientinnen, die bereit waren, diese orthodoxen Ansichten über die Krankheit zu hinterfragen, kostete es viele Jahre, bevor es ihnen gelang, den Trend dieses Irrglaubens umzukehren. Sie mussten nicht nur zuverlässige Evidenz zusammentragen, um den Mythos des «Mehr ist besser» zu bannen, sondern auch noch den Spott ihrer Kollegen und den Widerstand hoch angesehener Praktiker ertragen.

Auch heute noch sind die Angst und der Glaube, dass mehr besser sein müsse, der Motor unserer Behandlungsentscheidungen, selbst wenn es keine Belege für ihren Nutzen gegenüber einfacheren Behandlungsansätzen gibt und bekanntermaßen beträchtliche schädliche Wirkungen mit ihnen einhergehen, darunter auch das Risiko, an der Behandlung selbst zu sterben. Diese Einstellung treibt manche Patientinnen und ihre Ärzte auch heute noch dazu, die «traditionelle» verstümmelnde Operation zu wählen. Andere entscheiden sich, selbst wenn einfachere Behandlungen ausreichend wären, für eine hochdosierte Chemotherapie mit ihren bekannten unangenehmen wie auch schmerzhaften Nebenwirkungen oder für Herceptin, das schwerwiegende Herzbeschwerden hervorrufen kann (s. Kap. 1). Wie kommt das?

Drastische Therapien sind nicht immer am besten

«Diejenigen von uns, die Krebserkrankungen behandeln, können sich sehr leicht vorstellen, dass eine drastischere Behandlung zu besseren Ergebnissen führt. Um die Patienten vor sinnlosen Risiken und den frühen oder späten Nebenwirkungen unnötig aggressiver Therapien zu schützen, kommt den randomisierten Studien, in denen drastische mit weniger radikalen Therapien verglichen werden, eine entscheidende Bedeutung zu. Ein solcher Vergleich ist ethisch vertretbar, weil diejenigen, denen ein möglicher Nutzen vorenthalten wird, gleichzeitig auch vor möglichen unnötigen Schäden bewahrt bleiben – und niemand weiß, welche Richtung die Entwicklung am Ende nimmt.»

Brewin T in Rees G, ed. *The friendly professional: selected writings of Thurstan Brewin*. Bognor Regis: Eurocommunica, 1996.

Verstümmelnde Operationen

Bis Mitte des 20. Jahrhunderts stellte die Operation die wichtigste Behandlungsmethode bei Brustkrebs dar. Dies gründete auf der Annahme, dass der Krebs langsam und geordnet wächst und sich zunächst vom Tumor in der Brust auf die örtlichen (lokalen) Lymphknoten, etwa in der Achselhöhle, ausbreitet. Je radikaler und rascher der Tumor operiert würde, so die Argumentation, umso größer sei die Chance, die Ausbreitung des Krebses aufzuhalten. Die Therapie bestand aus einer ausgedehnten «lokalen» Operation, d.h. aus einem operativen Eingriff an oder nahe der Brust. Man mag diesen

Eingriff zwar lokal genannt haben, doch ist eine radikale Mastektomie alles andere als lokal: Sie umfasst die Entfernung eines großen Teils des Brustmuskels und des Lymphknotengewebes aus der Achselhöhle und der Brust selbst.

Die klassische radikale Mastektomie (nach Halsted)
<p>Die im späten 19. Jahrhundert von William Halsted entwickelte radikale Mastektomie war bis in das dritte Quartal des 20. Jahrhunderts die am häufigsten durchgeführte Operation bei Brustkrebs. Dabei entfernte der Operateur nicht nur die gesamte Brust, sondern auch den großen Brustmuskel (Musculus pectoralis major), der die Brustwand überzieht. Damit der Operateur einen leichteren Zugang zur Achselhöhle (Axilla) hatte, um dort die Lymphknoten und das umgebende Fettgewebe ausräumen zu können, wurde zudem auch noch der kleine Brustmuskel (Musculus pectoralis minor) entfernt.</p>
Die ausgedehnt radikale Mastektomie
<p>Der Glaube, mehr sei besser, veranlasste radikale Operateure dazu, sogar noch ausgedehntere Operationen vorzunehmen, bei denen auch die Lymphknotenketten unter dem Schlüsselbein und die Mammaria-interna-Lymphknoten unter dem Brustbein entfernt wurden. Um an die Mammaria-interna-Lymphknoten zu gelangen, mussten mehrere Rippen entfernt und das Brustbein mit einem Meißel gespalten werden. Damit noch nicht zufrieden, gingen einige Operateure sogar so weit, auch noch den Arm auf der betroffenen Seite zu amputieren und im ganzen Körper diverse Drüsen (Nebennieren, Hypophyse, Eierstöcke) herauszuschneiden, um die Produktion von Hormonen zu unterdrücken, von denen man annahm, dass sie die Ausbreitung des Tumors begünstigten.</p> <p>Wenn eine Frau solche Operationen überlebte, behielt sie einen stark verstümmelten Brustkorb zurück, was sich auch durch Kleidung nur schwer verbergen ließ. War die Operation auf der linken Seite durchgeführt worden, dann blieb dort nur eine dünne Hautschicht zur Abdeckung des Herzens übrig.</p> <p style="text-align: right;">Nach: Lerner BH. <i>The breast cancer wars: hope, fear and the pursuit of a cure in twentieth-century America</i>. New York; Oxford University Press, 2003.</p>

Nichtsdestotrotz bemerkten etliche aufmerksame Brustkrebsspezialisten, dass diese zunehmend verstümmelnden Operationen anscheinend überhaupt keinen Einfluss auf die brustkrebsbedingten Sterblichkeitsraten hatten (**Abb. 4**). Deshalb stellten sie eine ganz andere Theorie auf: Danach breitete sich der Brustkrebs nicht von der Brust über die nahe gelegenen Lymphknoten aus, sondern es handelte sich dabei von Anfang an um eine systemische (d.h. den ganzen Körper betreffende) Krankheit. Mit anderen Worten: Sie gingen davon aus, dass zum Zeitpunkt der Entdeckung des Knotens in der Brust anderswo im Körper bereits Krebszellen vorhanden

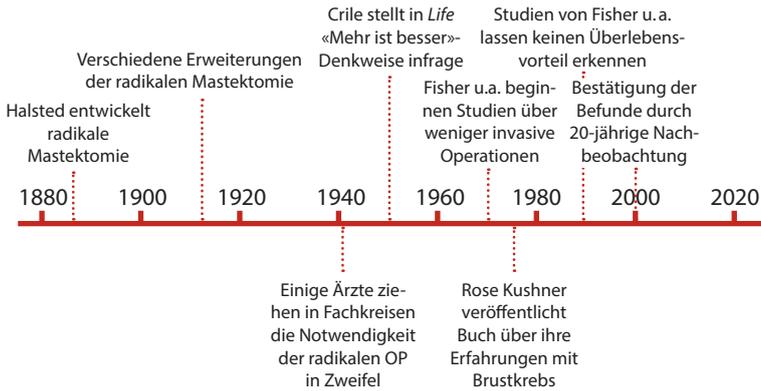


Abbildung 4: Infragestellung des «Mehr ist besser»-Ansatzes in der Brustkrebschirurgie.

sein müssten (s.u.). Wenn diese Annahme zuträfe, so ihre Vermutung, dann wäre die operative Entfernung des Tumors mit einem angemessenen Sicherheitssaum von gesundem Gewebe sowie die zusätzliche Durchführung einer lokalen Strahlentherapie für die betroffenen Frauen nicht nur angenehmer, sondern könnte auch genauso wirksam sein wie eine radikale Operation. Auf dieser neuen Theorie der Brustkrebsausbreitung beruhte zu der Zeit auch die Einführung von «systemischen Therapien» – d.h. von Behandlungen, die auf die Produktion oder Entstehung von Krebszellen an anderen Stellen im Körper abzielten.

Als direkte Folge dieser neuen Sichtweise traten die Ärzte für einen eingeschränkten Eingriff, die sogenannte Lumpektomie ein, in deren Rahmen der Tumor und ein Saum aus umgebendem gesundem Gewebe entfernt wurden. Im Anschluss an die Lumpektomie wurde eine Strahlentherapie und bei einigen Frauen auch noch eine Chemotherapie durchgeführt. Die Befürworter der Lumpektomie stießen jedoch auf gewaltige Widerstände, als sie diesen neuen Ansatz mit der Radikaloperation verglichen. Manche Ärzte waren sehr fest von der einen und manche sehr fest von der anderen Methode überzeugt, und auch die Patientinnen verlangten lautstark nach der einen oder der anderen Behandlung. Die Folge war, dass geraume Zeit verstrich, bis die entscheidenden Belege für die Vor- und Nach-

teile des neuen Behandlungsansatzes im Vergleich zum alten erbracht werden konnten.

Ungeachtet dieser Schwierigkeiten wurde letztlich die chirurgische Überbehandlung aber infrage gestellt, und zwar sowohl von Operateuren, die angesichts des fragwürdigen Nutzens für ihre Patientinnen nicht mehr bereit waren, mit dieser Praxis fortzufahren, als auch von entschlossenen Frauen, die nicht länger gewillt waren, sich verstümmelnden Eingriffen zu unterziehen.

Den Anfang machte Mitte der 1950er-Jahre der amerikanische Chirurg George Crile, als er mit seinen Bedenken über den «Mehr ist besser»-Ansatz an die Öffentlichkeit ging. Überzeugt davon, dass es keinen anderen Weg gab, um die Ärzteschaft zu einem kritischeren Denken anzuregen, veröffentlichte Crile in einem Beitrag zur bekannten Zeitschrift *Life* einen Appell.¹ Er traf den richtigen Ton: Die Debatte innerhalb der Ärzteschaft wurde damit in die Öffentlichkeit getragen und war nicht mehr auf rein akademische Kreise beschränkt. Daraufhin entwickelte Bernard Fisher, ein weiterer amerikanischer Chirurg, der mit Kollegen aus anderen Fachgebieten zusammenarbeitete, eine Reihe von aussagekräftigen Experimenten zur Untersuchung der Biologie des Krebses. Ihre Ergebnisse ließen darauf schließen, dass sich Krebszellen noch vor der Entdeckung des Primärtumors tatsächlich über den Blutstrom im ganzen Körper verbreiten können. So gesehen machten aggressive Operationen wenig Sinn, wenn sich bereits an anderen Stellen im Körper Krebszellen befinden konnten.

Während Crile klinisches Urteilsvermögen einsetzte, um weniger radikale lokale Therapien zu empfehlen und umzusetzen, arbeiteten Fisher und eine immer größer werdende Gruppe von Wissenschaftlern an einem formelleren und rigoroseren Ansatz. Sie versuchten, den Wert der radikalen Operation anhand der bekanntesten fairen (unverzerrten) Methode, und zwar anhand von randomisierten Studien (s. Kap. 6), zu beweisen bzw. zu widerlegen. Mithilfe solcher Studien wollten sie die medizinische Fachwelt und die Öffentlichkeit von dem einen oder dem anderen Therapieansatz überzeugen. Wie Fisher 1971 zudem geradeheraus erklärte, hätten Chirurgen die ethische und moralische Verpflichtung, ihre Theorien zu testen, indem sie solche Studien durchführten. Die 20-jährige Nachbeob-

achtung (Follow-up) von Fishers Studien hat zweifellos gezeigt, dass für die radikale Mastektomie im Vergleich zur Lumpektomie mit anschließender Strahlentherapie kein Vorteil – gemessen am Risiko, vorzeitig zu versterben – nachgewiesen werden konnte.²

Randomisierte Zuteilung – eine einfache Erklärung

«Randomisierung soll systematische Fehler (Bias) auf ein Minimum begrenzen und sicherstellen, dass sich die Patienten in jeder Behandlungsgruppe hinsichtlich aller bekannten und unbekanntem Faktoren möglichst ähnlich sind. Dadurch wird gewährleistet, dass alle Unterschiede, die man bei der/den interessierenden Zielgröße(n) zwischen den Gruppen findet, durch die Unterschiede in der Wirkung der Behandlung (im Therapieeffekt) bedingt und nicht auf die Unterschiede zwischen den Patienten zurückzuführen sind, welche die jeweilige Behandlung erhalten haben.

Randomisierung schaltet die Möglichkeit aus, dass ein Arzt die eine Therapie – bewusst oder unbewusst – einer ganz bestimmten Art von Patient zuteilt und die andere einem anderen Patiententyp oder dass eine Patientengruppe die eine Behandlung wählt, während eine andere Gruppe sich für die andere Behandlung entscheidet.»

Harrison J. Presentation to Consumers' Advisory Group for Clinical Trials, 1995.

Auch Forscher in anderen Ländern führten randomisierte Studien (s. Kap. 6) durch, in denen die brusterhaltende Therapie mit der radikalen Mastektomie verglichen wurde. Zu nennen sind hier beispielsweise Hedley Atkins und Kollegen Anfang der 1960er-Jahre in Großbritannien und später dann Veronesi et al. in Italien. Das Gesamtbild, das sich dabei ergab, bestätigte Fishers Ergebnisse: Es gab auch nach 20 Jahren Nachbeobachtung keine Belege dafür, dass die radikale Mastektomie zu einem längeren Überleben führte.³ Weitere randomisierte Studien – wurden in Schweden und Italien sowie in Großbritannien und den USA durchgeführt, um zahlreiche andere Behandlungsformen miteinander zu vergleichen, beispielsweise Strahlentherapie nach Operation im Vergleich zu alleiniger Operation oder kurzzeitige gegenüber langzeitigen Chemotherapien.

Insgesamt stützten die Ergebnisse aus diesen frühen Studien und aus detaillierten Laborstudien die Theorie, dass es sich beim Brustkrebs tatsächlich um eine systemische (d.h. den ganzen Körper betreffende) Erkrankung handelt, bei der sich die Krebszellen über die Blutbahn ausbreiten, bevor ein Knoten in der Brust nachweisbar ist.⁴ Weltweit ließen sich immer mehr Ärzte durch die wachsende

Evidenz davon überzeugen, dass eine radikale Operation mehr Schaden anrichtet als nützt. Und in den letzten Jahrzehnten des 20. Jahrhunderts begannen sich auch die Einstellungen der Patientinnen und der Öffentlichkeit zu wandeln. Angeführt durch die Arbeit von Patientenaktivisten wie Rose Kushner (s. Kap. 11) in den USA und anderswo, fanden sich auf der ganzen Welt besser informierte Patientengruppen zusammen, um den «Mehr ist besser»-Ansatz in der chirurgischen Brustkrebsbehandlung und den damit oftmals einhergehenden ärztlichen Paternalismus kritisch zu hinterfragen.

Diese auf breiter Front stattfindenden Aktivitäten von Patientinnen und Ärzten stellten die chirurgische Überbehandlung der vergangenen Jahre fast überall erfolgreich infrage. Angesichts dessen ist es nahezu unglaublich, dass sich trotzdem noch immer Berichte über die Durchführung unnötiger und verstümmelnder Brustoperationen finden: So wurden etwa 2003 in Japan noch mehr als 150 radikale Brustoperationen ausgeführt.⁵

Bis zum Jahr 1985 war es allein aufgrund der ungeheuren Anzahl von Brustkrebsstudien zu allen möglichen Aspekten der Therapie sehr schwer, sich hinsichtlich der neusten Ergebnisse auf dem Laufenden zu halten. Um dieses Problem zu lösen, fassten Richard Peto und Kollegen in Oxford alle Studienergebnisse mit sämtlichen Informationen über alle Frauen, die an den vielen Studien teilgenommen hatten, in einer ganzen Reihe von systematischen Übersichtsarbeiten (systematische Reviews; s. Kap. 8) zusammen.⁶ Mittlerweile werden die systematischen Reviews zu Brustkrebstherapien regelmäßig auf den neuesten wissenschaftlichen Stand gebracht und publiziert.^{7, 8}

Knochenmarktransplantation

Dennoch bedeutete das Ende der verstümmelnden Chirurgie nicht auch das Ende der «Mehr ist besser»-Denkweise – im Gegenteil. Während der letzten 20 Jahre des 20. Jh. wurde ein neuer Ansatz für die Therapie des Brustkrebses eingeführt: eine hochdosierte Chemotherapie mit anschließender Knochenmarktransplantation oder «Stammzellrettung». Ein Bericht in der *New York Times* aus dem Jahr 1999 fasste die Gründe für diesen Ansatz wie folgt zusammen:

Ärzte entnehmen der Patientin etwas Knochenmark oder rote Blutzellen und verabreichen ihr dann toxische Medikamente in sehr hohen Dosen, die das verbleibende Knochenmark zerstören. Dahinter steht die Hoffnung, dass diese hohen Dosen die Krebszellen vernichten und das zuvor entnommene Knochenmark bei seiner Rückführung in den Körper schnell genug nachwächst, damit die Patientin nicht an einer Infektion stirbt. Eine Variante dieser Maßnahme, bei der Knochenmarkspenden verwendet werden, hatte sich über lange Zeit bei Blutkrebs bewährt, aber nur deshalb, weil der Krebs das Knochenmark befallen hatte, das dann ersetzt wurde. Die Anwendung dieser Behandlung bei Brustkrebs beruht auf einer ganz anderen – und ungeprüften – Argumentation.⁹

Vor allem in den USA bedrängten Tausende von verzweifelten Frauen Ärzte und Krankenhäuser, diese ausgesprochen unangenehme Behandlung bei ihnen vorzunehmen, obwohl nicht weniger als fünf von 100 Patientinnen an den Folgen der Behandlung verstarben. Viele tausend Dollar wurden ausgegeben, und nicht wenige Patientinnen bezahlten die Therapie aus eigener Tasche. Einige Krankenversicherer gaben dem öffentlichen Druck nach und übernahmen bei einigen Patientinnen die Kosten, obwohl keine Belege für die Wirksamkeit der Behandlung vorlagen. Viele Krankenhäuser und Kliniken wurden durch diese Einnahmen reich. 1998 erzielte eine Krankenhausgesellschaft Einkünfte in Höhe von 128 Millionen Dollar, die größtenteils aus ihren Krebszentren stammten, in denen Knochenmarktransplantationen angeboten wurden. Für US-amerikanische Ärzte war dies eine lukrative Einkommensquelle, die ihnen zu Ansehen verhalf und ein weites Publikationsfeld eröffnete. Die hartnäckige Nachfrage vonseiten der Patientinnen heizte den Markt an. Unter den privat geführten US-amerikanischen Krankenhäusern, welche die Behandlungen anboten, herrschte ein intensiver Wettbewerb, der sogar so weit ging, dass mit Sonderangeboten geworben wurde. In den 1990er-Jahren wurde diese Behandlung in den USA sogar von medizinischen Lehrzentren angeboten, die versuchten, Patienten für klinische Studien zu rekrutieren. Diese fragwürdigen Programme waren für die Krebszentren zu einem wahren Goldesel geworden.

Der uneingeschränkte Zugang zu einer solchen ungeprüften Therapie hatte noch einen weiteren schwerwiegenden Nachteil: Es standen nicht genug Patientinnen für Studien zur Verfügung, in denen diese Therapien mit Standardbehandlungen verglichen werden sollten. Infolgedessen dauerte es deutlich länger als vorgesehen, bis verlässliche Antworten aus fairen Studien vorlagen.

Doch trotz der Schwierigkeiten, unter diesem Druck zu unverzerrten wissenschaftlichen Belegen (Evidenz) zu gelangen, wurden etliche klinische Studien durchgeführt und andere Evidenz kritisch geprüft. Ein systematischer Review, in dem die akkumulierten Studienergebnisse der konventionellen Chemotherapie mit denen einer hochdosierten Chemotherapie und anschließender Knochenmarktransplantation als allgemeiner Behandlungsstrategie bei Brustkrebs verglichen wurden, kam 2004 zu der Erkenntnis, dass es keine überzeugenden Belege für einen Nutzen der Chemotherapie plus Knochenmarktransplantation gab.^{10, 11}

Der Kampf um unverzerrte Evidenz

«Die Wissenschaftler waren davon ausgegangen, dass es ungefähr drei Jahre dauern würde, um ca. 1 000 Frauen in die beiden Studien aufzunehmen. Stattdessen benötigte man dazu sieben Jahre ... Das ist nicht weiter überraschend ... Patientinnen, die an klinischen Studien teilnehmen, müssen eine Einverständniserklärung unterzeichnen, in der ihre trostlose Prognose explizit benannt und darauf hingewiesen wird, dass es keinerlei Belege dafür gibt, dass eine Knochenmarktransplantation besser ist als die Standardtherapien. Um in eine Studie aufgenommen zu werden, muss man sich diesen Tatsachen stellen, und das ist in keinem Fall einfach. Findet eine solche Transplantation außerhalb von Studien statt, die auch über eine Kontrollgruppe von Patientinnen verfügen (also außerhalb von sogenannten randomisierten Studien), dann würden begeisterte Ärzte der Patientin vielleicht weismachen, dass eine Transplantation ihr Leben retten könne. Patienten haben ein Recht auf die Wahrheit, doch werden sie verständlicherweise nicht gerade Ärzte aufsuchen, die ihnen ihre Hoffnung rauben.»

Nach Kolata G, Eichenwald K. Health business thrives on unproven treatment, leaving science behind. New York Times Special Report, 2. Oktober 1999.

Den Gedanken wagen, weniger zu tun

Mehr ist also nicht immer besser – und diese Botschaft hat auch heute nichts von ihrer Bedeutung verloren. Heute blicken Frauen mit metastasiertem (gestreutem) Brustkrebs hoffnungsvoll auf Therapien wie Herceptin (s.o. und Kap. 1). Und doch bietet Herceptin diesen Patientinnen bestenfalls eine geringe Chance auf ein längeres Leben – das sich zuweilen nur in Tagen oder Wochen bemisst – auf Kosten von schwerwiegenden Nebenwirkungen oder manchmal sogar eines durch die Therapie selbst bedingten Todes.^{12,13} Diese Tendenz zur Überbehandlung ist auch am anderen Ende des Brustkrebspektrums zu beobachten. So werden hier beispielsweise bei Frauen, bei denen im Rahmen des Mammographie-Screenings Krebsvorstufen wie das duktales Carcinoma in situ (DCIS) entdeckt werden (s. Kap. 4), übermäßige und oftmals unnötige Behandlungen angewendet, obwohl es durchaus sein kann, dass ein solches DCIS auch ohne Behandlung ein Leben lang nie Probleme verursachen wird. Inzwischen wird die Notwendigkeit einer routinemäßigen Entfernung der Lymphknoten in der Achselhöhle, die mit unangenehmen Komplikationen am betroffenen Arm wie etwa einem Lymphödem (s. Kap. 5) einhergehen kann, zunehmend infrage gestellt, da ihre zusätzliche Anwendung zu anderen Behandlungen das Überleben nicht zu verbessern scheint.¹⁴

- Eine intensivere Therapie ist nicht zwangsläufig von Vorteil und kann manchmal auch mehr schaden als nützen.

4

Früher ist nicht zwangsläufig besser

In den ersten drei Kapiteln haben wir dargelegt, dass nicht ausreichend geprüfte Therapien ernsthaften Schaden anrichten können. In diesem Kapitel geht es nun um die Untersuchung von anscheinend gesunden Menschen zur Erkennung von Krankheitsfrühzeichen, das sogenannte Screening. Screening klingt vernünftig – wie sonst könnte man schwerwiegende Krankheitsfolgen besser abwenden und gesund bleiben? Doch auch wenn Screeninguntersuchungen bei einigen Krankheiten sinnvoll sind, können sie nicht nur nützen, sondern auch schaden.

Wir werden hier auf verschiedene Krankheitsbeispiele zurückgreifen, an denen wir zeigen wollen, warum eine frühzeitige Diagnose besser sein kann, aber nicht immer besser sein muss, warum zahlreiche Screenings keinen oder nur einen unbestimmten Nutzen aufweisen und wie die Vorteile von Screeninguntersuchungen oftmals übermäßig angepriesen und ihre schädlichen Effekte heruntergespielt oder gänzlich ignoriert werden.

Vom Menschen zum Patienten

Aus Menschen, die «positiv» getestet werden, macht das Screening zwangsläufig Patienten – eine Veränderung, die man nicht so ohne weiteres wegsteckt. «Wenn ein Patient einen Arzt um Hilfe bittet, wird der Arzt das Bestmögliche für ihn tun. Der Arzt ist nicht für Wissenslücken in der Medizin verantwortlich. Veranlasst er jedoch Screeningmaßnahmen, dann stellt sich die Situation ganz anders dar. Unserer Ansicht nach sollte der Arzt über schlüssige Belege dafür verfügen, dass das Screening den natürlichen Krankheitsverlauf bei einem signifikanten Anteil der gescreenten Personen verändern kann.»

Cochrane AL, Holland WW. Validation of screening procedures.
British Medical Bulletin 1971; 27: 3–8.

Das Screening von gesunden Menschen sollte nicht leichtfertig erfolgen; es hat immer auch gewichtige Nachteile, die uns ein umsichtiges Vorgehen abverlangen. Eine Screeninguntersuchung ist eine medizinische Maßnahme (Intervention). Aber nicht nur das; schon das Angebot eines Screenings stellt eine Intervention dar. Selbst bei jemandem, der dieses Angebot ablehnt, hinterlässt es einen nagenden Zweifel, ob er die «richtige» Entscheidung getroffen hat oder nicht – das liegt in der menschlichen Natur. Etwas ganz anderes ist es, wenn einem ein Screening gar nicht erst angeboten wird.

Ein Screening sollte gesunden Menschen, die dadurch Gewissheit erlangen oder einer Behandlung zugeführt werden können, nur angeboten werden, wenn es fundierte Belege dafür gibt, dass das Screening (a) bei annehmbaren Kosten mehr nützt als schadet und (b) im Rahmen eines hochwertigen und gut organisierten Programms angeboten wird (s. u.).¹

Screening ist viel mehr als ein einmaliger Test. Menschen, die zur Teilnahme an einem Screening eingeladen werden, benötigen ausreichend unverzerrte, relevante Informationen, damit sie entscheiden können, ob sie das Angebot annehmen wollen oder nicht, d. h., sie müssen wissen, worauf sie sich einlassen (s. u.).²

Tatsächlich sind Screeninguntersuchungen dann und nur dann sinnvoll, wenn für die aufzudeckende Erkrankung auch eine wirksame Behandlung existiert.

Lehren aus dem Neuroblastom-Screening

Die Erfahrungen mit dem Screening auf ein Neuroblastom – eine seltene Krebsart, die vorwiegend bei kleinen Kindern auftritt – sind in mehrfacher Hinsicht lehrreich. Dieser Tumor befällt die Nervenzellen in verschiedenen Teilen des Körpers. Die Überlebensraten der betroffenen Kinder hängen u. a. davon ab, welcher Körperteil betroffen ist, wie weit sich der Tumor bei der Diagnosestellung schon ausgebreitet hat und wie alt das Kind ist. Die 5-Jahres-Gesamtüberlebensrate bei der Diagnose von Kindern im Alter zwischen 1 und 4 Jahren beträgt rund 55%.³ Eine merkwürdige Eigenschaft des Neuroblastoms ist, dass es zu den wenigen Krebsarten gehört, die

auch ohne Behandlung manchmal vollkommen verschwinden – ein Phänomen, das man als Spontanheilung oder Spontanregression bezeichnet.⁴

Aus vier Gründen war das Neuroblastom ein verlockendes Screeningziel: Man wusste erstens, dass Kinder, bei denen die Diagnose vor dem ersten Lebensjahr gestellt wird, eine bessere Prognose haben als Kinder, bei denen das Neuroblastom später diagnostiziert wird; zweitens war bekannt, dass es Kindern mit fortgeschrittener Erkrankung deutlich schlechter ergeht als Kindern mit einem Neuroblastom im Frühstadium; dass es drittens einen einfachen und kostengünstigen Screeningtest gibt, bei dem man nur die feuchte Windel abtupfen und bestimmte Substanzen im Urin messen muss; und dass man viertens durch diesen Test neun von zehn Kindern mit einem Neuroblastom identifizieren kann.⁵

Das Massenscreening auf ein Neuroblastom bei Säuglingen im Alter von sechs Monaten wurde erstmals 1985 in Japan eingeführt. Während der ersten drei Jahre der landesweiten Screeningaktion wurden mehr als 337 Säuglinge diagnostiziert, von denen 1990 nach Therapie noch 97 % am Leben waren. Aber auch 20 Jahre später gab es keine Belege dafür, dass das Neuroblastom-Screening die Anzahl der Kinder, die an dieser Krebsart verstarben, reduzierte. Wie konnte das sein?

Als man die Evidenz, auf deren Grundlage das Screening in Japan eingeführt und gefördert worden war, genauer unter die Lupe nahm, stellte sich heraus, dass sie ernsthafte Schwachstellen aufwies – doch hatte man auch gleich eine Erklärung bei der Hand. Die beeindruckende Zahl von 97 % Überlebenden veranschaulicht den Effekt eines Phänomens, das man fachsprachlich als «Length-Time-Bias» oder «Überdiagnose-Bias» bezeichnet. Es bedeutet, dass Screening sich am besten zur Erkennung von langsam fortschreitenden Krankheiten eignet (in diesem Fall langsam wachsender Tumoren). Schnell wachsende Tumoren werden dagegen durch ein solches Screening eher seltener entdeckt; sie äußern sich beim Säugling vielmehr in Krankheitssymptomen – beispielsweise einer Schwellung im Abdomen (Bauchraum) –, deretwegen die Eltern ohnehin meist unverzüglich ärztlichen Rat einholen. Diese schnell wachsenden Tumoren sind sehr viel gefährlicher als die langsam wachsen-

den. Langsam wachsende Neuroblastome haben im Allgemeinen eine gute Prognose; auch eine Spontanregression ist möglich (s. o.).⁶

So gesehen hätten die 337 mittels Screening diagnostizierten Fälle größtenteils ohnehin eine gute Prognose gehabt; dagegen waren die Säuglinge mit der potenziell ungünstigsten Prognose nicht in dieser Gruppe enthalten. Und natürlich wurden durch das Screening auch Neuroblastome entdeckt, die von allein wieder verschwunden wären. Ohne das Screening hätte niemand jemals von der Existenz dieser Tumoren erfahren; das Screening hat durch diese Überdiagnostizierung aus den betroffenen Säuglingen Patienten gemacht, die daraufhin unnötigen behandlungsbedingten Leiden ausgesetzt wurden.

Zudem hatte bei der Analyse der ermutigenden Ergebnisse aus kleinen Studien, die den Anlass zur Durchführung eines landesweiten Screenings in Japan gegeben hatten, anfangs die Überlebensdauer vom Zeitpunkt ab der Neuroblastomdiagnose im Mittelpunkt gestanden und nicht die Überlebensdauer seit der Geburt. Dies ist wichtig, weil eine frühere Diagnose nicht automatisch dazu führt, dass die Patienten länger leben – sie leben lediglich längere Zeit mit einem Krankheitsetikett. Anders ausgedrückt: Die Kinder scheinen ihre Diagnose länger zu «überleben», weil die «Krankheitsuhr» früher zu ticken beginnt (**Abb. 5**). Dies ist ein Beispiel für eine andere Art von systematischem Fehler (Bias), die man als «Leadtime-Bias» (Vorlaufzeit-Bias) bezeichnet und der sich dadurch ausschalten lässt, dass man die Ergebnisse nach Geburtsdatum der Kinder und nicht nach ihrem Alter bei der Diagnosestellung auswertet.

Anhand unverzerrter wissenschaftlicher Belege aus klinischen Studien, die unter Beteiligung von insgesamt drei Millionen Kindern in Kanada und Deutschland durchgeführt worden waren, konnten die Wissenschaftler keinerlei Nutzen für dieses Screening nachweisen, dafür stießen sie aber auf offensichtliche schädliche Auswirkungen.⁷ Dazu zählten ungerechtfertigte operative Eingriffe und Chemotherapien, die beide mit schwerwiegenden unerwünschten Wirkungen einhergehen können. Im Lichte dieser Evidenz wurde das Neuroblastom-Screening bei Säuglingen in Japan 2004 gestoppt.

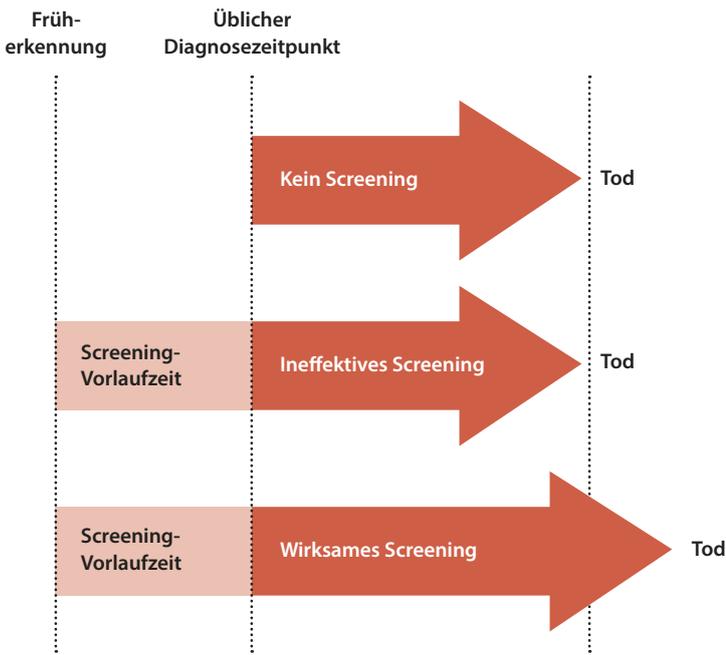


Abbildung 5: Länger leben mit dem Etikett «krank».

Glauben Sie bloß nicht, dass Früherkennung sich lohnt

«Das Neuroblastom-Screening zeigt anschaulich, wie leicht man dem Irrtum erliegen kann, dass Screeninguntersuchungen nützen, nur weil eine Krankheit dadurch frühzeitig entdeckt wird ... Die beiden Studien belegen, dass das Neuroblastom-Screening nicht nur nutzlos war, sondern darüber hinaus auch zu einer «Überdiagnostizierung» geführt hat und auch Tumoren identifiziert haben muss, die spontan wieder verschwunden wären. Beide Studien berichten über Kinder in der gescreenten Gruppe, die während der Behandlung unter schwerwiegenden Komplikationen litten ... Es bleibt zu hoffen, dass diese Erfahrungen Berücksichtigung finden, wenn die Einrichtung weiterer Screeningprogramme in Erwägung gezogen wird – beispielsweise ein Prostatakarzinom-Screening.»

Morris JK. Screening for neuroblastoma in children.
Journal of Medical Screening 2002; 9: 56.

Allerdings belegen die japanischen Ergebnisse, wie oben erwähnt, dass die gescreenten Säuglinge vom Zeitpunkt der Diagnosestellung an länger überlebten; die von der Geburt an gerechnete Überlebensdauer war nicht analysiert worden. Vor der Einführung des

Screeningprogramms im australischen Bundesstaat New South Wales wertete deshalb ein australischer Spezialist die japanischen Ergebnisse erneut aus, dieses Mal auf der Grundlage des Geburtsdatums der Säuglinge anstelle des Datums der Diagnosestellung: Seine Analyse ergab keinen Unterschied zwischen den Überlebensraten der gescreenten und der nicht gescreenten Säuglinge. Dieses Ergebnis überzeugte die Behörden in New South Wales, sodass sie das geplante Screeningprogramm nicht weiter verfolgten. Den Kindern blieben dadurch unnötige Leiden und dem Gesundheitswesen überflüssige Kosten erspart.

Nutzen und Schaden gegeneinander abwägen

Es gibt zahlreiche Beispiele für sinnvolle Screeninguntersuchungen. Am weitesten verbreitet ist bei Erwachsenen vermutlich die in der Grundversorgung routinemäßig durchgeführte Untersuchung auf Risikofaktoren für Herzerkrankungen und Schlaganfall. Vieles spricht dafür, dass Bluthochdruck, ein hoher Cholesterinspiegel und Rauchen das Risiko für diese Krankheiten erhöhen und dass man Herzinfarkten und Schlaganfällen vorbeugen kann, indem man Personen mit diesen Risikofaktoren identifiziert, berät und entsprechend behandelt.

Phenylketonurie-Screening: eindeutige Vorteile

Neugeborene werden routinemäßig auf Phenylketonurie (PKU), eine erbliche Stoffwechselstörung, untersucht. Babys mit PKU sind nicht in der Lage, die Aminosäure Phenylalanin abzubauen, die in alltäglichen Nahrungsmitteln wie Milch, Fleisch, Fisch und Eiern vorkommt. Bleibt die Krankheit unbehandelt, reichert sich das Phenylalanin im Blut an und führt zu schwerwiegenden irreversiblen Hirnschädigungen. Für den PKU-Test müssen nur wenige Tropfen Blut aus der Ferse des Säuglings entnommen und im Labor untersucht werden. Wenn dieser «Fersen-Pricktest» positiv ausfällt und die Diagnose durch weitere Untersuchungen bestätigt wird, werden die betroffenen Babys mit einer Spezialdiät behandelt, die ihnen zu einer normalen Entwicklung verhilft.

Screening auf Bauchaortenaneurysma: Einsatz mit Bedacht

Am anderen Ende des Altersspektrums kann das Screening auf ein Bauchaortenaneurysma von Vorteil sein. Die Aorta ist das wichtigste Blutgefäß im Körper (Hauptschlagader). Sie zieht vom Herzen durch die Brust (den Thorax) in den Bauchraum (das Abdomen). Bei manchen Menschen wird die Bauchaortenwand im Alter schwächer, sodass es allmählich zu einer Aussackung kommt – man spricht in diesem Fall von einem Aneurysma, einer Gefäßerkrankung, die nur selten Symptome verursacht und am häufigsten bei Männern ab 65 Jahren auftritt. Große Aneurysmen können ohne Vorwarnung schließlich platzen, sodass Blut in den Bauchraum fließt, was oftmals zum Tod führt.⁸

Solche Belege für die Häufigkeit von Aneurysmen bei älteren Männern können den Ausgangspunkt für die Einführung eines Screeningprogramms bilden. In Großbritannien beispielsweise wird Männern (aber nicht Frauen) ab dem 65. Geburtstag ein Ultraschallscreening angeboten. Ultraschallaufnahmen (Scans) können große Aneurysmen darstellen, sodass die betroffenen Männer fachärztlich beraten und (meist operativ) behandelt werden können. Männer mit kleineren Aneurysmen werden mithilfe weiterer Ultraschalluntersuchungen überwacht; Männer, deren Aorta keine Aussackungen zeigt, benötigen kein weiteres Screening. Von entscheidender Bedeutung ist sowohl die Screening- als auch die Operationsqualität. Da die Aneurysmachirurgie einen größeren Eingriff darstellt, würde bei hohen Komplikationsraten mehr Männern geschadet als geholfen werden.

Brustkrebs-Screening: etabliert, aber nach wie vor umstritten

Da die routinemäßige Brustuntersuchung mittels Mammographie-Screening in vielen Ländern etabliert ist, könnte man wohl annehmen, es lägen fundierte Belege dafür vor, dass die Vorteile der Mammographie ihre Nachteile überwiegen. Ein US-amerikanischer Gesundheitsexperte brachte dies 2010 mit folgenden Worten zum Ausdruck:

Kein Screeningtest ist jemals gründlicher untersucht worden. In den vergangenen 50 Jahren haben mehr als 600 000 Frauen an zehn randomisierten Studien mit jeweils einer Nachbeobachtung von ungefähr zehn Jahren teilgenommen.» Aber, so merkte er weiter an: «Angesichts dieser außergewöhnlichen Forschungsanstrengungen entbehrt es doch nicht einer gewissen Ironie, dass das Mammographie-Screening unter Medizinern nach wie vor eines der strittigsten Themen darstellt.»⁹

Woran liegt das? Ein wesentlicher Grund ist, dass es den Frauen von Screeninganbietern und Patientengruppen als vernünftig «verkauft» wird. In der Broschüre, die Frauen mit der Einladung zum Brustkrebs-Screening erhalten, werden die Vorteile hervorgehoben, während über die Nachteile, Grenzen und Folgen der Mammographie hinweggegangen wird.¹⁰ Dabei führt die Mammographie nicht nur zu einer frühzeitigen Diagnose, sondern auch, genauso wie beim Prostatakarzinom (s. unten), dazu, dass Krebsarten diagnostiziert werden, die den Patienten zu Lebzeiten niemals gesundheitliche Probleme bereitet hätten. Und natürlich muss man unweigerlich auch mit falsch positiven Ergebnissen rechnen, d.h. der diagnostische Test kann positiv ausfallen, obwohl die Erkrankung gar nicht vorliegt.

Die zuverlässigsten Belege stammen aus der systematischen Überprüfung der Ergebnisse klinischer Studien, in denen die Frauen nach dem Zufallsprinzip (randomisiert) einem Screening oder keinem Screening zugeteilt wurden. Und die Ergebnisse sind sehr aufschlussreich. Wie sie zeigen, müssen 2 000 Frauen regelmäßig über zehn Jahre gescreent werden, damit eine von ihnen insofern von diesem Screening profitiert, dass sie nicht an Brustkrebs verstirbt. Gleichzeitig werden aber aus zehn gesunden Frauen infolge des Screenings «Krebspatientinnen», die unnötigerweise behandelt werden. Bei diesen Frauen hat die Mammographie tatsächlich Gewebeveränderungen nachgewiesen, die aber so langsam (oder sogar gar nicht) wachsen, dass sie sich niemals zu einem echten Karzinom entwickelt hätten. Diesen gesunden Frauen wird dann ein Teil ihrer Brust oder sogar die ganze Brust entfernt, und danach erhalten sie häufig eine Strahlen- und manchmal auch noch eine Chemotherapie.¹¹

Des Weiteren kommt es bei 200 von 2000 gescreenten Frauen zu einem falschen Alarm durch falsch positive Befunde, die weiter abgeklärt werden müssen: Die psychische Belastung, unter der die Betroffenen stehen, bis und auch noch nachdem sie wissen, ob sie an Krebs erkrankt sind, kann sehr schwerwiegend sein. Darüber hinaus wird Frauen die Mammographie oft zusammen mit Empfehlungen zur Selbstuntersuchung der Brust oder zu Brustbewusstsein angeraten – beides Methoden, die nachweislich ebenfalls mehr Schaden anrichten als nutzen.¹²

Ein britischer Gesundheitsexperte hat darauf hingewiesen, dass das Potenzial für einen individuellen Nutzen des Mammographie-Screenings sehr gering ist. Er merkte dazu an: «Das ist nicht allen klar. Zum Teil ist dies auf die Verschleierung seitens der Organisatoren von Mammographiediensten zurückzuführen, die der Meinung sind, dass sie die Vorteile herausstreichen müssen, um eine passable Compliance [mit dem Screening] sicherzustellen.» In einer Bewertung der im Jahr 2010 verfügbaren Evidenz erklärte er: «Mammographie rettet Leben – dies trifft in stärkerem Maße für ältere Frauen zu; aber es birgt auch gewisse Nachteile.» Mit den Nachteilen meint er die Überdiagnostizierung und die falsch positiven Befunde. Kritisch bemerkte er, dass die unvoreingenommene vollständige Untersuchung aller Einzelergebnisse aus neueren Screeningstudien noch aussteht.¹³ Während wir auf eine solche unparteiische Evaluation noch warten, erhalten Frauen auch weiterhin Einladungen zum Mammographie-Screening. Zumindest müssen ihnen dann aber hinreichend ausgewogene Informationen zur Verfügung gestellt werden, aufgrund derer sie (zusammen mit ihrer Familie und, wenn sie das wünschen, ihrem Arzt) entscheiden können, ob sie an diesem Screening teilnehmen möchten – oder nicht.

Prostatakarzinom-Screening: deutliche Nachteile bei unklaren Vorteilen

Das Prostatakarzinom, die weltweit zweithäufigste Krebsart bei Männern¹⁴, lässt sich grob in zwei Typen unterteilen. Manche Männer leiden an einer aggressiven Form der Krankheit; diese gefährli-

chen Karzinome breiten sich rasch aus und gehen mit einer hohen Sterblichkeit (Mortalität) einher. Viele der Erkrankten haben aber ein langsam wachsendes Karzinom, das für sie zu Lebzeiten niemals zu einer Bedrohung führen würde. Im Idealfall würde man bei einer Screeninguntersuchung die gefährlichen Karzinome entdecken (und hoffen, dass sie behandelbar sind), nicht aber die langsam wachsenden. Der Grund ist, dass die Behandlung beider Formen des Prostatakarzinoms belastende Nebenwirkungen wie Inkontinenz und Impotenz mit sich bringt – ein hoher Preis also, den man auch dann bezahlen muss, wenn das Karzinom eigentlich gar keine Probleme verursachen würde.¹⁵

Die Mehrzahl der an einem Prostatakarzinom erkrankten Männer weist erhöhte Blutkonzentrationen einer Substanz auf, die man prostataspezifisches Antigen (PSA) nennt. Allerdings gibt es keinen eindeutigen Grenzwert, mit dessen Hilfe man Männer mit Prostatakrebs von Männern ohne Prostatakrebs unterscheiden könnte¹⁶, und nicht weniger als einer von fünf Männern mit einem behandlungsbedürftigen Prostatakarzinom hat normale PSA-Spiegel. Zudem ist das PSA, entgegen seinem Namen, alles andere als «spezifisch», denn auch gutartige Prostatatumoren, Infektionen und sogar einige frei verkäufliche Schmerzmittel können zu einem Anstieg der PSA-Spiegel führen. Allein aus diesen Gründen weist das PSA als Screeningtest gravierende Schwächen auf.

Und doch wird die routinemäßige Durchführung von PSA-Tests bei gesunden Männern von Ärzten, Patientengruppen und Unternehmen, die diese Tests verkaufen, leidenschaftlich befürwortet und ist in vielen Ländern auch verbreitet eingeführt worden. Die das PSA-Screening befürwortende Lobby hat sich besonders lautstark in den USA zu Wort gemeldet, wo alljährlich schätzungsweise 30 Millionen Männer in dem Glauben getestet werden, dies sei sinnvoll. Welche Belege gibt es also für die Behauptung, dass die Früherkennung von Prostatakrebs mittels PSA-Screening das Behandlungsergebnis der Betroffenen verbessert, und was weiß man über die mit diesem Test verbundenen Nachteile?

Inzwischen ist auch hochwertige Evidenz für die Vor- und Nachteile des PSA-Screenings verfügbar. Im Jahr 2010 wurden die Ergebnisse aller relevanten Studien systematisch ausgewertet. Diese Be-

Überdiagnostizierung des Prostatakarzinoms

«Das Prostatakarzinom gilt als Paradebeispiel für Überdiagnostizierung. Das bedeutet nicht, dass es nicht Männer gibt, die durch die Früherkennung vor einem vorzeitigen, durch Prostatakrebs bedingten Tod bewahrt werden. Aber ... es gibt kaum eine Möglichkeit, vorab festzustellen, welche Männer von einem Screening profitieren und welche sich unnötigerweise einer Therapie unterziehen – oftmals mit einschneidenden unerwünschten Folgen für ihr Leben. Das wesentliche Problem besteht darin, dass wir durch das Prostatakrebs-Screening und die Untersuchungen sehr viel mehr Prostatakarzinome entdecken als jemals zuvor, und so merkwürdig das auch anmuten mag: Viele dieser Karzinome würden niemals zu einer Lebensbedrohung werden. In der Vergangenheit hätten diese Männer niemals erfahren, dass sie an Prostatakarzinom erkrankt sind. Sie wären nicht wegen ihres Prostatakarzinoms, sondern mit ihm, d. h. an etwas anderem, verstorben. Durch den Nachweis all dieser indolenten (schmerzlosen) Prostatakarzinome erhalten heutzutage mehr Männer die Diagnose Prostatakrebs als jemals zuvor – daher auch der Begriff «Überdiagnose». Dies ist das Grunddilemma, vor dem jeder Mann steht, der einen solchen Screeningtest in Erwägung zieht.»

Chapman S, Barratt A, Stockler M. *Let sleeping dogs lie? What men should know before getting tested for prostate cancer*. Sydney: Sydney University Press, 2010: S. 25

Der Entdecker des PSA meldet sich zu Wort

«Die Popularität des PSA-Tests hat ein immens kostspieliges Public-Health-Desaster ausgelöst. Dies ist ein Punkt, dessen ich mir schmerzlich bewusst bin – denn ich war es, der das PSA 1970 entdeckt hat. ... Amerikaner geben enorm viel Geld für Untersuchungen zur Früherkennung von Prostatakrebs aus. Die jährlichen Kosten für das PSA-Screening belaufen sich auf mindestens 3 Milliarden Dollar; ein Großteil dieser Summe wird von der öffentlichen Krankenversicherung *Medicare* und der Veteranen-Versorgungsbehörde *Veterans Benefit Administration* aufgebracht. Das Prostatakarzinom mag zwar in der Presse eine hohe Aufmerksamkeit genießen, aber sehen wir uns einmal die Zahlen dazu an: Bei amerikanischen Männern liegt die Lebenszeitwahrscheinlichkeit, mit einem Prostatakarzinom diagnostiziert zu werden, bei 16%; die Wahrscheinlichkeit, daran zu versterben, beträgt dagegen nur 3%. Das liegt daran, dass die Mehrzahl der Prostatakarzinome langsam wächst. Mit anderen Worten: Männer, die das Glück haben, ein hohes Alter zu erreichen, sterben sehr viel eher mit als wegen Prostatakrebs. Und selbst dann ist der Test kaum effizienter, als eine Münze zu werfen. Seit vielen Jahren versuche ich nun schon klarzustellen, dass der PSA-Test ein Prostatakarzinom nicht nachweisen kann. Was aber noch viel wichtiger ist: Er kann nicht zwischen den beiden Typen des Prostatakarzinoms unterscheiden – dem, das tödlich verläuft, und dem, das nicht tödlich ist.»

Ablin RJ. The great prostate mistake. *New York Times*, 10. März 2010.

wertung ergab, dass das PSA-Screening die Wahrscheinlichkeit einer Prostatakrebsdiagnose erhöhte (was ja auch zu erwarten ist), dass es aber keine wissenschaftlichen Belege für einen Einfluss auf die krebsbedingte Sterblichkeitsrate oder die Gesamtsterblichkeit gibt.¹⁷

Wendet sich das Blatt also nun gegen das PSA-Screening? Ginge es nach Richard Ablin, dem Entdecker des PSA, dann würde genau das eintreten, was er bereits seit vielen Jahren propagiert. 2010 schrieb er dazu in einem Beitrag:

Ich hätte mir niemals träumen lassen, dass meine vier Jahrzehnte zurückreichende Entdeckung zu einem solchen profitgesteuerten Desaster im Gesundheitswesen führen würde. Die Ärzteschaft muss sich der Realität stellen und diesen unangemessenen Einsatz des PSA-Screenings stoppen. Das würde Milliarden Dollar einsparen helfen und Millionen Männern unnötige, kräftezehrende Behandlungen ersparen.

Zumindest sollte jeder Mann über die Grenzen des Tests und seine möglichen unerwünschten Folgen beraten werden, bevor er sich einem PSA-Test unterzieht. Eine Gruppe von Experten hat darauf hingewiesen, dass «[Männer] darüber aufgeklärt werden sollten, dass der Test [ihnen] keine Auskunft darüber geben kann, ob sie an einer lebensbedrohlichen Krebserkrankung leiden, dass er sie aber durch einen Dschungel von Tests und Behandlungen führen könnte, auf die sie vielleicht lieber verzichtet hätten.»¹⁸

Lungenkrebs-Screening: früh, aber nicht früh genug?

Durch ein Screening kann die Krankheit früher diagnostiziert werden, aber sie wird nicht immer früh genug nachgewiesen, um einen Unterschied zu bewirken (**Abb. 6**).

Einige Krebsarten wie beispielsweise das Lungenkarzinom breiten sich im Körper aus, bevor der Patient Symptome zeigt und das Vorhandensein der Krebserkrankung durch Untersuchungen nachweisbar ist. Versuche, Lungenkrebs mithilfe von Thorax-Röntgen nachzuweisen, veranschaulichen dieses Problem (s. Stadium B in **Abb. 6**). In den 1970er-Jahren ergaben mehrere großangelegte Stu-

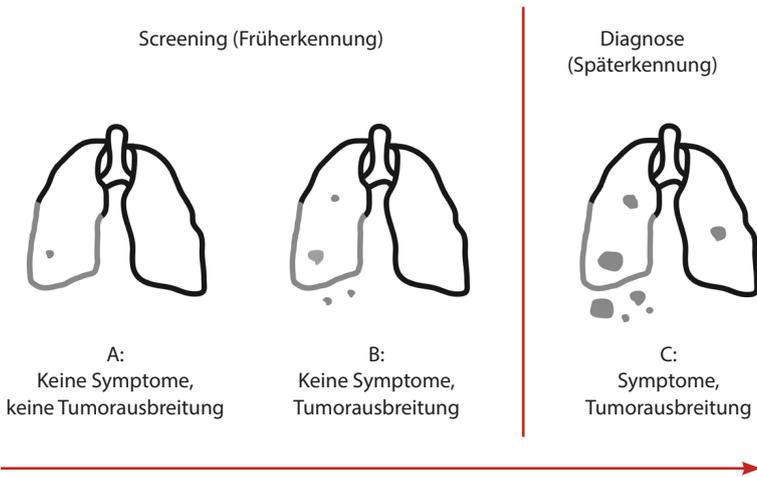


Abbildung 6: Wachstum und Ausbreitung des Lungenkarzinoms bei starken Rauchern.

dien mit Starkrauchern, dass die Krebserkrankungen zwar früher entdeckt wurden, es ließ sich aber nicht nachweisen, dass dadurch auch die Anzahl der lungenkrebsbedingten Todesfälle abnahm. Die in den Röntgenaufnahmen sichtbaren Lungenkarzinome hatten sich bereits über die Lunge hinaus ausgebreitet. Somit lebten diese Patienten länger mit ihrer Krebsdiagnose und wurden auch früher behandelt, doch es gab keinen Anhaltspunkt dafür, dass dieser Umstand sich auf ihre Lebenserwartung auswirkte.

In einer neueren, großen randomisierten Studie, an der 53 000 aktuelle und ehemalige starke Raucher teilnahmen, wurde das Röntgen-Thorax-Screening mit einem besonderen Verfahren der Computertomographie (CT), der sog. Spiral-CT, verglichen. Zwei Patientengruppen wurden drei jährlichen Screeninguntersuchungen zugeteilt. Mithilfe der Spiral-CT konnte das Lungenkarzinom in einem noch früheren Stadium diagnostiziert werden als mit dem Röntgen-Thorax, und bei einem kleinen Anteil der Patienten wurde die Diagnose früh genug gestellt (Stadium A in Abb. 6), sodass eine nutzbringende Behandlung eingeleitet werden konnte (346 lungenkrebsbedingte Todesfälle in der Gruppe mit Spiral-CT gegenüber

Wie Screenings verkauft werden

«Screening zu verkaufen kann ziemlich einfach sein: Angst verbreiten, indem man das Risiko übertreibt. Hoffnung schüren, indem man den Screeningnutzen übertreibt. Und auf keinen Fall die Nachteile erwähnen. Besonders einfach ist dies beim Thema Krebs – denn keine Diagnose ist gefürchteter. Und wir alle kennen das Mantra: Früherkennung ist der beste Schutz. Wenn Sie Zweifel äußern, dann könnte es Ihnen passieren, dass man Ihnen rät, Ihren Geisteszustand untersuchen zu lassen: «Wenn Sie eine Frau und über 35 Jahre alt sind, sollten Sie sich zur Mammographie anmelden. Es sei denn, Sie sind noch immer nicht davon überzeugt, wie wichtig das ist. Dann könnte es aber sein, dass Sie nicht nur Ihre Brust untersuchen lassen sollten.» (Plakat der *American Cancer Society*).

Botschaften, die uns ein Screening verkaufen wollen, sind allgegenwärtig. In den Nachrichten hören wir regelmäßig Geschichten von berühmten Personen, die uns versichern, dass sie der Früherkennung einer Krebserkrankung ihr Leben zu verdanken haben. Selten sind dagegen Geschichten über Menschen zu lesen, die durch Überdiagnostizierung und Überbehandlung Schaden genommen haben. In gängigen Zeitschriften lesen wir emotional aufgeladene, aber vollkommen unrepräsentative Geschichten über junge Frauen mit Brustkrebs und ihre Angst davor, daran zu sterben und ihre kleinen Kinder zurücklassen zu müssen. Medizinische Zentren nutzen das Screening als Geschäftsstrategie, indem sie, um Patienten anzulocken, kostenlose Untersuchungen anbieten. Öffentliche Werbebotschaften – wie der oben zitierte Slogan der *American Cancer Society* – sprechen für sich selbst.»

Woloshin S, Schwartz LM. Numbers needed to decide. *Journal of the National Cancer Institute* 2009; 101: 1163–1165.

425 in der Gruppe mit Röntgen-Thorax). Doch ging dieses günstige Behandlungsergebnis zulasten eines großen Anteils von Studienteilnehmern, bei denen die Lungenkrebsdiagnose irrtümlich gestellt wurde. Insgesamt verstarben im Laufe der achtjährigen Nachbeobachtungszeit pro 1 000 starke Raucher, die jedes Jahr drei Röntgen- oder CT-Untersuchungen erhalten hatten, drei Patienten weniger an Lungenkrebs. Aber 13 starben trotz Früherkennung an Lungenkrebs, und 233 erhielten einen falsch positiven Befund, der eine weitere Diagnostik erforderlich machte.¹⁹

Genetische Tests: manchmal sinnvoll, oftmals unzuverlässig

Vor nicht allzu langer Zeit waren «genetische Tests» mehr oder weniger auf die generell seltenen monogenen Erkrankungen beschränkt – beispielsweise auf die in der Kindheit einsetzende und zu Muskelschwund führende Duchenne-Muskeldystrophie oder die Huntington-Krankheit, eine fortschreitende Erkrankung des Nervensystems, die normalerweise im Alter zwischen 35 und 40 Jahren auftritt. Um solche Erkrankungen zu diagnostizieren, werden genetische Tests durchgeführt. Man kann sie im Hinblick auf die Familienplanung aber auch zum Screening gesunder Menschen einsetzen, bei denen die Familienanamnese vermuten lässt, dass die fragliche Krankheit bei ihnen mit einer überdurchschnittlich hohen Wahrscheinlichkeit auftreten könnte.

Die Mehrzahl der erblichen Erkrankungen lässt sich allerdings nicht auf ein einzelnes defektes Gen zurückführen. In der Regel ist der Ausbruch dieser Krankheiten davon abhängig, inwieweit in mehreren Genen vorhandene Risikovarianten miteinander in Wechselwirkung treten und wie diese mit Umweltfaktoren zusammenwirken. Nur wenn es zu einer «kritischen» Kombination von genetischen Risikovarianten und Umweltfaktoren kommt, bricht die Krankheit aus.¹

Obwohl es durchaus kompliziert ist, die Mehrzahl der Krankheiten auf abnormale (aberrante) Gene zurückzuführen, preisen die Medien und Befürworter von *Direct-to-Consumer*-(DTC-) Gentests (die dem Verbraucher direkt, ohne Umweg über eine medizinische Beratung, angeboten werden) den angeblichen Wert und die Einfachheit der Erstellung eines solchen genetischen Risikoprofils an. Zur DNA-Analyse braucht man lediglich eine Speichelprobe an ein Unternehmen einzusenden, das solche Tests gegen Bezahlung durchführt und Ihnen dann Ihr genetisches Profil zusendet. Doch die Informationen, die Sie so erhalten, werden Ihnen – bzw. Ihrem Arzt – kaum helfen, vernünftige Vorhersagen über Ihr Erkrankungsrisiko zu treffen, ganz zu schweigen davon, dass sie Ihnen auch nicht sagen, was man im Fall der Fälle – wenn überhaupt – tun könnte. Dieser «Do-it-yourself»-Ansatz erfüllt eindeutig nicht die Kriterien für ei-

nen sinnvollen Screeningtest (s. unten). Stattdessen kann das Ergebnis Sie aber sehr wohl beunruhigen und Ihre Entscheidungsfindung erschweren; zudem kann es auch umfangreichere Folgen haben – beispielsweise für Ihre Angehörigen. Um es mit den Worten Ray Moynihans, eines australischen Gesundheitsjournalisten, zu sagen:

Für alle, die über die schleichende Medikalisation unseres Lebens besorgt sind, eröffnet der Markt für genetische Tests mit Sicherheit eine neue Front, an der durch eine anscheinend harmlose Technologie aus gesunden Menschen ängstliche Patienten werden können, deren Menschsein durch multiple genetische Prädispositionen für Krankheit und vorzeitigen Tod damit neu definiert wird.²⁰

Pokern Sie nicht mit Ihren Genen

«Auf der Grundlage des Wissens um eine einzelne (oder auch einige wenige) Genvariante(n) zu handeln, ist ungefähr so, als würden Sie Ihr gesamtes Geld auf ein Pokerblatt setzen, obwohl Sie erst eine Karte gesehen haben. Sie wissen weder, was für ein Blatt an genetischen Faktoren Ihnen ausgeteilt wurde, noch wissen Sie, welche Auswirkungen Ihre Umwelt haben wird, und hier haben wir es nicht nur mit fünf Karten, sondern mit mehr als 20 000 Genen und vielen tausend Umweltfaktoren zu tun. Zudem kann die Wirkung eines Gens durch den Einfluss Ihrer Lebensführung, Ihrer Familiengeschichte oder durch das Vorhandensein anderer, schützender Gene aufgehoben werden. Viele von uns tragen defekte Gene in sich, ohne dass es jemals zum Auftreten einer Krankheit kommt.»

Sense About Science. Making sense of testing: a guide to why scans and other health tests for well people aren't always a good idea. London: Sense About Science 2008, S. 7. Erhältlich unter www.senseaboutscience.org

Was man mit Screening erreichen will und warum Evidenz so wichtig ist

Wie die bisher angeführten Beispiele zeigen, lohnt es sich, einen Moment innezuhalten, bevor man vorschnell für ein flächendeckendes Screening eintritt. Denken wir noch einmal über die wichtigsten Eigenschaften von Screeningprogrammen nach und rufen uns in Erinnerung, was sie eigentlich bezwecken sollen. Die Menschen, denen ein Screening angeboten wird, zeigen weder Symptome noch Zeichen der Krankheit, auf die sie getestet werden sollen,

oder sie haben diese noch nicht bemerkt – sie haben wegen der fraglichen Erkrankung noch keine ärztliche Hilfe in Anspruch genommen. Das Screening von Einzelpersonen oder von Bevölkerungsgruppen verfolgt den Zweck, das Sterblichkeitsrisiko oder das Risiko für eine spätere Erkrankung zu reduzieren, indem man ihnen eine Untersuchung anbietet, mit der sich diejenigen identifizieren lassen, die von einer Behandlung profitieren könnten.^{1,21}

Die grundlegenden Kriterien für die Bewertung des Nutzens von Screeninguntersuchungen wurden 1968 in einem Bericht der Weltgesundheitsorganisation niedergelegt²² und stetig weiter verfeinert, um unserer modernen Gesundheitsversorgung Rechnung zu tragen. Menschen, die zu einem Screening eingeladen werden, benötigen ausreichende und ausgewogene Informationen über das Untersuchungsangebot – darunter seine möglichen Nachteile, Konsequenzen und Grenzen wie auch seine potenziellen Vorteile – sodass eine informierte Entscheidung getroffen werden kann. Die wichtigsten Aspekte lassen sich im Wesentlichen wie folgt zusammenfassen: Ein Screening sollte nur durchgeführt werden,

- wenn die gesuchte Erkrankung ein relevantes allgemeines Gesundheitsproblem darstellt (wenn es sich beispielsweise um eine schwere Krankheit handelt und/oder wenn eine große Anzahl von Menschen davon betroffen ist)
- wenn die Krankheit in einem Frühstadium erkennbar ist
- wenn es einen validen und zuverlässigen Test für die betreffende Erkrankung gibt, der für die Menschen, denen das Screening angeboten wird, akzeptabel ist
- wenn eine wirksame und akzeptable Behandlung für diese Krankheit zur Verfügung steht, sodass ein Screening die Prognose wahrscheinlich günstig beeinflussen kann
- wenn es sich um eine hochwertige Screeningmaßnahme handelt, die in dem Bereich, in dem sie angeboten wird, kosteneffektiv ist
- wenn die Informationen, die den Betroffenen zur Verfügung gestellt werden, unvoreingenommen sind, auf soliden wissenschaftlichen Daten basieren und die möglichen Nachteile (z.B.

Überdiagnosen, die zu Überbehandlung führen) wie auch die potenziellen Vorteile klar dargestellt sind

- wenn die Einladung zum Screening nicht mit Zwang verbunden ist – d. h. wenn sie einen Hinweis darauf enthält, dass es gerechtfertigt sein kann, die Einladung nicht anzunehmen
- wenn die Wahrscheinlichkeit eines körperlichen oder seelischen Schadens für den Empfänger des Screeningangebots wahrscheinlich geringer ist als seine Nutzenwahrscheinlichkeit
- wenn ausreichende Möglichkeiten für die Diagnose und Therapie der durch das Screening nachgewiesenen gesundheitlichen Störungen verfügbar sind.

Diese Kriterien unterstreichen noch einmal unsere am Anfang dieses Kapitels formulierte Botschaft: dass nämlich alle Entscheidungen über die Einführung eines Screeningprogramms auf der Grundlage hochwertiger Evidenz getroffen werden sollten, und zwar nicht nur bezüglich seiner Wirksamkeit, sondern auch hinsichtlich seines Schadenspotenzials.

Ist überhaupt jemand normal?

Untersuchungen mittels Ganzkörper-CT

Von Privatkliniken wird zur Untersuchung von Schädel, Hals, Brust, Bauch und Becken u. a. die Ganzkörper-Computertomographie (CT) angeboten. Sie wird dem Konsumenten direkt angeboten und normalerweise ohne Rücksprache mit dem Hausarzt der Betroffenen durchgeführt. Ganzkörper-Scans werden häufig als Möglichkeit beworben, einer potenziellen Erkrankung immer einen Schritt voraus zu sein, davon ausgehend, dass ein «normaler» Befund beruhigend wirkt. Doch die Scans sind nicht nur teuer, sondern es ist auch nicht belegt, dass sie bei Menschen ohne Krankheitssymptome oder Krankheitszeichen überhaupt einen gesundheitlichen Gesamtnutzen haben.

Der Screening-Zirkus

2009 erfuhr ein frisch pensionierter Professor für Neurologie, der sich schon längere Zeit für das Thema Schlaganfallprävention interessierte, von seinen Nachbarn, dass sie eine Broschüre erhalten hatten, mit der sie zu einem Screening für Schlaganfall und andere Komplikationen kardiovaskulärer Erkrankungen eingeladen wurden. Die von einem Gefäßscreening-Unternehmen stammende Broschüre lud sie dazu ein, sich für eine Reihe von Untersuchungen (zum Preis von € 170) in die örtliche Kirche zu begeben. Neugierig geworden – nicht zuletzt, weil einige der Informationen in der Broschüre faktisch falsch waren –, beschloss er, sich selbst dorthin zu begeben.

«Als erstes stand eine Untersuchung auf Aortenaneurysma [Erweiterung der Hauptarterie, die das Blut vom Herzen weg befördert] mittels Ultraschall auf dem Programm. Sie wurde von einer Frau durchgeführt, die sich nicht auf ein Gespräch über die Folgen eines etwaigen Aneurysmabefunds einlassen wollte. Als Nächstes kamen Blutdruckmessungen am Knöchel und am Arm an die Reihe, «wegen eventueller Probleme mit meinem Kreislauf... gefolgt von einem kleinen Bonus, der nichts mit meinen Gefäßen zu tun hatte: ein Osteoporose-Screening meines Sprunggelenks. Dann gab es noch eine ... Elektrokardiographie zum Nachweis von «Problemen mit den beiden oberen Kammern meines Herzens» ... Das Schlusslicht bildete ein Ultraschall der Karotis [Halsarterie], um zu prüfen, «ob sich Plaques [Ablagerungen] gebildet haben». Auf meine Frage, welche Folgen das haben könne, antworteten sie mir, dass sich Blutgerinnsel bilden und zu einem Schlaganfall führen könnten. Als ich nachhakte, wie man das behandeln könne, bekam ich vage Bemerkungen über blutverdünnende Medikamente zu hören, aber nichts über eine operative Therapie. Erst als ich direkt fragte, ob vielleicht auch eine Operation möglich sei, wurde diese Frage tatsächlich bejaht. «Könnte das gefährlich sein?», erkundigte ich mich gutgläubig. Die Antwort lautete, dass die Risiken von einer gründlichen diagnostischen Abklärung abhingen, die meine Hausärztin durchführen müsse und mit der ich etwaige auffällige Befunde dieser Untersuchungen besprechen sollte.

All diese Untersuchungen (ausgenommen das Aortenaneurysma-Screening) wurden ohne Sichtschutz durchgeführt ... Es schien kein Arzt anwesend zu sein, und das Team ließ keine Absicht erkennen bzw. war nicht bereit, mit mir die Folgen falsch positiver oder falsch negativer Befunde, die prognostischen Konsequenzen echter Auffälligkeiten oder die Risiken und Vorteile etwaiger Behandlungen zu erörtern.

Es war bloß ein Screening, nicht mehr und nicht weniger, in rein kommerzieller Absicht angeboten: Die Ergebnisse würden mir innerhalb von 21 Werktagen ins Haus flattern, und es bliebe meiner Hausärztin überlassen, sich um die emotionalen und körperlichen Folgen etwaiger richtiger oder falscher auffälliger Befunde zu kümmern, auch wenn sie diese Untersuchungen nicht selbst angefordert hatte. ... Bei anfälligen Menschen weckt dieser ganze Screening-Zirkus wahrscheinlich Angst, ohne dass darüber gesprochen wird, was passiert, wenn Auffälligkeiten nachgewiesen werden, und ohne dass auch nur die geringste Verantwortung für die Folgen übernommen wird.»

Warlow C. The new religion: screening at your parish church.
BMJ 2009; 338: b1940

Zudem gehen diese Untersuchungen mit einer nicht unerheblichen Strahlenbelastung einher: etwa 400-mal mehr als beim Röntgen-Thorax, d.h. so viel mehr, dass das britische Komitee zur Untersuchung von Strahlenrisiken durch die Umwelt (COMARE) 2007 «Dienstleistern», die symptomfreien (asymptomatischen) Personen Ganzkörper-CT-Untersuchungen anboten, dringend empfahl, davon Abstand zu nehmen.

Nach Beratungen kündigte die britische Regierung 2010 den Erlass strengerer Regeln für den Einsatz von Ganzkörper-Scans an. Auch die US-amerikanische *Food and Drug Administration* machte die Öffentlichkeit darauf aufmerksam, dass diese Scans bei Gesunden keinen nachweislichen Nutzen haben und kommentierte das mit folgenden Worten:

Viele Menschen sind sich nicht darüber im Klaren, dass eine Untersuchung mittels Ganzkörper-CT ihnen nicht unbedingt den erhofften «Seelenfrieden» verschafft oder die Informationen liefert, die ihnen helfen würden, einem gesundheitlichen Problem vorzubeugen. So kann es beispielsweise durchaus vorkommen, dass ein auffälliger Befund nichts Ernstes bedeutet und dass ein normaler Befund ungenau ist.^{23, 24, 25}

Einen Mittelweg finden

Einen Mittelweg zwischen der übereifrigen Suche nach Krankheiten und dem Versagen bei der Identifizierung derjenigen Menschen zu finden, die von einer Früherkennung profitieren könnten, ist niemals leicht und führt unweigerlich zu unpopulären Entscheidungen. Alle Gesundheitssysteme müssen sparsam mit ihren Ressourcen umgehen, wenn die ganze Bevölkerung davon profitieren soll. Dieses grundlegende Prinzip bedeutet mit Sicherheit, dass sich Screeningprogramme bei ihrer Einführung nicht nur auf solide Evidenz stützen können müssen, sondern dass man sie auch laufend überwachen muss, um zu kontrollieren, ob sie auch noch dann sinnvoll sind, wenn neue wissenschaftliche Daten vorliegen und sich die Umstände vielleicht geändert haben. Ernsthaft zu prüfen ist die Frage, ob Screeningprogramme großen Teilen der Bevölkerung angeboten oder stärker auf Personen mit einem erhöhten Erkrankungsrisiko zugeschnitten werden sollten.

- Eine frühere Diagnose führt nicht zwangsläufig zu besseren Behandlungsergebnissen; manchmal verschlimmert sie die Lage.
- Screeningprogramme sollten nur eingeführt werden, wenn es fundierte wissenschaftliche Belege für ihre Wirksamkeit gibt.
- Manchmal ist es besser, auf die Einführung eines Screeningprogramms zu verzichten.
- Menschen, die zu einer Screeninguntersuchung eingeladen werden, brauchen ausgewogene Informationen.
- Die Vorzüge eines Screenings werden oftmals zu sehr angepriesen.
- Die Nachteile eines Screenings werden häufig heruntergespielt oder ignoriert.
- Eine gute Kommunikation über die Vor- und Nachteile sowie die Risiken eines Screenings ist unerlässlich.

5

Wie man mit Unsicherheit im Hinblick auf Therapieeffekte umgeht

In diesem Kapitel wollen wir uns mit den Unsicherheiten befassen, mit denen die angeblichen Wirkungen neuer oder auch bereits bekannter Behandlungen fast unweigerlich behaftet sind. So würden wohl nur wenige Menschen den routinemäßigen Einsatz von Sauerstoff bei Herzinfarkt-Patienten infrage stellen. Jedoch gibt es keine soliden Daten dafür, dass diese Maßnahme tatsächlich hilft, und sogar einige Anhaltspunkte dafür, dass sie schaden könnte. Diese Unsicherheit wurde niemals wirklich angemessen untersucht¹, und auch viele andere Behandlungseffekte sind umstritten.

Dramatische Behandlungseffekte: selten und leicht erkennbar

Nur selten ist die Datenlage so eindeutig, dass alle Zweifel an der Wirksamkeit einer Therapie ausgeräumt werden können.² In solchen Fällen fällt der Therapieeffekt dann oftmals dramatisch aus und setzt unmittelbar ein. Nehmen wir z.B. das Kammerflimmern, eine Herzrhythmusstörung, bei der die Muskelkontraktion in den Ventrikeln (Herzkammern) vollkommen ungeordnet verläuft. Kammerflimmern ist ein medizinischer Notfall: Der Tod kann innerhalb von Minuten eintreten. Um den normalen Herzrhythmus wiederherzustellen, wird aus einem Defibrillator ein direkter elektrischer Stromimpuls an das Herz abgegeben. Ist diese Maßnahme erfolgreich, kann man den Effekt praktisch sofort sehen.

Als weitere dramatische Effekte (s.a. Kap. 6, S. 114) sind das Absaugen von Eiter aus Abszessen zur Schmerzlinderung, die Bluttransfusion bei Schock infolge schwerer Blutungen sowie die Gabe von Insulin bei Diabetes zu nennen. Bis in die 1920er-Jahre wurden Diabetiker nicht alt, mussten viel leiden und siechten mit unkontrollierbar hohen Blutzuckerspiegeln dahin. Sehr schnell kam es deshalb nach den ersten positiven Ergebnissen von Tierversuchen zur Anwendung von Insulin beim Menschen – mit herausragendem Erfolg: Das Ansprechen der Betroffenen auf diese Therapie galt damals fast als ein Wunder. Ein weiteres Beispiel aus dieser Zeit war die Gabe von Leber – die sich später als Quelle von Vitamin B₁₂ herausstellte – zur Behandlung von Patienten mit perniziöser Anämie. Bei dieser damals tödlichen Form der Blutarmut (Anämie) fiel die Anzahl der Erythrozyten (roten Blutkörperchen) allmählich auf katastrophal niedrige Werte ab, was den Patienten eine geisterhafte Blässe verlieh und sie stark schwächte. Als man diesen Patienten Leberextrakt gab, erholten sie sich rasch. Heutzutage wird Vitamin B₁₂ bei dieser Form von Anämie routinemäßig verordnet.

Auch zu Beginn dieses Jahrhunderts gab es ähnlich dramatische Ergebnisse zu beobachten, die durch einige Beispiele veranschaulicht werden sollen.

Laserbehandlung von Feuermalen

Die auch als «Portweinflecken» bezeichneten Geburts- oder Feuermale entstehen durch dauerhaft und krankhaft erweiterte Blutgefäße in der Haut. Häufig treten sie im Gesicht auf. Sie bilden sich nicht zurück, sondern werden oft dunkler, wenn das Kind heranwächst, was sehr entstellend wirken kann. Im Laufe der Zeit wurden verschiedene Behandlungen ausprobiert: u. a. Kältetherapie, Operation und Bestrahlung, die jedoch kaum Wirkungen gezeigt haben und mit zahlreichen Nebenwirkungen einhergingen. Erst die Einführung der Lasertherapie brachte eindrucksvolle Ergebnisse: Bei den meisten Feuermalen ist eine Besserung meist schon nach einer einzigen Laserbehandlung sichtbar. Schäden am angrenzenden Gewebe der Haut durch die Hitzestreuung aus dem Laser sind vorübergehender Natur.^{2, 3}

Imatinib bei chronischer myeloischer Leukämie

Beeindruckende Ergebnisse waren auch bei Patienten mit chronischer myeloischer Leukämie nach der Gabe des Arzneimittels Imatinib zu beobachten.^{4,5}

Vor der Einführung von Imatinib gegen Ende der 1990er-Jahre hatte dieser Leukämietyp nur sehr schlecht auf die Standardtherapien angesprochen. Als das neue Medikament zunächst bei Patienten getestet wurde, die auf die Standardtherapie nicht ansprachen, besserte sich ihre Prognose deutlich. Imatinib beeinflusst den Krankheitsverlauf positiv und scheint das Leben der Betroffenen im Vergleich zur Zeit vor der Einführung dieses Medikaments bei im Großen und Ganzen nur leichten Nebenwirkungen beträchtlich zu verlängern. Inzwischen gilt Imatinib als Behandlungsoption der ersten Wahl.

Entfernung von Fremdkörpern aus der Nase

Auch Maßnahmen ohne großen apparativen Aufwand (sogenannte Low-Tech-Verfahren) können dramatische Effekte haben. Manchmal stecken sich Kleinkinder kleine Gegenstände (z. B. Plastikspielzeug oder Perlen) in die Nase. Häufig schaffen sie es aber nicht, durch die Nase zu schnäuzen, um den Fremdkörper wieder auszustoßen. Die im Englischen anschaulich als «Mother's Kiss», also «mütterlicher Kuss», bezeichnete Methode zur Entfernung von Fremdkörpern – bei der Mutter oder Vater dem Kind das freie Nasenloch zuhalten, während sie/er dem Kind in den Mund pustet – ist nicht nur ausgesprochen einfach, sondern obendrein noch sehr effektiv.^{2,6}

Eine neue Behandlung von Erdbeer-Hämangiomen

Gelegentlich werden Therapien, die dramatische Wirkungen zeigen, auch rein zufällig entdeckt. Ein gutes Beispiel dafür ist die Therapie des Hämangioms. Dieser «Blutschwamm» tritt bei Säuglingen auf und ist, ähnlich wie die Portweinflecken, ebenfalls auf eine Fehlbildung unreifer Blutgefäße zurückzuführen. Hämangiome sind Wucherungen kleiner, neu gebildeter Blutgefäße. Meist ist die Haut,

üblicherweise an Kopf und Hals, betroffen; Hämangiome können aber auch in Körperorganen wie der Leber vorkommen. Die Hauterscheinungen (wegen ihres hellroten, erhabenen Erscheinungsbildes auch Erdbeerläsionen genannt) sind normalerweise nicht gleich bei der Geburt sichtbar, sondern treten im Allgemeinen etwa in der ersten Lebenswoche auf. Nach den ersten drei Lebensmonaten, in denen sie meist sehr schnell wachsen, verlangsamt sich die Wachstumsgeschwindigkeit. In den meisten Fällen bilden sie sich bis zum Alter von etwa fünf Jahren von selbst zurück; dabei kann ein schwach rosafarbenes Mal oder überschüssige Haut zurückbleiben.

Manche Hämangiome müssen wegen ihrer Lage jedoch behandelt werden – z. B., wenn sie ein Auge überdecken oder die Nase blockieren. Eine Behandlung kann aber auch wegen anderer Komplikationen erforderlich werden: Ausgeprägte geschwürartige (ulzerierte) Hämangiome können sich infizieren, oder bei Patienten mit großflächigen Gewebeeränderungen kann sich eine Herzschwäche (Herzinsuffizienz) entwickeln, weil das Herz so viel Blut durch die in diesem Blutschwamm befindlichen Blutgefäße pumpen muss.

Bis vor kurzem galten Steroide bei problematischen Blutgefäßen als medikamentöse Behandlung der ersten Wahl. 2008 konnten Ärzte jedoch mit einer anderen Therapie, auf die sie rein zufällig gestoßen waren, aufsehenerregende Ergebnisse erzielen. Bei einem Säugling mit einem riesigen Hämangiom, das fast das ganze Gesicht und das rechte Auge überdeckte, hatten sie zunächst Steroide eingesetzt. Als das Baby trotz dieser Maßnahme aber eine Herzinsuffizienz entwickelte, leiteten sie die Therapie mit Propranolol ein, dem Standardmedikament für diese Erkrankung. Zu ihrem großen Erstaunen besserte sich binnen 24 Stunden auch das Aussehen des Hämangioms. Innerhalb von einer Woche war der Tumor so weit geschrumpft, dass das Baby ein Augenlid öffnen konnte. Nach sechsmonatiger Behandlung hatte sich das Hämangiom zurückgebildet. Im darauffolgenden Jahr wendeten die Ärzte Propranolol bei einem Dutzend weiterer Kinder mit ähnlich großem Erfolg an. Auch andere Ärzte erzielten bei einer kleinen Anzahl von Kindern eindrucksvolle Ergebnisse mit diesem Medikament. Mittlerweile wird Propranolol bei einer größeren Anzahl von Säuglingen genauer untersucht.^{7,8}

Schrittweise Erfolge sind nicht gerade schlagzeilenträchtig

«Die Wissenschaft selbst eignet sich nicht besonders gut für die Schlagzeilen: Ihrem Wesen nach ist sie besser im Feuilleton aufgehoben, denn ihre Fortschritte äußern sich in der Regel nicht in plötzlichen, bahnbrechenden Durchbrüchen. Vielmehr schreitet die Wissenschaft eher über sich stufenweise entwickelnde Ideen und Theorien voran. Untermauert werden diese Theorien durch eine ganze Reihe von Ergebnissen, die aus den verschiedensten Fachgebieten stammen und den unterschiedlichsten Erklärungsstufen zuzuordnen sind. Und trotzdem sind die Medien immer nur auf «neue Durchbrüche» fixiert.»

Goldacre B. *Bad Science*. London: Fourth Estate, 2008, S. 219.

Moderate Behandlungseffekte: normal und nicht ganz so offensichtlich

Die meisten Therapien haben allerdings keine so dramatischen Effekte. Um sie genau einzuschätzen, bedarf es fairer Tests. Gelegentlich kommt es auch vor, dass eine Therapie unter bestimmten Umständen eine dramatische Wirkung entfaltet, in anderen Situationen dagegen nicht.

Auch wenn die Wirksamkeit von Vitamin B₁₂ zur Behandlung der perniziösen Anämie (s. o.) unbestritten ist, wird bis heute darüber diskutiert, ob die Patienten vierteljährliche oder häufigere Injektionen benötigen. Diese Frage wird man nur mithilfe von sorgfältig kontrollierten Tests und durch Vergleich der einzelnen Optionen beantworten können. Ein weiteres Beispiel ist der Hüftgelenkersatz, durch den es zu einer dramatischen Schmerzlinderung kommt. Die jeweiligen Vorzüge der verschiedenen Typen von künstlichen Hüftgelenken sind dagegen nicht so offensichtlich, können aber dennoch relevant sein: So unterliegen einige Kunstgelenke einem schnelleren Verschleiß als andere. Auch im Zusammenhang mit der Lasertherapie von Portweinmalen (s. o.) können wir noch viel lernen. Diese Behandlung gilt zwar nach wie als der «Goldstandard», doch wird weiterhin darüber geforscht, warum manche Hautmale nach einigen Jahren wieder nachdunkeln und welche Auswirkungen die verschiedenen Typen von Lasern haben: Möglicherweise besteht ein Zusammenhang mit der Kühlung der Haut während des Eingriffs.^{9, 10}

Und während die Gabe von Aspirin unmittelbar bei der Diagnosestellung das Sterberisiko von Herzinfarkt-Patienten erheblich senkt, so hängt die Antwort auf die Frage, ob die Einnahme von Aspirin im Hinblick auf die Herzinfarkt- und Schlaganfallprävention mehr schadet als nützt, doch auch davon ab, ob die Patienten an einer zugrunde liegenden Herz-Kreislauf- (kardiovaskulären) Erkrankung leiden. Die Vorteile (Senkung des Risikos für Herzinfarkt, Schlaganfall und kardiovaskulär bedingten Tod) müssen gegen die Risiken (insbesondere den durch Blutungen in das Gehirn verursachten Schlaganfalltyp sowie gastrointestinale Blutungen) abgewogen werden. Bei Patienten, die bereits an einer Herz-Kreislauf-Erkrankung leiden, wiegen die Vorteile des Medikaments eindeutig schwerer als seine Risiken. Bei ansonsten gesunden Menschen dagegen wird das Blutungsrisiko nicht eindeutig durch die mit Aspirin verbundenen Vorteile aufgewogen (s. Kap. 7).¹¹

Wenn Ärzte sich nicht einig sind

Bei vielen Erkrankungen und Beschwerden besteht eine erhebliche Unsicherheit, inwieweit die verfügbaren Therapien wirksam sind oder welche Behandlung für welchen Patienten am besten ist. Das hindert manche Ärzte allerdings nicht daran, eine eindeutige Meinung dazu zu vertreten, obwohl diese von Arzt zu Arzt durchaus sehr unterschiedlich ausfallen kann. Dies kann im Hinblick auf die Behandlungen, die für eine bestimmte Erkrankung verordnet werden, zu beträchtlichen Unterschieden führen.

Als Iain Chalmers, einer der britischen Autoren dieses Buchs, sich in den 1990er-Jahren während eines Urlaubs in den USA das Sprunggelenk brach, musste er dort einen orthopädischen Chirurgen aufsuchen, der das Bein vorübergehend in eine Schiene legte. Nach Abklingen der Schwellung sollte dann für sechs Wochen ein Unterschenkelgips angelegt werden. Nach Großbritannien zurückgekehrt, suchte er einige Tage später die örtliche Unfallchirurgie auf, wo der britische orthopädische Chirurg die Empfehlung seines amerikanischen Kollegen ohne Umschweife verwarf. Das Bein einzugipsen sei, so meinte er, vollkommen unangebracht. Angesichts

dieser offensichtlichen fachärztlichen Unsicherheit erkundigte sich Iain Chalmers, ob er an einem kontrollierten Vergleich teilnehmen könne, um herauszufinden, welche Behandlung besser sei.

Der britische Chirurg antwortete, kontrollierte Studien wären etwas für Leute, die sich nicht sicher seien, ob sie Recht hätten oder nicht – er dagegen habe keinen Zweifel, dass er Recht habe.

Wie kann es zu derartigen Unterschieden in der fachlichen Beurteilung kommen, und wie soll der Patient sich in einem solchen Fall verhalten? Beide Chirurgen waren sich sicher, dass sie sich für die richtige Behandlung entschieden hatten. Und doch lassen ihre weit auseinandergelassenen Ansichten deutlich die innerhalb der orthopädischen Chirurgie insgesamt bestehende Unsicherheit über die beste Behandlung eines gewöhnlichen Knochenbruchs erkennen. Gab es solide Belege dafür, welche Behandlung besser war? Und wenn ja, war sie einem der beiden Chirurgen oder keinem von ihnen bekannt? Oder wusste überhaupt niemand, welche Behandlung die bessere war (**Abb. 7**)?

Vielleicht lag der Unterschied zwischen den beiden Chirurgen ja darin, welchen Stellenwert sie den jeweiligen Behandlungsergebnissen zumaßen: Der amerikanische Chirurg hatte vielleicht eher die Schmerzlinderung im Sinn – deshalb die Empfehlung einer Gipschiene –, während sein britischer Kollege vielleicht Bedenken wegen des etwaigen Muskelschwunds hatte, zu dem es kommt, wenn eine Gliedmaße durch einen Gipsverband immobilisiert wird. Wenn dies zutraf, warum wurde Iain Chalmers dann von keinem der beiden Chirurgen gefragt, welches Behandlungsergebnis für ihn

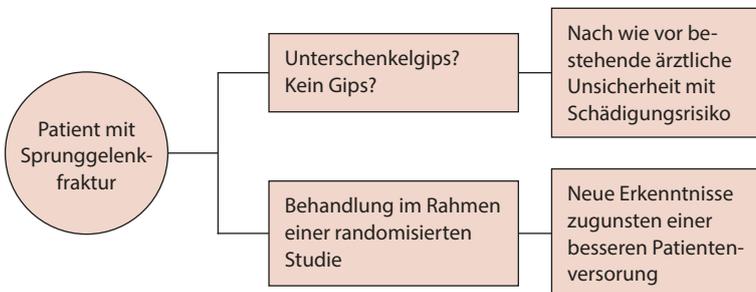


Abbildung 7: *Wie soll der Arzt entscheiden?*

als Patient wichtiger war? Und auch heute, zwei Jahrzehnte später, besteht nach wie vor Unsicherheit über die Behandlung dieser doch sehr häufig vorkommenden Art von Verletzung.¹²

In diesem Zusammenhang stellen sich verschiedene Fragen. Erstens: Lagen zuverlässige Ergebnisse aus klinischen Studien (Evidenz) vor, anhand derer man die beiden sehr unterschiedlichen Ansätze, die von den zwei Chirurgen empfohlen wurden, hätte vergleichen können? Wenn ja, enthielt diese Evidenz Angaben zu den jeweiligen Auswirkungen auf die Behandlungsergebnisse (z. B. Schmerzlinderung oder weniger Muskelschwund), die für Iain oder für andere Patienten mit möglicherweise anderen Präferenzen wichtig sein könnten? Und was passiert, wenn es keine Evidenz mit den zur Beantwortung dieser Fragen nötigen Informationen gibt?

Sich mit Unsicherheiten auseinandersetzen: eine Frage auf Leben und Tod

«Wenn man sich nicht mit den Unsicherheiten über die Effekte von Therapien auseinandersetzt, dann kann dies in ganz massivem Umfang zu Leiden und Todesfällen führen, die eigentlich vermeidbar wären. So hätten hunderttausenden von Frauen Leiden und Tod erspart werden können, wenn man Diazepam und Phenytoin bei ihrer Einführung zur Behandlung von Eklampsie (d. h. bei ihrer Zulassung als Antikonvulsiva) mit dem seit Jahrzehnten in Gebrauch befindlichen Magnesiumsulfat verglichen hätte. Und wenn die Wirkungen systemischer Steroide zur Behandlung traumatischer Hirnverletzungen überprüft worden wären, bevor diese Therapie breite Anwendung fand, hätten zehntausende von unnötigen Todesfällen verhindert werden können. Dies sind nur zwei Beispiele von vielen, die man anführen könnte, um zu veranschaulichen, warum Ärzte von Berufs wegen die Pflicht haben, sich an der Klärung von Unsicherheiten im Hinblick auf Behandlungseffekte zu beteiligen.»

Chalmers I. Addressing uncertainties about the effects of treatments offered to NHS patients: whose responsibility?
Journal of the Royal Society of Medicine 2007; 100: 440.

Manche Ärzte wissen genau, was sie tun müssen, wenn es keine zuverlässigen Informationen über die Wirkungen alternativer Behandlungen gibt. Sie sind darauf eingestellt, mit ihren Patienten über diese Unsicherheit zu sprechen. So meinte ein Arzt, der sich auf die Versorgung von Schlaganfall-Patienten spezialisiert hat, die Forschungsergebnisse würden zwar zeigen, dass es für seine Patienten besser wäre, wenn sie auf einer speziellen Schlaganfallstation (Stroke Unit) versorgt würden, dass es aber – bei vielen Typen von Patienten

– nach wie vor unklar sei, ob sie Blutgerinnsel auflösende Medikamente erhalten sollten oder nicht (s.a. Kap. 11, S. 195). Bei der Besprechung der verschiedenen Behandlungsmöglichkeiten erklärte er seinen Patienten, dass diese Medikamente mehr nutzen als schaden können. Bei manchen Patienten könne aber auch der Schaden den Nutzen überwiegen. Einem Patienten, für den das Verhältnis von Nutzen zu Schaden dieser Therapie nicht eindeutig zu klären war, empfahl er diese Behandlung nur unter der Voraussetzung, dass sie im Rahmen eines sorgfältig kontrollierten Vergleichs durchgeführt würde, der dazu beitragen könnte, diese Unsicherheit zu verringern.¹³ Bei den Blutgerinnsel auflösenden Medikamenten bestehen hinsichtlich mehrerer Aspekte nach wie vor Unsicherheiten.¹⁴

Die Auseinandersetzung mit Unsicherheiten zeugt von Professionalität

«Eines der wichtigsten Merkmale von Professionalität ... sollte die Fähigkeit sein, Unsicherheiten in der Medizin zu identifizieren und sich mit ihnen auseinanderzusetzen. Tagtäglich sind Ärzte im Hinblick auf die Entstehung von Krankheiten (Pathogenese) sowie ihre Diagnostik und Therapie mit Unsicherheiten konfrontiert. Und doch werden diese Unsicherheiten, die allen drei Gebieten ärztlicher Tätigkeit innewohnen, nur selten explizit zur Kenntnis genommen. Manche Ärzte fühlen sich sogar unwohl, wenn sie – vor allem gegenüber Patienten – Unsicherheiten zugeben müssen. Unsicherheit dient der medizinischen Forschung aber auch als wichtiger Impuls zur Verbesserung der menschlichen Gesundheit. Es wird in Zukunft immer wichtiger werden, dass Ärzte sich mit sämtlichen für ihr jeweiliges Fachgebiet relevanten Forschungsergebnissen vertraut machen, damit sie wissen, wo weiterhin Unsicherheiten bestehen und woran gerade geforscht wird oder woran geforscht werden müsste, um diese Unsicherheiten abzubauen. Ein Zeichen für Professionalität wird alles in allem zukünftig ein Forschungsbewusstsein zum Wohle des Patienten sein. Einige Ärzte werden sich aktiv an dieser Forschung beteiligen, alle anderen sollten sie zumindest fördern, ihre Patienten gegebenenfalls aktiv in die medizinischen Forschungsvorhaben einzubinden versuchen und die Ergebnisse dieser Forschung in ihrer ärztlichen Praxis umsetzen.»

Aus: Medical Research Council response to Royal College of Physicians consultation on medical professionalism. 2005

Koffein zur Behandlung von Atemproblemen bei Frühgeborenen

Große Unterschiede bei den Behandlungen, die bei einer bestimmten Krankheit angewendet werden, sind ein deutlicher Hinweis auf fachliche Unsicherheiten bezüglich der jeweiligen Vorzüge verschied-

dener Therapien. Und eingefahrene Gepflogenheiten können bedeuten, dass es sehr lange dauert, bevor solche Unsicherheiten durch faire Tests zerstreut werden. Ein besonders treffendes Beispiel ist die Anwendung von Koffein bei Frühgeborenen. Frühgeborene Babys haben häufig Schwierigkeiten mit der richtigen Atmung, und manchmal hören sie auch für sehr kurze Zeit auf zu atmen. Von diesen Problemen, die man auch als Frühgeborenen-Apnoe bezeichnet, sind sehr viele Säuglinge betroffen, die vor der 34. Schwangerschaftswoche geboren werden. Als Ende der 1970er-Jahre der Nachweis erbracht wurde, dass die Behandlung mit Koffein die Anzahl der Apnoe-Episoden reduziert, wurde sie von etlichen Kinderärzten angewendet.

Dennoch blieben die Wirkungen von Koffein umstritten. Obwohl faire Tests gezeigt hatten, dass die Anzahl der Apnoe-Episoden durch die Gabe von Koffein abnahm, waren viele Kinderärzte der Meinung, dass diese Episoden nicht schwerwiegend genug waren, um die Anwendung von Koffein zu rechtfertigen; und manche hatten Bedenken, dass sie bei diesen winzigen Babys möglicherweise nicht ohne Risiko für ihre Gesundheit war. Das bedeutete, dass einige Säuglinge mit Koffein behandelt wurden und andere nicht. Als diese weitverbreitete Unsicherheit schließlich mehr als 30 Jahre nach der Einführung dieser Therapie in einer großen internationalen Studie untersucht wurde, stellte sich heraus, dass diese einfache Therapie nicht nur die Atemschwierigkeiten minderte, sondern auch – was sehr wichtig war – die Wahrscheinlichkeit eines Langzeitüberlebens ohne Zerebralparese und ohne Verzögerung der kindlichen Entwicklung signifikant verbesserte. Hätte man sich bereits bei Einführung der Therapie mit dieser Unsicherheit auseinandergesetzt, wäre es bei weniger Babys zur Entwicklung von Behinderungen gekommen.^{15, 16}

Antibiotika bei vorzeitigen Wehen

Wenn medizinische Behandlungen, von denen man sich vorteilhafte Effekte erhofft und die für harmlos gehalten werden, durch faire Tests untersucht werden, kann sich herausstellen, dass keine von beiden Annahmen den Tatsachen entspricht. Ärzte verordnen Thera-

prien in der besten Absicht, vor allem, wenn sie ihren Patienten in einer verzweifelten Situation damit Hoffnung geben können. Eine Annahme lautete beispielsweise, dass eine «stumme» (klinisch kaum nachweisbare) Infektion vorzeitige Wehen und eine Frühgeburt auslösen könne. Diese Theorie verleitete Ärzte dazu, Schwangeren Antibiotika zu verschreiben – in der Hoffnung, dass sie die Schwangerschaft dann fortsetzen könnten. Niemand dachte ernsthaft daran, dass diese Anwendung von Antibiotika schwerwiegende Probleme verursachen würde. Es gibt sogar Hinweise darauf, dass die betroffenen Frauen selbst mit Antibiotika behandelt werden wollten – nach dem Motto: «Lassen Sie es uns versuchen; es kann ja nicht schaden.»

Als man diese Therapie letzten Endes in einem fairen Test untersuchte, hatten die Ergebnisse eindeutige Konsequenzen für die klinische Praxis: Zunächst einmal ließen sich keinerlei Vorteile feststellen. Darüber hinaus ergab die langfristige Nachbeobachtung der Babys in dieser Studie, dass diejenigen, die im Mutterleib Antibiotika ausgesetzt gewesen waren, eher Zerebralparesen und Probleme mit der Sprache, dem Sehvermögen und dem Laufen aufwiesen als die Säuglinge in den Vergleichsgruppen. In all den Jahrzehnten, in denen Schwangeren Antibiotika verordnet worden waren, ohne dass ihre Effekte durch faire Tests ausreichend belegt gewesen wären, waren diese Risiken von Antibiotika unerkannt geblieben. Wie so häufig trugen diejenigen, die im «normalen» klinischen Alltag eine nicht ausreichend bewertete Behandlung erhielten, eher Schäden da-

Ärzte im Gespräch über das Rätselraten beim Verordnen von Behandlungen

In einem fiktiven Gespräch zwischen zwei Ärzten stellte ein Allgemeinmediziner Folgendes fest: «Ein Großteil unserer Tätigkeit besteht aus Rätselraten, und ich kann mir nicht vorstellen, dass wir beide uns dabei besonders wohlfühlen. Die einzige Möglichkeit herauszufinden, ob etwas wirkt, ist eine richtige Studie, aber die Hürden sind sehr hoch. Was tun wir also? Wir tun, wonach uns gerade der Sinn steht. Und ich bin sicher, dass das manchmal auch völlig in Ordnung ist – immerhin haben wir als Ärzte ja Erfahrung und all das. Den Rest der Zeit liegen wir damit vermutlich genauso oft richtig wie falsch. Doch weil das, was wir da tun, von niemandem als Studie bezeichnet wird, unterliegt es auch keinerlei Vorschriften und keiner von uns kann was daraus lernen.»

In Anlehnung an Petit-Zeman S. Doctor, what's wrong?
Making the NHS human again. London: Routledge, 2005, S. 79–80.

von als diejenigen, denen dieselbe Therapie im Rahmen einer wissenschaftlichen Studie verordnet wurde. Anders ausgedrückt: Diejenigen, welche die Medikamente nicht im Rahmen eines fairen Tests erhielten, waren generell einem höheren Risiko ausgesetzt.^{17, 18, 19}

Brustkrebstherapie

Die Behandlung von Brustkrebs (s. Kap. 3) stellt ein weiteres Beispiel für fachliche Unsicherheit dar. Im Hinblick auf den Einsatz von Operation, Strahlentherapie und Chemotherapie bestehen erhebliche Meinungsunterschiede. Es ist ungeklärt, wie man Mammakarzinome in einem sehr frühen Stadium und sogenannte «Pseudokarzinome» der Brust am besten behandelt. Das Gleiche gilt auch für die Frage, wie viele Lymphknoten im Idealfall aus der Achselhöhle entfernt werden müssen oder ob überhaupt Lymphknoten entfernt werden sollten.²⁰ Und als ob das noch nicht genug wäre, so sind auch Fragen, die für die Patientinnen von besonderem Interesse sind, wie etwa die Linderung der therapiebedingten Müdigkeit (Fatigue) oder die beste Möglichkeit zur Behandlung von Lymphödemem des Arms – eine belastende und beeinträchtigende Folge der Operation und der Bestrahlung der Achselhöhle – noch immer nicht angemessen untersucht worden.

Unsicherheiten in Bezug auf Therapieeffekte thematisieren

Wie geht es nun weiter? Ärzte müssen Zugriff auf Wissensquellen haben, in denen sie die beste aktuelle Evidenz über eine Therapie finden, die auf gebündelter Sachkenntnis und systematischen Übersichtsarbeiten (Reviews) aller vorhandenen zuverlässigen Studien beruht. Wenn sie das Gefühl haben, dass danach noch immer Unsicherheit bezüglich einer bestimmten Therapie herrscht, müssen sie bereit sein, diesen Umstand mit ihren Patienten zu besprechen, und ihnen den Grund dafür erklären. Patienten und Ärzte können die verschiedenen Möglichkeiten dann gemeinsam erörtern; dabei gilt es die Präferenzen des Patienten zu berücksichtigen. Und natürlich

können solche Gespräche weitere Unsicherheiten wecken, die erkannt und besprochen werden müssen. Nur die gemeinsame Erkenntnis, dass noch Unsicherheiten bestehen, ermöglicht einen kontinuierlichen Fortschritt, durch den medizinische Behandlungen zielgerichteter und sicherer gemacht werden können. So gesehen ist Unsicherheit eine Voraussetzung für den Fortschritt und nicht das Eingeständnis einer «Niederlage». Diese positive Einstellung gegenüber dem Umgang mit Unsicherheiten spiegelt sich mittlerweile auch in verschiedenen ärztlichen Leitfäden wider. In Großbritannien wird Ärzten in der jüngsten Fassung des vom *General Medical Council* herausgegebenen Leitfadens *Good Medical Practice* der Rat erteilt, dass sie im Zuge der Erhaltung und Verbesserung ihrer Leistungen «ihren Beitrag dazu leisten müssen, Unsicherheiten bezüglich der Effekte von Behandlungen zu klären».²¹ Hierzu müssen Patienten und Ärzte mit dem Ziel zusammenarbeiten, bessere Studien zu planen (s. Kap. 11).

Medizinische Versorgung im Rahmen fairer Tests anbieten

Was sollte also passieren, wenn relevante Unsicherheiten über die Wirkungen neuer oder alter Therapien bestehen, die nicht richtig geprüft wurden? Eine offensichtliche Antwort lautet, unserem obigen Beispiel des Arztes zu folgen, der sich um Schlaganfall-Patienten kümmert: sich mit den bestehenden Unsicherheiten auseinandersetzen, indem man unzureichend bewertete Therapien ausschließlich im Rahmen wissenschaftlicher Studien anbietet, die darauf ausgelegt sind, mehr über ihre erwünschten und unerwünschten Wirkungen herauszufinden.

Ein Medizinethiker hat das einmal so ausgedrückt:

Wenn wir uns unsicher sind, welche spezifischen Vorzüge jeweils [unterschiedliche] Behandlungen haben, dann können wir auch bezüglich der Vorzüge der Anwendung einer einzelnen Behandlungsmethode – etwa bei der Behandlung eines einzelnen Patienten – keine Gewissheit

haben. Deshalb scheint es irrational und ethisch nicht vertretbar zu sein, vor Abschluss einer geeigneten Studie auf der einen oder der anderen Methode zu beharren. Die Antwort auf die Frage: «Welches ist die für den Patienten beste Behandlung?» lautet daher: «Die Studie». Die Studie ist die Behandlung. Hat das etwas mit Herumexperimentieren zu tun? Ja. Aber was ich damit meine, ist: eine Entscheidung unter Unsicherheit treffen plus Daten sammeln. Spielt es eine Rolle, dass es sich dabei um eine «zufällige» Entscheidung handelt? Logischerweise nein. Denn welche bessere Methode könnte es geben, um sich, wenn Unsicherheit herrscht, zu entscheiden?»²²

Die Verabreichung von Therapien im Rahmen fairer Tests kann für Patienten im Hinblick auf das Behandlungsergebnis einen tief greifenden Unterschied ausmachen. Die Geschichte der kindlichen Leukämien liefert dafür ein sehr drastisches Beispiel. Bis in die 1960er-Jahre verstarb nahezu jedes Kind mit Leukämie kurz nach der Diagnosestellung. Heutzutage überleben etwa 85 von 100 Kindern. Dies konnte erreicht werden, weil die Mehrzahl der Kinder mit Leukämie an randomisierten Studien teilnahm, in denen die jeweilige Standardtherapie mit einer neuen Variante dieser Therapie verglichen wurde.²³ Bei den meisten Kindern mit einer Krebserkrankung wird die beste Behandlungsoption deshalb durch die Teilnahme an solchen Studien ausgewählt.

Können Patienten mit Unsicherheit umgehen?

«Wo stehen wir also bezüglich des Umgangs mit Unsicherheiten gegenüber Therapieeffekten? ... Ungeachtet der allgemeinen Anerkennung, dass Patienten in der medizinischen Forschung und bei gesundheitlichen Entscheidungen Partner sind, sind manche Ärzte von der Komplexität der Diskussion über therapeutische Unsicherheiten entnervt. Manche haben einfach Angst, dass sie dadurch beim Patienten Besorgnis auslösen: eine zweifellos aufrichtige Sorge, die aber dennoch bevormundend wirkt. Andere versuchen, ihr Handeln unter dem Aspekt der Ausgewogenheit zwischen zwei ethischen Argumenten zu rechtfertigen – erstreckt sich die ethische Pflicht, die Wahrheit zu sagen, auch auf das Ansprechen von Unsicherheiten, oder soll die moralische Verpflichtung, den Patienten vor emotionalen Belastungen zu schützen, mehr Gewicht haben? Sind die Patienten bereit, mit Unsicherheiten zu leben? Das müssen wir herausfinden. Vielleicht sind die Menschen sehr viel belastbarer, als Ärzte meinen.»

Evans I. More nearly certain.

Journal of the Royal Society of Medicine 2005; 98: 195–6.

Wenn keine solche Studie verfügbar ist, sollten die Ergebnisse, die aus der Anwendung neuer und nicht getesteter Therapien resultieren, in standardisierter Form dokumentiert werden – beispielsweise mithilfe einer Checkliste, die verschiedene Punkte umfasst: etwa Labor- oder andere Untersuchungen, die zur Diagnose einer Krankheit herangezogen werden, sowie die Untersuchungen, die durchgeführt werden, um die Auswirkungen einer Therapie zu bewerten. Ein solcher Untersuchungsplan könnte – wie es auch bei klinischen Studien der Fall sein sollte – in einer Datenbank registriert werden (s. Kap. 8). Auf diese Weise könnten die Ergebnisse in den vorhandenen Wissensfundus eingespeist werden – zum Vorteil der Patienten, welche die nicht getestete Therapie erhalten, wie auch aller anderen Patienten. Riesige Summen sind bereits in Informationstechnologie-Systeme für die Gesundheitsversorgung investiert worden, die problemlos dazu genutzt werden könnten, um diese Informationen zum Nutzen der Patienten wie auch der Öffentlichkeit zu erfassen (s. a. Kap. 11).²⁴

Wenn Unsicherheiten über die Wirkungen von Behandlungen effektiver und effizienter angegangen werden sollen, wird es zu Veränderungen kommen müssen. Mit einigen dieser Veränderungen – insbesondere der stärkeren Beteiligung der Patienten – werden wir uns später noch befassen (s. Kap. 11 und 12). Einen Punkt – den wir bereits weiter oben schon kurz angesprochen haben – möchten wir an dieser Stelle aber noch besonders hervorheben. Wenn es nicht genügend Informationen über die Wirkungen von Therapien gibt, kann man die Erkenntnisse auch dadurch verbessern, dass man dafür sorgt, dass Ärzte eine solche Therapie so lange nur im Rahmen einer formalen Bewertung anbieten, bis man mehr über ihren Wert und ihre möglichen Nachteile in Erfahrung gebracht hat. Doch im Grunde wird eine solche risikobegrenzende Vorgehensweise durch manche der vorherrschenden Standpunkte, darunter z. B. auch in der Reglementierung der Forschung (s. Kap. 9), eher behindert.

Über dieses Problem hat sich vor mehr als 30 Jahren schon ein britischer Kinderarzt geärgert: Wenn er der Hälfte seiner Patienten eine Behandlung zukommen lassen wollte (um etwas über die Wirkungen dieser Behandlung herauszufinden, indem er in einem

kontrollierten Vergleich bei der einen Hälfte der Patienten die neue Behandlung und bei der anderen Hälfte die bereits bekannte Therapie anwandte) bräuchte er – so seine prägnante Feststellung – eine Genehmigung. Nicht erforderlich sei eine solche Genehmigung dagegen, wenn er allen seinen Patienten dieselbe Behandlung als Standardverordnung gäbe.²⁵ Dieser unlogische Doppelstandard taucht immer wieder auf und schreckt diejenigen Ärzte ab, die ihren Beitrag zum Abbau von Unsicherheiten in Bezug auf Therapieeffekte leisten wollen. Im Großen und Ganzen kann dies dazu führen, dass Ärzte davon abgehalten werden, aus ihren Erfahrungen mit der Versorgung von Patienten neues Wissen zu generieren. Wie der amerikanische Soziologe Charles Bosk einmal bemerkte: «Alles ist erlaubt, solange wir versprechen, nicht aus unseren Erfahrungen zu lernen.»

Die Fähigkeit, Unsicherheiten zu erklären, verlangt aufseiten der Ärzte sicherlich Geschick und ein gewisses Maß an Demut. Viele fühlen sich unwohl, wenn sie potenzielle Teilnehmer an einer klinischen Studie darüber aufklären müssen, dass niemand weiß, welche Behandlung die beste ist. Doch die öffentliche Meinung hat sich gewandelt: Mit arroganten Ärzten, die «Gott spielen», wird kurzer Prozess gemacht. Wir müssen uns auf die Ausbildung von Ärzten konzentrieren, die sich nicht schämen zuzugeben, dass sie auch nur Menschen sind und die Unterstützung und Mitwirkung von Patienten an der Forschung brauchen, um Behandlungsentscheidungen auf eine sicherere Grundlage stellen zu können (s. Kap. 11 und 12).

Das größte Hindernis stellt für viele Ärzte und Patienten die mangelnde Vertrautheit mit den Gegebenheiten solcher fairen Therapietests dar – ein Problem, mit dem wir uns als Nächstes befassen wollen (s. Kap. 6).

- Dramatische Therapieeffekte kommen selten vor.
- Sehr häufig herrscht Unsicherheit über die Wirkungen von Therapien.
- Kleine Unterschiede in den Wirkungen unterschiedlicher Behandlungen sind normal, und es ist wichtig, diese Unterschiede zuverlässig aufzudecken.
- Wenn bei einer relevanten Unsicherheit hinsichtlich der Wirkungen einer Therapie niemand die Antwort kennt, muss etwas unternommen werden, um diese Unsicherheit abzubauen.
- Es könnte sehr viel mehr getan werden, damit auch Patienten ihren Beitrag zur Verringerung von Unsicherheiten über Therapieeffekte leisten können.

6

Faire Tests von Therapien

Wahrscheinlich sind viele Leser mit den Prinzipien, die fairen Tests medizinischer Therapien zugrunde liegen, nicht vertraut. Dabei sind sie gar nicht besonders kompliziert. Genau genommen beruht auch ein Großteil unseres intuitiven Alltagsverständnisses auf diesen Prinzipien. Trotzdem lernt man sie nicht in der Schule, und häufig sind sie auch noch in unnötig komplizierte sprachliche Formulierungen verpackt. Infolgedessen scheuen viele Menschen das Thema, in der Annahme, es ginge über ihr Begriffsvermögen hinaus. Wir hoffen, dass dieses und die beiden folgenden Kapitel Sie davon überzeugen werden, dass Sie die wichtigsten Prinzipien bereits kennen und somit ohne weiteres verstehen werden, warum sie so wichtig sind. Leser, die sich näher mit diesen Themen befassen möchten, finden zusätzliche Informationen unter www.testingtreatments.org sowie in der *James Lind Library* (www.jameslindlibrary.org).

Warum wir faire Therapietests brauchen

Natürliche Heilkräfte

Viele gesundheitliche Probleme verschlimmern sich meistens, wenn man sie nicht behandelt, und einige verschlimmern sich auch trotz Behandlung. Manche Krankheiten bessern sich aber auch von allein – man sagt in einem solchen Fall, sie seien «selbstlimitierend». Ein Wissenschaftler, der an der Überprüfung eines Behandlungsansatzes für Erkältungen beteiligt war, hat dies einmal folgendermaßen ausgedrückt: «Behandelt man eine Erkältung energisch, dann erholt sich der Patient innerhalb von sieben Tagen. Lässt man der Erkältung ihren Lauf, fühlt er sich nach einer Woche wieder wohl.» Oder

etwas zynischer ausgedrückt: «Die Natur heilt, der Arzt kassiert das Honorar.» Aber natürlich kann sich ein Krankheitsbild unter Therapie auch verschlechtern.

Gerade weil es vorkommt, dass Menschen häufig auch ohne irgendeine spezielle Behandlung von einer Krankheit genesen, müssen bei der Überprüfung medizinischer Therapien auch der «natürliche» Verlauf und die Folgen der unbehandelten Krankheit berücksichtigt werden. Sicher haben Sie auch schon einmal Halsschmerzen, Magenkrämpfe oder einen ungewöhnlichen Hautausschlag gehabt. Diese Beschwerden klingen häufig von selbst, d.h. ohne besondere ärztliche Behandlung, wieder ab. Wenn Sie aber deswegen behandelt worden sind, dann haben Sie vielleicht geglaubt, dass das Abklingen der Symptome auf die Behandlung zurückzuführen war (selbst wenn es sich dabei um eine unwirksame Behandlung gehandelt haben sollte). Kurzum: Die Kenntnis des natürlichen Verlaufs einer Krankheit, einschließlich der Wahrscheinlichkeit, dass sie von selbst heilt (Spontanremission), kann sowohl unnötigen Behandlungen wie auch dem falschen Glauben an Heilmittel vorbeugen, deren Wirksamkeit nicht bewiesen ist.

Falsch verstandene Heilmittel

« ... zum einen gilt es angeblich als bewiesen, dass sich Taucher, und zwar nicht wenige von ihnen, von der Taucherkrankheit als geheilt ansehen, wenn sie Tabak zu sich nehmen, und zum anderen, dass kein Mensch dadurch jemals Schaden genommen habe. Zum einen wohnt diesem Argument ein großes Missverständnis inne, zum anderen eine ungeheuerliche Absurdität: ... wenn ein Kranker sich auf dem Höhepunkt seiner Krankheit befindet und in diesem Moment Tabak zu sich genommen hat und wenn seine Krankheit danach ihren natürlichen Verlauf nimmt und nachlässt und der Patient folglich wieder gesundet, ja dann ist es fürwahr der Tabak, der dieses Wunder bewirkt hat.»

James Stuart, King of Great Britaine, France and Ireland. A counterblaste to tobacco. In: *The workes of the most high and mightie prince, James*. Published by James, Bishop of Winton, and Deane of his Majesties Chappel Royall. London: printed by Robert Barker and John Bill, printers to the Kings most excellent Majestie, 1616: S. 214–222.

Gerade wenn die Krankheitssymptome periodisch kommen und gehen, ist es schwer, die Wirkungen medizinischer Behandlungen näher zu bestimmen. Patienten mit Arthritis beispielsweise suchen vor allem dann ärztliche Hilfe, wenn sie einen besonders schweren

Schub haben. Ein solcher Schub hält aber naturgemäß selten längere Zeit an. Ganz gleich, ob es sich bei der Behandlung, welche die Patienten dann erhalten, um eine wirksame oder unwirksame schulmedizinische oder auch alternative Therapie handelt: Nach der Behandlung nehmen die Schmerzen der Patienten ab, und zwar einfach deshalb, weil der Schub abklingt. Verständlicherweise neigen Ärzte und Patienten aber dazu, solche Besserungen auf die Behandlung zurückzuführen, selbst wenn sie möglicherweise gar nicht dafür verantwortlich ist.

Die positiven Wirkungen von Optimismus und Wunschdenken

Heute wissen wir Einiges über die psychologischen Gründe dafür, dass Menschen dazu neigen, eine Besserung ihrer Krankheit der Behandlung zuzuschreiben, die sie erhalten haben. Wir alle neigen zu dem Glauben, dass, wenn ein Ereignis auf das andere folgt, dann das erste Ereignis für das zweite verantwortlich ist. Ferner neigen wir dazu, Muster zu sehen, wo keine sind: Dies ist ein Phänomen, das sich oft und bei so unterschiedlichen Dingen wie dem Münzenwerfen, bei Aktienkursen und Korbwürfen beim Basketball beobachten lässt. Wir alle sind auch anfällig für ein Problem, das man als Bestätigungsbias (engl. *confirmation bias*) bezeichnet: Wir sehen, was wir zu sehen erwarten: «Glauben heißt sehen». Jedes Argument, das unsere Überzeugungen stützt, stärkt unser Vertrauen darauf, dass wir Recht haben. Umgekehrt erkennen oder akzeptieren wir Informationen, die unseren Ansichten widersprechen, nicht so ohne weiteres und neigen deshalb dazu, die Augen davor zu verschließen, oftmals unbewusst.

Die meisten Patienten und Ärzte hoffen natürlich, dass die medizinische Behandlung hilft. Sie kommen vielleicht zu dem Schluss, dass etwas wirkt, einfach weil dies mit ihrer Überzeugung, dass es wirken *sollte*, übereinstimmt. Nach Informationen, die ihren Überzeugungen zuwiderlaufen, wird gar nicht erst gesucht, oder sie werden verworfen. Diese psychologischen Muster erklären auch, warum Patienten, die daran glauben, dass ihnen eine Behandlung helfen wird, durchaus auch eine Besserung ihres Zustands erleben können – selbst wenn die Behandlung in Wirklichkeit gar keinen aktiven

Glauben heißt sehen

Der britische Arzt Richard Asher stellte dazu in einem Vortrag Folgendes fest: «Wenn Sie inbrünstig an Ihre Therapie glauben können, selbst wenn kontrollierte Studien zeigen, dass sie im Grunde nutzlos ist, dann erzielen Sie sehr viel bessere Behandlungsergebnisse, Ihren Patienten geht es deutlich besser und auch Ihr Einkommen sieht viel besser aus. Mit diesem Phänomen lässt sich meiner Meinung nach zum einen der bemerkenswerte Erfolg erklären, den einige weniger talentierte, doch leichtgläubigere Mitglieder unseres Berufsstandes haben, zum anderen aber auch die tiefe Abneigung gegen Statistiken und kontrollierte Studien, die moderne und erfolgreiche Ärzte für gewöhnlich an den Tag legen.»

Asher R. Talking sense (Lettsomian lecture, 16 Feb, 1959). *Transactions of the Medical Society of London*, vol LXXV, 1958–59. Reproduced in: Jones, FA, ed. *Richard Asher talking sense*. London: Pitman Medical, 1972.

Wirkstoff enthält (eine «Scheinbehandlung», oft auch als «Placebo» bezeichnet). Auch nach der Einnahme einer aus Zucker bestehenden Tablette, nach einer Injektion mit Wasser, nach einer Behandlung mit inaktivierten elektrischen Geräten oder nach einer Operation, bei der nicht mehr als ein kleiner Schnitt gesetzt und wieder zugenäht wurde, haben Patienten schon über eine Besserung berichtet.

Warum der erste Eindruck nicht alles ist

Aber warum reicht es nicht, wenn Patienten glauben, dass ihnen etwas hilft? Warum ist es wichtig, keine Mühen und Kosten für Forschungsarbeiten zu scheuen, um die Wirkungen medizinischer Behandlungen auf formalere Art und Weise zu bewerten und vielleicht herauszufinden versuchen, ob und wenn ja, *wie* sie den Patienten geholfen haben? Dafür gibt es mindestens zwei Gründe. Zum einen lenken Therapien, die nicht wirken, unsere Aufmerksamkeit von wirksamen Therapien ab. Zum anderen haben viele (wenn nicht sogar die meisten) Therapien unerwünschte Nebenwirkungen, von denen einige vorübergehend sind, einige länger anhalten und manche noch gänzlich unerkannt sind. Wenn Patienten diese Therapien nicht anwenden, bleiben sie auch von diesen unerwünschten Wirkungen verschont. Deshalb sollte man versuchen herauszufinden,

welche Behandlungen höchstwahrscheinlich gar nicht helfen oder welche Behandlungen eventuell mehr schaden als nützen. Zudem kann uns die Forschung wichtige Hinweise darauf geben, wie Therapien wirken, und so Möglichkeiten zur Entwicklung besserer und sichererer Therapien aufzeigen.

Die Erforschung von Therapieeffekten spielt überall eine Rolle, besonders aber in Gesellschaften, die um die gerechte Verteilung der Ressourcen im Gesundheitswesen unter allen Patienten bemüht sind. Beispiele dafür sind der britische *National Health Service* (NHS) oder die *US Veterans Health Administration*. In diesen Systemen müssen ständig Entscheidungen darüber getroffen werden, welche Therapien – gemessen an den im Gesundheitswesen zwangsläufig nur begrenzt verfügbaren Ressourcen – ihr Geld wert sind. Wenn einige Patienten Behandlungen erhalten, deren Wirksamkeit nicht belegt ist, kann dies zur Folge haben, dass anderen Patienten Behandlungen vorenthalten werden müssen, die sich als wirksam erwiesen haben. In Deutschland wurde für solche Nutzenbewertungen 2004 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen eingerichtet (www.iqwig.de).

Die vorangegangenen Ausführungen sollen aber nicht den Eindruck erwecken, als seien die Eindrücke und Vorstellungen von Patienten und Ärzten bezüglich der Wirkungen medizinischer Therapien unwichtig. Tatsächlich sind sie häufig der Ausgangspunkt für die formale Untersuchung von anscheinend vielversprechenden neuen Therapien. Den Beobachtungen von Ärzten und Patienten durch formale Forschung nachzugehen, führt manchmal zur Aufdeckung sowohl schädlicher als auch nützlicher Therapieeffekte. Ein Beispiel dafür ist die Patientin aus Kap. 2 (S. 48) und ihre Vermutung, dass das seltene Vaginalkarzinom ihrer Tochter vielleicht durch das Medikament Diethylstilbestrol (DES) ausgelöst worden war, mit dem sie selbst 20 Jahre zuvor während ihrer Schwangerschaft behandelt worden war. Und weder der Patient, der die unerwartete Nebenwirkung des neuen Medikaments erwähnte, das ihm wegen seines erhöhten Blutdrucks verschrieben worden war, noch sein Arzt hätten sich wohl zu dem Zeitpunkt träumen lassen, dass diese Beobachtung die Grundlage für den Verkaufsschlager Sildenafil (Viagra) legen würde.

Solche – manchmal wichtigen – Einzelbeobachtungen eignen sich aber nicht als ausreichende Basis, um daraus fundierte Schlussfolgerungen über Therapieeffekte zu ziehen, ganz zu schweigen von Therapieempfehlungen für andere Patienten.

Was also sind faire Tests von Therapien?

Wie die meisten von uns wissen, kann es ein Fehler sein, Medienberichte über irgendwelche medizinischen Fortschritte für bare Münze zu nehmen. Doch die traurige Wahrheit ist, dass man auch bei Therapieberichten in anscheinend seriösen Fachzeitschriften auf der Hut sein muss. Irreführende und überzogene Behauptungen über medizinische Therapien sind auch hier keine Seltenheit, und daher ist es wichtig, dass man die Zuverlässigkeit solcher Behauptungen beurteilen kann.

Wenn wir Berichte über die Wirkungen irgendwelcher Therapien einfach unbesehen glauben, gehen wir zwei Risiken ein: Wir könnten fälschlicherweise den Schluss ziehen, dass eine hilfreiche Therapie nutzlos oder sogar gefährlich ist. Oder wir folgern ebenfalls fälschlicherweise, dass eine an sich nutzlose oder sogar gefährliche Therapie wirksam ist. Faire Therapietests sind daher dazu ausgelegt, zuverlässige Informationen über die Wirkungen von Behandlungen zu gewinnen, indem sie:

1. Gleiches mit Gleichem vergleichen, um Einflussfaktoren (systematische Fehler, Bias), die zu Ergebnisverzerrungen führen, zu reduzieren
2. den Zufallsfaktor berücksichtigen
3. alle zusätzlichen relevanten, zuverlässigen wissenschaftlichen Erkenntnisse in die Bewertung einfließen lassen.

In diesem und den nächsten beiden Kapiteln wollen wir uns mit diesen drei wichtigsten Eigenschaften fairer Tests befassen.

Gleiches mit Gleichem vergleichen: entscheidend ist der Vergleich

Bei allen fairen Therapiestudien sind Vergleiche von zentraler Bedeutung. Patienten und Ärzte vergleichen manchmal im Stillen die jeweiligen Vorteile zweier Therapien und gewinnen dabei den Eindruck, dass sie selbst bzw. andere Patienten auf eine bestimmte Behandlung anders ansprechen als auf frühere Behandlungen. Manchmal werden auch Vergleiche angestellt, die mehr den Charakter eines Versuchs haben. Schon im 9. Jh. verglich der persische Arzt Al-Razi das Behandlungsergebnis von Meningitis-Patienten, die mit Aderlässen behandelt worden waren, mit dem Ergebnis von Meningitis-Patienten, die keine Aderlässe erhalten hatten, um herauszufinden, ob Aderlässe bei Meningitis helfen könnten.

In der Regel werden Therapien getestet, indem man Gruppen von Patienten vergleicht, die unterschiedlich behandelt wurden. Wenn Therapievergleiche fair sein sollen, dann muss gewährleistet sein, dass Gleiches auch mit Gleichem verglichen wird: dass die Behandlung, welche die Patienten jeweils erhalten haben, den einzigen systematischen Unterschied zwischen den Vergleichsgruppen darstellt. Diese Erkenntnis ist keineswegs neu: Bevor James Lind 1747 an Bord der HMS *Salisbury* seinen Vergleich der sechs Skorbut-Therapien durchführte, traf er erstens eine sorgfältige Auswahl der Patienten, die sich in einem vergleichbaren Stadium dieser oftmals tödlichen Krankheit befanden; zweitens vergewisserte er sich, dass die Patienten dieselbe Nahrung erhielten; und drittens sorgte er dafür, dass sie unter ähnlichen Bedingungen untergebracht waren (s. Kap. 1, S. 32). Lind hatte erkannt, dass auch andere Faktoren als die Therapien selbst einen Einfluss auf die Heilungschancen seiner Patienten haben können.

Eine Möglichkeit, einen solchen Test *unfair* zu gestalten, hätte zum Beispiel darin bestanden, eine der empfohlenen Behandlungen – sagen wir, die Schwefelsäure, zu der das *Royal College of Physicians* in London riet – nur Patienten zu verabreichen, die zunächst weniger krank erschienen und sich noch im Frühstadium der Krankheit befanden, und eine andere Behandlung – beispielsweise die Zitrusfrüchte, die von manchen Seeleuten empfohlen wurden

– nur Patienten zu geben, die bereits dem Tod nahe waren. In diesem Fall hätte die Schwefelsäure besser abgeschnitten, obwohl es sich eigentlich um die schlechtere Therapie handelte. Zu solchen systematischen Fehlern kann es kommen, wenn nicht dafür gesorgt wird, dass in allen wichtigen Punkten Gleiches auch mit Gleichem verglichen wird.

Vergleiche von Therapien mit dramatischen Effekten

Manchmal sprechen Patienten so drastisch auf eine Behandlung an, dass man auch ohne sorgfältig durchgeführte Tests zuverlässige Schlussfolgerungen ziehen kann (s. Kap. 5, S. 89–95).³ Bei einem Patienten mit einem Lungenkollaps (Pneumothorax) sorgt das Einführen einer Hohlnadel in den Brustraum und das Absaugen der eingeschlossenen Luft so unmittelbar für Linderung, dass die Vorteile dieser therapeutischen Maßnahme auf der Hand liegen. Andere Beispiele für dramatische Wirkungen sind die Gabe von Morphinum bei Schmerzen, die Gabe von Insulin bei einem diabetischen Koma und der künstliche Hüftgelenkersatz bei arthritsbedingten Schmerzen. Aber auch die unerwünschten Wirkungen von Behandlungen können dramatische Formen annehmen. So verursachen Medikamente manchmal schwerwiegende, sogar tödliche allergische Reaktionen; als Beispiel für dramatische Nebenwirkungen können auch die durch Thalidomid hervorgerufenen seltenen Fehlbildungen der Gliedmaßen angeführt werden (s. Kap. 1, S. 34).

Allerdings kommen solche dramatischen Wirkungen von Therapien, ganz gleich, ob positiver oder negativer Art, nur selten vor. Die meisten Therapieeffekte fallen schwächer aus. Trotzdem ist es gut, über sie Bescheid zu wissen. Beispielsweise bedarf es sorgfältig durchgeführter Tests, um festzustellen, welche Dosierschemata bei Morphinum wirksam und sicher sind, oder um herauszufinden, ob gentechnisch hergestelltes Insulin Vorteile gegenüber tierischem Insulin aufweist oder ob ein neu auf den Markt gebrachtes künstliches Hüftgelenk, das zwanzigmal teurer ist als die kostengünstigste Variante, diese Zusatzkosten – im Hinblick auf das, worauf es dem Patienten ankommt – auch wert ist. Unter diesen alltäglichen Um-

ständen müssen wir alle unfairen (verzerrten) Vergleiche vermeiden und mit ihnen auch die falschen Schlüsse, die sich daraus ergeben können.

Vergleiche von Therapien mit mäßigen, aber wichtigen Effekten

Vergleich gegenwärtig behandelter Patienten mit Patienten in der Vergangenheit

Manchmal vergleichen Wissenschaftler die Patienten, die sie gegenwärtig behandeln, mit augenscheinlich ähnlichen Patienten, die wegen derselben Krankheit früher anders behandelt wurden. Solche Vergleiche können verlässliche Erkenntnisse liefern, wenn die Behandlungseffekte dramatisch sind – beispielsweise wenn eine neue Therapie bewirkt, dass einige Patienten eine Krankheit überleben, die früher in der Regel tödlich verlaufen ist. Wenn die Unterschiede zwischen den Therapien jedoch nicht dramatisch sind, es sich jedoch trotzdem lohnt, sie zu kennen, können solche Vergleiche, bei denen «historische Kontrollen» verwendet werden, aber Probleme verursachen. Auch wenn Wissenschaftler statistische Anpassungen (Adjustierungen) und Analysen durchführen, um möglichst sicherzustellen, dass sie wirklich Gleiches mit Gleichem vergleichen, so ist es in diesen Analysen jedoch unmöglich, auch den in den Vergangenheit nicht erfassten, aber relevanten Charakteristika von Patienten Rechnung zu tragen. Schlussendlich können wir uns also niemals völlig sicher sein, dass wirklich Gleiches mit Gleichem verglichen wurde.

Diese Probleme lassen sich dadurch veranschaulichen, dass man die Ergebnisse zu ein und derselben Therapie vergleicht, die ähnlichen Patienten zu unterschiedlichen Zeitpunkten verabreicht wurde. Sehen wir uns dazu eine Analyse von 19 Patienten mit fortgeschrittenem Lungenkarzinom an. Darin wurde die jährliche Sterblichkeitsrate von ähnlichen Patienten verglichen, die mit exakt der gleichen Therapie, aber zu verschiedenen Zeitpunkten behandelt wurden. Obwohl man nur geringfügige Unterschiede in den Sterblichkeitsraten erwartet hätte, unterschieden sie sich beträchtlich voneinander: Die Sterblichkeitsraten schwankten zwischen ei-

ner Besserung um 24 % und einer Verschlechterung um 46 %.⁴ Diese Unterschiede waren eindeutig nicht darauf zurückzuführen, dass sich die Behandlung geändert hatte – sie war gleich geblieben – oder dass die Patienten sich nachweislich unterschieden – auch das traf nicht zu. Vielmehr spiegelten die unterschiedlichen Sterblichkeitsraten vermutlich entweder nicht erkannte Unterschiede zwischen den Patienten wider oder aber andere, im Zeitverlauf aufgetretene, aber nicht erfasste Veränderungen (z. B. bessere Pflege oder bessere Infektionskontrolle), die bei den Vergleichen nicht berücksichtigt werden konnten.

Vergleich zwischen anscheinend ähnlichen Patientengruppen

Um die Wirkungen von Therapien zu beurteilen, vergleicht man auch die Erfahrungen und Behandlungsergebnisse von augenscheinlich ähnlichen Patientengruppen, die im selben Zeitraum zufällig unterschiedliche Therapien erhalten haben. Aber auch diese Methode kann grob irreführend sein. Ähnlich wie bei den Vergleichen mit «historischen Kontrollen» besteht auch hier das Problem darin, dass man wissen muss, ob sich die Personengruppen, die unterschiedliche Therapien erhalten haben, bereits vor Einleitung der Behandlung hinreichend ähnlich waren, damit ein aussagekräftiger (valider) Vergleich überhaupt möglich ist. Wie bei den «historischen Kontrollen» können die Wissenschaftler auch hier wieder statistische Adjustierungen und Analysen durchführen, um dafür zu sorgen, dass diese Gleichheit gewährleistet ist. Das funktioniert aber nur dann, wenn relevante Charakteristika der Patienten in den Vergleichsgruppen erfasst und berücksichtigt wurden. Diese Bedingungen sind aber nur sehr selten erfüllt, sodass solche Analysen stets mit einer gewissen Vorsicht zu genießen sind. Ihnen einfach zu glauben kann echte Tragödien auslösen.

Ein aufschlussreiches Beispiel dafür ist die Hormonersatztherapie (HRT). Frauen, die während und nach der Menopause mit HRT behandelt worden waren, wurden mit anscheinend ähnlichen Frauen ohne HRT verglichen. Diese Vergleiche ließen darauf schließen, dass HRT das Herzinfarkt- und Schlaganfallrisiko verringerte: Das wäre eine willkommene Nachricht gewesen, wenn sie sich bewahr-

heitet hätte. Leider war dies nicht der Fall. Spätere Vergleiche, die vor Beginn der Therapie entwickelt worden waren, um sicherzustellen, dass die Vergleichsgruppen ähnlich sind, ergaben, dass die HRT genau den gegenteiligen Effekt hatte – tatsächlich erhöhte sie die Anzahl der Herzinfarkte und Schlaganfälle (s. Kap. 2, S. 49f.). In diesem Fall war der offensichtliche Unterschied in den Herzinfarkt- und Schlaganfallraten auf den Umstand zurückzuführen, dass die Frauen mit HRT generell gesünder waren als die Frauen, die keine HRT erhielten – es hatte nichts mit der HRT an sich zu tun. Wissenschaftliche Untersuchungen, bei denen nicht gewährleistet ist, dass Gleiches wirklich mit Gleichem verglichen wird, können dazu führen, dass Zehntausende von Menschen Schaden erleiden.

Wie die Erfahrungen mit der HRT zeigen, wird der Vergleich von Gleichem mit Gleichem am besten dadurch gewährleistet, dass die Vergleichsgruppen noch vor Therapiebeginn zusammengestellt werden. Die Gruppen müssen aus Patienten bestehen, die sich nicht nur im Hinblick auf die bekannten und zu untersuchenden Faktoren wie Alter und Krankheitsschweregrad ähnlich sind, sondern auch in Bezug auf Faktoren, die zwar nicht untersucht werden, die aber dennoch einen Einfluss auf die Genesung von einer Erkrankung ausüben können. Das können die Ernährung, der Beruf und andere soziale Faktoren oder Angst vor der Krankheit oder vor den vom Arzt vorgeschlagenen Behandlungen sein. Es ist immer schwierig und häufig sogar unmöglich, sicher zu sein, dass sich die Behandlungsgruppen ähneln, wenn sie erst nach Therapiebeginn zusammengestellt werden.

Die kritische Frage lautet also: Spiegeln Unterschiede in den Behandlungsergebnissen Unterschiede in den Wirkungen der miteinander verglichenen *Behandlungen* wider oder Unterschiede zwischen den *Patienten* in den Vergleichsgruppen?

Unverzerrte prospektive Zuteilung zu verschiedenen Therapien

Im Jahre 1854 demonstrierte Thomas Graham Balfour, ein in einem Militärwaisenhaus tätiger Armeearzt, wie man Behandlungsgruppen bilden kann, um zu gewährleisten, dass Gleiches mit Gleichem

verglichen wird. Balfour wollte herausfinden, ob das aus der Schwarzen Tollkirsche gewonnene «Belladonna», wie manche Leute behaupteten, Kinder vor Scharlach schützen könne. Um also «zu vermeiden, dass ihm Selektion unterstellt würde», wie er es formulierte, teilte er die Kinder *abwechselnd* (*alternierend*) entweder der Gruppe zu, die das Medikament erhalten sollte, oder der Gruppe, die das Medikament nicht erhielt.⁵ Die Anwendung dieser alternierenden Zuteilung (Allokation) oder irgendeiner anderen unverzerrten Art der Vergleichsgruppenbildung gehört zu den Schlüsselmerkmalen fairer Therapietests. Es erhöht die Wahrscheinlichkeit, dass sich die Vergleichsgruppen ähnlich sind, und zwar nicht nur hinsichtlich der bekannten und zu untersuchenden Faktoren, sondern auch in Bezug auf die nicht untersuchten Faktoren, welche die Genesung von einer Krankheit beeinflussen und für die unmöglich statistische Adjustierungen vorgenommen werden können.

Um eine faire (unverzerrte) Zuteilung zu verschiedenen Therapien zu erreichen, müssen diejenigen, die solche fairen Tests entwerfen, unbedingt dafür sorgen, dass Ärzte und Patienten nicht wissen oder vorhersehen können, wie die nächste Zuteilung aussehen wird. Wenn sie diese Zuteilung nämlich kennen, könnten sie – bewusst oder unbewusst – in Versuchung kommen, eine bestimmte Therapie auszuwählen. Wenn ein Arzt beispielsweise weiß, dass der nächste für die Teilnahme an einer klinischen Studie vorgesehene Patient ein Placebo (eine Scheinbehandlung) erhalten wird, könnte er versuchen, einen schwerer erkrankten Patienten von der Teilnahme an der Studie abzubringen, und auf einen weniger kranken Patienten warten. Selbst wenn also ein unverzerrtes Zuteilungsschema erstellt wurde, findet eine unverzerrte Zuteilung zu den Behandlungsgruppen nur statt, wenn die anstehenden Zuteilungen in diesem Schema erfolgreich vor den Personen geheim gehalten werden, die darüber entscheiden, ob ein Patient an der Studie teilnimmt oder nicht. Auf diese Weise kann niemand vorhersagen, welche Behandlung als nächstes zugeteilt wird, und es wird niemand dazu verführt, vom unverzerrten Zuteilungsschema abzuweichen.

Diese verdeckte Behandlungszuteilung wird normalerweise erstens durch die Erstellung von Zuteilungsschemata erreicht, die weniger vorhersehbar sind als die einfache wechselweise Zuordnung –

zum Beispiel dadurch, dass die Zuteilung auf Zufallszahlen beruht – und zweitens durch Geheimhaltung dieses Schemas. Um die Zuteilungsschemata geheim zu halten, kann man verschiedene Methoden anwenden. So kann die Zufallszuteilung beispielsweise bei einem Patienten, der sich für die Teilnahme an der Studie eignet, «aus der Ferne» – d.h. per Telefon oder Computer – erfolgen (**Abb. 8**). Eine weitere Möglichkeit stellt die Verwendung einer Serie von nummerierten Umschlägen dar, die jeweils eine Zuteilung enthalten. Ist ein Patient für die Teilnahme geeignet, wird jeweils der nächste Umschlag in dieser Serie geöffnet, um zu sehen, welcher Gruppe er zugeweiht wird. Damit dieses System funktioniert, müssen die Umschläge undurchsichtig sein, damit die Ärzte nicht «schummeln» können, indem sie den Umschlag gegen das Licht halten, um die Zuteilung darin zu erkennen.

Dieses Vorgehen gilt heutzutage als wesentliches Merkmal fairer Therapietests. Studien, in denen für die Zuteilung zu den Therapien das Zufallsprinzip verwendet wird, heißen «randomisierte Studien» (s. den Kasten in Kap. 3, S. 61).



Abbildung 8: Verdeckung der Behandlungszuteilung in einer Studie durch telefonische Randomisierung.

Möglichkeiten der unverzerrten (zufälligen) Zuteilung in Therapievergleichen

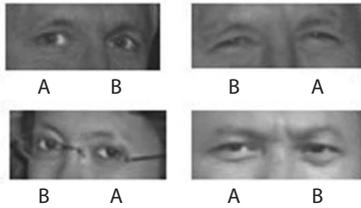
Die Zufallszuteilung zu Therapievergleichen kann unterschiedlich eingesetzt werden (**Abb. 9**). So kann sie zum Beispiel dazu verwendet werden, um unterschiedliche Therapien, die derselbe Patient zu unterschiedlichen Zeiten in einer zufälligen Reihenfolge erhält, miteinander zu vergleichen – in diesem Fall spricht man von einer «randomisierten Studie mit Crossover-Design». Um beispielsweise zu beurteilen, ob einem einzelnen Patienten mit anhaltendem trockenem Husten die Inhalation eines Medikaments hilft, könnte die Studie für die Dauer von mehreren Monaten angelegt werden. In einigen zufällig ausgewählten Wochen würde der Patient einen medikamentenhaltigen Inhalator benutzen, und während der übrigen Wochen bekäme er einen identisch aussehenden Inhalator, der kein Medikament enthält. Die Erkenntnisse aus der Forschung auf diese Weise maßgeschneidert auf einzelne Patienten anzuwenden, ist, falls machbar, natürlich wünschenswert. Aber es gibt auch viele Situationen, in denen solche Crossover-Studien einfach nicht möglich sind. So kann man beispielsweise weder unterschiedliche chirurgische Eingriffe auf diese Weise miteinander vergleichen, noch die Behandlungen bei einmaligen akuten Gesundheitsproblemen wie z. B. starken Blutungen nach einem Autounfall.

Die Zufallszuteilung kann auch dazu eingesetzt werden, um verschiedene Therapien zu vergleichen, die bei demselben Patienten an unterschiedlichen Körperteilen angewendet werden. Bei einer Hautkrankheit etwa wie einem Ekzem oder der Schuppenflechte (Psoriasis) kann nach dem Zufallsprinzip ausgewählt werden, welche der befallenen Hautflächen mit einer medikamentenhaltigen Salbenzubereitung behandelt werden sollen und welche mit einer Salbenzubereitung ohne aktive Wirkstoffe. Oder bei der Behandlung einer beidseitigen Augenerkrankung könnte ein Auge nach dem Zufallsprinzip zur Behandlung ausgewählt und das Ergebnis anschließend mit dem unbehandelten Auge verglichen werden.

Eine weitere Anwendung der Zufallszuteilung ist der Vergleich unterschiedlicher Therapien, die verschiedene Populationen oder Personengruppen erhalten, beispielsweise alle Personen, die be-

	Mo	Di	Mi	Do	Fr
1. Maiwoche	A	B	A	B	B
2. Maiwoche	A	A	B	B	A
3. Maiwoche	B	A	B	A	A
4. Maiwoche	B	A	B	B	A
5. Maiwoche	A	B	A	B	B

Randomisierung von Tagen im Leben einer Person



Randomisierung von rechtem und linkem Auge von Personen



Randomisierung von Gemeinden innerhalb einer bestimmten Region



Randomisierung von Einzelpersonen innerhalb einer Gruppe

Abbildung 9: Verschiedene Möglichkeiten der randomisierten Zuteilung.

stimmte Grundversorgungskliniken oder Krankenhäuser aufsuchen. Solche Vergleiche nennt man «cluster- (oder gruppen-) randomisierte Studien». Um beispielsweise die Auswirkungen des allgemeinen Krankenversicherungsprogramms in Mexiko zu beurteilen, untersuchten Wissenschaftler nach der Matched-Pairs-Technik 74 vergleichbare (gematchte) Patienten-Paare aus den medizinischen Einzugsgebieten – Cluster, die zusammen 118 000 Haushalte in sieben Bundesstaaten repräsentierten. Pro gematchtem Paar wurde nach dem Zufallsprinzip jeweils einer dem Versicherungsprogramm zugeteilt.⁶

Die bei weitem häufigste Anwendung der Zufallszuteilung dient jedoch der Entscheidung, welcher Patient welche Behandlung erhalten soll.

Nachverfolgung aller Teilnehmer an Therapievergleichen

Nachdem man sich die Mühe gemacht hat, Vergleichsgruppen zusammenzustellen, um zu gewährleisten, dass Gleiches mit Gleichem verglichen wird, ist es wichtig, systematischen Fehlern (Bias) vorzubeugen, die auftreten würden, wenn man bei der Auswertung der Studie diejenigen Patienten ignorieren würde, die einen von der tatsächlichen Planung abweichenden Studienverlauf hatten. Soweit möglich sollten alle den Vergleichsgruppen zugeteilten Patienten nachverfolgt und in der Hauptanalyse der Ergebnisse derjenigen Gruppe berücksichtigt werden, der sie anfangs zugeteilt wurden, unabhängig davon, welche Behandlung sie später (wenn überhaupt) tatsächlich erhalten haben. Dies nennt man eine «Intention-to-Treat»-Analyse. Geschieht dies nicht, ist der Vergleich von Gleichem mit Gleichem nicht mehr gewährleistet.

Auf den ersten Blick mag es unlogisch erscheinen, Gruppen miteinander zu vergleichen, in denen einige Patienten nicht die Therapien erhalten haben, denen sie zugeteilt wurden. Missachtet man diesen Grundsatz jedoch, können die Studien unfair werden und die Ergebnisse in die Irre führen. Patienten beispielsweise, die an einer teilweisen Verstopfung der hirnversorgenden Blutgefäße mit Schwindelanfällen leiden, haben ein überdurchschnittlich hohes Schlaganfallrisiko. Wissenschaftler haben eine Studie durchgeführt, um herauszufinden, ob eine Operation zur Beseitigung der Gefäßverstopfung bei diesen Patienten die Zahl späterer Schlaganfälle verringern kann. Zu Recht verglichen sie alle Patienten, die der Operation zugeteilt worden waren, und zwar unabhängig davon, ob sie den Eingriff überlebten oder nicht, mit all den Patienten, die der Gruppe ohne Operation zugeteilt worden waren. Hätten sie die Häufigkeit der Schlaganfälle nur bei den Patienten erfasst, welche die unmittelbaren Auswirkungen der Operation überlebten, hätten sie die wichtige Tatsache übersehen, dass die Operation selbst einen Schlaganfall verursachen und zum Tod führen kann und dass unter sonst gleichen Umständen bei den überlebenden Patienten in dieser Gruppe weniger Schlaganfälle auftreten. Das wäre ein unfairer Test der Operationswirkungen gewesen, weil deren Risiken mit in die Bewertung einfließen müssen.

Die in der **Abbildung 10** dargestellten Ergebnisse der operativen und der medikamentösen Therapie sind tatsächlich gleichwertig. Wenn jedoch die beiden der Operation zugeteilten Personen vor dem Eingriff versterben und bei der Auswertung unberücksichtigt bleiben, wird der Vergleich der beiden Gruppen verzerrt. Er würde fälschlich darauf schließen lassen, dass die Operation besser ist.

Wie man mit Abweichungen von der Behandlungszuteilung umgeht

Aus all den bisher in diesem Kapitel genannten Gründen wird deutlich, dass faire Therapietests sorgfältig geplant werden müssen. Die Dokumente, in denen diese Pläne ausgeführt werden, nennt man Studienprotokolle. Doch selbst die besten Pläne funktionieren nicht immer wie beabsichtigt: Manchmal weichen die Therapien, welche die Patienten tatsächlich erhalten, von den ursprünglich zugeteilten Therapien ab. Beispielsweise kann es vorkommen, dass Patienten

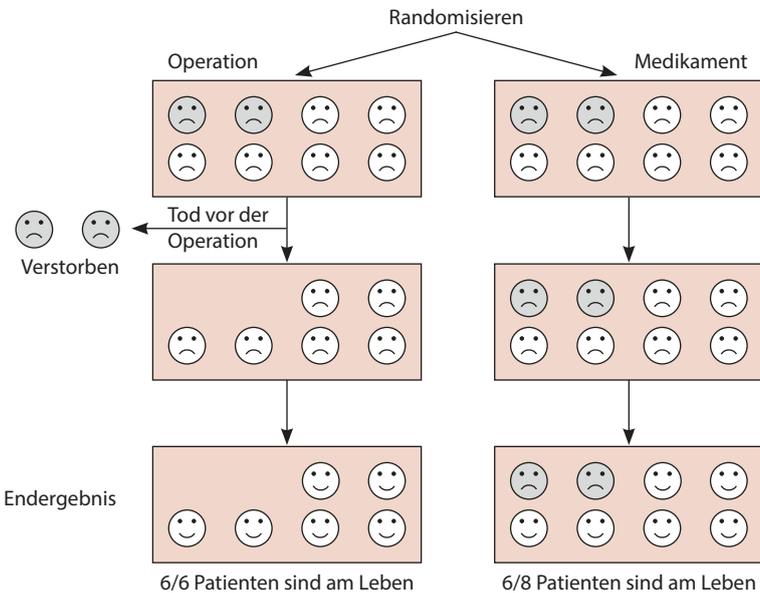


Abbildung 10: Warum alle randomisierten Patienten im Endergebnis berücksichtigt werden sollten («Intention-to-Treat»).

die Behandlungen nicht wie vorgesehen durchführen oder dass eine der Behandlungen nicht verabreicht wird, weil das benötigte Material oder das Personal dafür nicht verfügbar sind. Wenn solche Diskrepanzen bemerkt werden, muss man die Folgen berücksichtigen und umsichtig darauf eingehen.

Während der 1970er- und 1980er-Jahre erzielte man bemerkenswerte Fortschritte in der Behandlung von Kindern mit akuter lymphatischer Leukämie, dem in dieser Altersgruppe am häufigsten vorkommenden Leukämietyp. Allerdings war es ein Rätsel, warum amerikanische Kinder dabei erheblich besser abschnitten als britische Kinder, die allem Anschein nach exakt dieselbe Medikamentenabfolge erhielten.⁷ Bei einem Besuch in einem kalifornischen Kinderkrebszentrum stellte ein aufmerksamer britischer Statistiker fest, dass amerikanische Kinder mit Leukämie deutlich «aggressiver» mit Chemotherapie behandelt wurden als die betroffenen Kinder in Großbritannien. Die Therapie hatte fürchterliche Nebenwirkungen (Übelkeit, Infektion, Anämie, Haarausfall usw.). Wenn diese Nebenwirkungen für die Kinder allzu belastend wurden, dann neigten britische Ärzte und Krankenschwestern im Gegensatz zu ihren amerikanischen Kollegen dazu, die verordnete Therapie zu reduzieren oder zu pausieren. Dieses «sanftere Vorgehen» schien die Wirksamkeit der Therapie verringert zu haben und war wahrscheinlich ein Grund für den unterschiedlichen Behandlungserfolg in Amerika und Großbritannien.

Den Patienten bei der Befolgung der zugewiesenen Therapien helfen

Auch in anderer Hinsicht kann es zu Abweichungen zwischen den beabsichtigten und den tatsächlich verabreichten Behandlungen kommen, welche die Auswertung von Therapietests erschweren können. Den Teilnehmern an wissenschaftlichen Studien sollten medizinisch notwendige Behandlungen nicht vorenthalten werden. Wenn eine neue Therapie mit erhofften, aber unbewiesenen günstigen Wirkungen in einem fairen Test untersucht wird, sollte man deshalb allen teilnehmenden Patienten versichern, dass sie in jedem Fall alle bewährte und wirksame Therapien erhalten.

Wenn man weiß, wer in einer Studie wie behandelt wird, können sich daraus mehrere systematische Fehler ergeben. Ein solcher Fehler ist z. B., dass Patienten und Ärzte das Gefühl haben können, dass Teilnehmer, die der «neuen» Therapie zugeteilt sind, Glück haben, und dies kann sie unbewusst zu einer Überbewertung der Vorzüge dieser Therapie verleiten. Andererseits sind Patienten und Ärzte vielleicht der Meinung, dass die den «älteren» Therapien zugeteilten Teilnehmer benachteiligt sind, sodass ihre Enttäuschung dazu führen kann, dass sie etwaige positive Effekte unterbewerten. Das Wissen um die Zuteilung der Behandlungen kann Ärzte zudem dazu veranlassen, den Patienten, die den älteren Therapien zugeteilt wurden, eine zusätzliche Behandlung oder Pflege angeeignet zu lassen, um sie gewissermaßen für die Tatsache zu entschädigen, dass sie nicht einer der neueren, wenn auch ungeprüften Therapien zugeteilt wurden. Solche zusätzlichen Behandlungen bei Patienten in der einen, aber nicht in der anderen Vergleichsgruppe anzuwenden, erschwert die Beurteilung einer neuen Therapie und birgt die Gefahr, dass der Vergleich dadurch unfair wird und die Ergebnisse verfälscht werden. Eine Möglichkeit, wie man Unterschiede in Vergleichen zwischen beabsichtigten und tatsächlich verabreichten Therapien vermindern kann, ist dafür zu sorgen, dass die neueren und älteren Therapien möglichst gleich aussehen, gleich schmecken und gleich riechen.

Und genau das passiert beim Vergleich zwischen einer Therapie mit erhofften nützlichen Wirkungen und einer Therapie ohne aktive Wirkstoffe (einer Scheinbehandlung oder einem Placebo), die genauso aussieht, riecht, schmeckt und sich anfühlt wie die «richtige» Therapie. Ein solches Vorgehen nennt man «Verblindung» oder «Maskierung». Wenn diese «Verblindung» erreicht werden kann (und es gibt viele Fälle, in denen das nicht möglich ist), dann unterscheiden sich die Patienten in den beiden Vergleichsgruppen im Grunde nur in einer Hinsicht – und zwar darin, ob sie der neuen Therapie oder der Therapie ohne aktive Wirkstoffe zugeteilt wurden. Dadurch können auch die Ärzte und Pflegekräfte, die die Studienteilnehmer versorgen, weniger gut erkennen, ob ihre Patienten die neue Therapie erhalten oder nicht. Wenn weder Ärzte noch Patienten wissen, welche Therapie verabreicht wird, spricht man von

einer «Doppelblindstudie». Infolgedessen haben die Patienten in beiden Vergleichsgruppen eine ähnlich starke Motivation, die ihnen zugeteilte Behandlung zu befolgen. Zudem nimmt auch die Wahrscheinlichkeit zu, dass die Ärzte, die sich um sie kümmern, alle Patienten gleich behandeln.

Faire Erfassung des Behandlungsergebnisses

Einer der Gründe für die Anwendung von Scheinbehandlungen in Therapievergleichen ist zwar, Patienten und Ärzte bei der Befolgung der jeweils zugeteilten Behandlung zu unterstützen, doch es gibt noch einen triftigeren Grund für eine solche «Verblindung», und zwar die Reduktion von systematischen Fehlern bei der Auswertung der Behandlungsergebnisse.

Die Methode der Verblindung hat eine interessante Geschichte. So ordnete z. B. im 18. Jh. Ludwig XVI. von Frankreich an, Anton Mesmers Behauptung zu prüfen, dass der «animalische Magnetismus» (gelegentlich auch als «Mesmerismus» bezeichnet) vorteilhafte Wirkungen habe. Der König wollte wissen, ob die Wirkungen auf eine «echte (physikalische) Kraft» oder eher auf Kräfte der «Imagination» (Einbildung) zurückzuführen seien. In einem Behandlungstest teilte man den Teilnehmern, denen die Augen verbunden worden waren, mit, dass sie mit animalischem Magnetismus behandelt bzw. nicht damit behandelt würden, wobei gelegentlich genau das Gegenteil erfolgte. Die Teilnehmer gaben nur dann an, dass sie die Wirkungen der «Behandlung» gespürt hätten, wenn ihnen vorher gesagt worden war, dass sie diese Behandlung erhalten würden.

Eine Verzerrung der Ergebnisbewertung ist bei manchen Ergebnissen, die in der Studie gemessen werden (Outcomes) – beispielsweise Überleben – sehr unwahrscheinlich, da kaum Zweifel darüber aufkommen dürften, ob ein Patient gestorben ist oder nicht. Bei den meisten Ergebnissen dürfte die Bewertung allerdings mit einer gewissen Subjektivität einhergehen, weil die zu messenden Studienergebnisse auch das Auftreten von Symptomen wie Schmerzen und Angst bei den Patienten beinhalten sollten und dies oftmals auch tun.

Unter diesen nicht selten auftretenden Umständen stellt die Verblindung ein wünschenswertes Merkmal fairer Tests dar. In einer Studie über Therapien der multiplen Sklerose beispielsweise wurden alle Patienten zum einen von einem Arzt untersucht, der nicht wusste, ob die Patienten die neuen Medikamente oder ein Medikament ohne aktiven Wirkstoff erhalten hatten (d.h., der Arzt war «verblindet»), dann aber auch von einem Arzt, der über die Zuteilung der Patienten zur jeweiligen Vergleichsgruppe informiert war (d.h., der Arzt war «unverblindet»). Die von den «verblindeten» Ärzten durchgeführten Untersuchungen legten nahe, dass die neue Therapie nicht nützlich war, während die Beurteilungen der «unverblindeten» Ärzte darauf hindeuteten, dass die neue Therapie Vorteile brachte.⁸ Dieser Unterschied in der Beurteilung lässt darauf schließen, dass die neue Therapie nicht wirksam war und dass die Kenntnis der Behandlungszuteilung die «unverblindeten» Ärzte dazu verleitet hatte, das «zu sehen, woran sie glaubten» oder worauf sie hofften. Je stärker also die Komponente der Subjektivität bei der Beurteilung von Behandlungsergebnissen ist, umso wünschenswerter ist die Verblindung, um eine faire Therapiestudie zu gewährleisten.

Manchmal gelingt es sogar, Patienten darüber im Unklaren zu lassen, ob bei ihnen ein echter chirurgischer Eingriff vorgenommen wurde oder nicht. Dazu wurde eine Studie bei Patienten mit Kniegelenkarthrose durchgeführt. Hierbei zeigte der operative Eingriff, bei dem das arthrotisch veränderte Gelenk gespült und gereinigt wurde, keinen eindeutigen Vorteil gegenüber einem Vorgehen, bei dem unter Narkose oberhalb des Knies lediglich ein Hautschnitt gesetzt und «nur so getan wurde, als ob» der Gelenkraum danach gespült und gereinigt worden wäre.⁹

Häufig ist es aber schlicht unmöglich, Patienten und Ärzte gegenüber den jeweils zu vergleichenden Therapien zu verblinden – beispielsweise wenn eine chirurgische mit einer medikamentösen Therapie verglichen wird oder wenn ein Medikament eine charakteristische Nebenwirkung aufweist. Aber selbst bei Ergebnissen, bei denen sich systematische Fehler (Bias) einschleichen könnten – etwa bei der Feststellung der Todesursache oder bei der Beurteilung einer Röntgenaufnahme –, ist dies vermeidbar, wenn man dafür sorgt,

dass die Auswertung dieser Ergebnisse unabhängig durch Personen erfolgt, die nicht wissen, welche Behandlungen die einzelnen Patienten erhalten haben.

Untersuchungen zu unerwünschten Therapiewirkungen

Wie Vermutungen bezüglich unerwarteter Behandlungswirkungen entstehen

Ärzte bzw. Patienten sind häufig die ersten, bei denen der Verdacht auf unerwartete (ob nun positive oder negative) Behandlungseffekte aufkommt.¹⁰ Da die für eine Marktzulassung benötigten Therapiestudien nur wenige hundert bzw. wenige tausend Menschen einschließen, die über einige Monate behandelt werden, können meist nur vergleichsweise kurzzeitige und häufig auftretende Nebenwirkungen erfasst werden. Seltene Wirkungen und solche, die sich erst nach einer gewissen Zeit entwickeln, werden erst entdeckt, wenn die Therapien in größerem Umfang, über einen längeren Zeitraum und bei einem breiteren Patientenspektrum angewendet werden als bei den Teilnehmern der Zulassungsstudien.

In immer mehr Ländern – darunter auch in Großbritannien, den Niederlanden, Schweden, Dänemark, den USA und Deutschland – haben Ärzte und Patienten die Möglichkeit, mutmaßliche unerwünschte Arzneimittelwirkungen zu melden, denen dann offiziell nachgegangen werden kann.¹¹ In Deutschland ist hierfür das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) zuständig. Auch wenn sich keines dieser Meldesysteme bei der Identifizierung relevanter unerwünschter Reaktionen auf Arzneimittel besonders hervorgetan hat, so ließen sich in einzelnen Fällen doch Erfolge verbuchen. Als beispielsweise der Cholesterinsenker Rosuvastatin 2003 in Großbritannien auf den Markt kam, zeichnete sich aus den eingehenden Meldungen bald eine seltene schwerwiegende, unerwartete Nebenwirkung auf die Muskulatur ab, die sogenannte Rhabdomyolyse. Bei diesem Krankheitsbild kommt es zu einer ra-

schen Auflösung von Muskelfasern, deren Abbauprodukte schwerwiegende Nierenschäden hervorrufen können. Wie die weitere Untersuchung ergab, bestand das größte Risiko für diese Komplikation bei Patienten, die das Medikament in hohen Dosierungen einnahmen.

Untersuchung von Vermutungen über unerwartete Behandlungswirkungen

Oft stellen sich Vermutungen über unerwünschte Wirkungen als falscher Alarm heraus.¹⁰ Wie also soll man herausfinden, ob es sich um echte Nebenwirkungen handelt? Tests, mit denen man mutmaßliche unerwartete Wirkungen bestätigen oder verwerfen kann, müssen dieselben Prinzipien befolgen wie Tests, die der Identifizierung von erhofften erwarteten Therapieeffekten dienen. Und das bedeutet, dass man verzerrte Vergleiche vermeiden muss, um sicherzustellen, dass «Gleiches mit Gleichem verglichen wird», und es muss eine angemessene Anzahl von Fällen untersucht werden.

Das Yellow-Card-System

Das Yellow Card Scheme wurde 1964 in Großbritannien eingeführt, nachdem die Thalidomid-Tragödie gezeigt hatte, wie wichtig es ist, Probleme zu beobachten, die nach der Zulassung eines Medikaments auftreten. Empfänger dieser Meldung ist die Medicines and Healthcare Products Regulatory Agency (MHRA), welche die Ergebnisse auswertet. Jedes Jahr gehen bei der MHRA mehr als 20 000 Meldungen zu möglichen Nebenwirkungen ein. Anfangs konnten nur Ärzte Meldung erstatten, später wurden dann auch Pflegekräfte, Apotheker, Gerichtsmediziner, Zahnärzte, Röntgenassistenten und Optiker einbezogen. Seit 2005 haben auch Patienten und Betreuungspersonen die Möglichkeit, mutmaßliche unerwünschte Reaktionen zu melden. Die Meldungen können online unter www.yellowcard.gov.uk, per Post oder telefonisch erfolgen.

Eine Patientin fasste ihre Erfahrungen wie folgt zusammen: «Nebenwirkungen über das Yellow-Card-System melden zu können, gibt mir ein Gefühl der Kontrolle. Es bedeutet, dass man die Meldung selber machen kann, ohne darauf warten zu müssen, dass ein vielbeschäftigter Arzt das übernimmt ... Damit rückt der Patient in den Mittelpunkt der medizinischen Versorgung. Es ist ein Quantensprung in der Patiententeilhabe; es weist den Weg nach vorn und kennzeichnet einen grundlegenden Einstellungswandel.»

Bowser A. A patient's view of the Yellow Card Scheme. In: *Medicines & Medical Devices Regulation: what you need to know*. London: MHRA, 2008.

Abrufbar unter www.mhra.gov.uk

Wie bei den erhofften Behandlungseffekten kann man auch bei den unerwarteten Wirkungen die dramatischen leichter erkennen und bestätigen als die weniger dramatischen. Wenn das mutmaßliche unerwartete Behandlungsergebnis normalerweise nur selten auftritt, nach Anwendung einer Therapie aber recht häufig vorkommt, wird es sowohl dem Arzt als auch dem Patienten auffallen, dass etwas nicht stimmt. Gegen Ende des 19. Jh. erfuhr der Schweizer Chirurg Theodor Kocher von einem Hausarzt, dass eines der Mädchen, bei denen Kocher einige Jahre zuvor eine vergrößerte Schilddrüse (Struma) entfernt hatte, stumpfsinnig und lethargisch geworden war. Als er sich näher mit diesem Fall und anderen früheren Struma-Patienten befasste, die er operiert hatte, entdeckte er, dass die komplette Entfernung der vergrößerten Schilddrüse zu Kretinismus und Myxödemem geführt hatte – seltene, aber schwerwiegende Komplikationen, die, wie wir heute wissen, durch den Mangel an dem von der Schilddrüse produzierten Hormon bedingt sind.¹² Auch hinsichtlich der unerwarteten Wirkungen von Thalidomid (s. Kap. 1, S. 34) gab es Vermutungen, die bestätigt wurden, weil der Zusammenhang zwischen der Einnahme des Medikaments in der Schwangerschaft und der Geburt von Kindern ohne Gliedmaßen so dramatisch war. Von solchen Anomalien hatte man bis dahin noch nie gehört.

Weniger dramatische unerwartete Behandlungseffekte kommen manchmal in randomisierten Studien ans Licht, die auf die Beurteilung der jeweiligen Vorteile alternativer Therapien ausgelegt sind. Ein randomisierter Vergleich von zwei Antibiotika, die Neugeborenen zur Vorbeugung gegen Infektionen verabreicht wurden, ergab, dass eines der Medikamente die Verarbeitung von Bilirubin, einem Abbauprodukt aus der Leber, im Körper beeinträchtigte. Die Anreicherung dieses Abbauprodukts im Blut führte bei den Säuglingen, die das eine der beiden Vergleichsantibiotika erhalten hatten, zu Hirnschädigungen.¹³

Gelegentlich kann auch die weitere Auswertung von in der Vergangenheit durchgeführten randomisierten Studien zur Identifizierung von weniger dramatischen unerwünschten Wirkungen beitragen. Nachdem der Nachweis erbracht worden war, dass die Verabreichung des Medikaments Diethylstilbestrol (DES) an

Schwangere bei den Töchtern einiger der betroffenen Frauen die Ursache für eine seltene Krebserkrankung war, kam es zu Spekulationen über weitere potenzielle unerwünschte Wirkungen. Diese wurden entdeckt, indem man Kontakt zu den Söhnen und Töchtern der Frauen aufnahm, die an kontrollierten Studien teilgenommen hatten. In diesen Nachbeobachtungsstudien (Follow-up-Studien) fanden sich sowohl bei Männern als auch Frauen genitale Anomalien und Unfruchtbarkeit (Infertilität). Als in der jüngeren Vergangenheit bei Rofecoxib (Vioxx), einem neuen Arthritismedikament, der Verdacht aufkam, dass es Herzinfarkte auslösen könne, wurde ebenfalls durch eine detaillierte Untersuchung der Ergebnisse relevanter randomisierter Studien nachgewiesen, dass das Medikament diese unerwünschte Wirkung tatsächlich aufwies (s. Kap. 1, S. 37).¹⁴

Die Nachbeobachtung von Patienten, die an randomisierten Studien teilgenommen haben, bietet offensichtlich eine erfolgversprechende Möglichkeit, um sicherzustellen, dass, wenn man Vermutungen über unerwartete Therapieeffekte nachgeht, auch wirklich Gleiches mit Gleichem verglichen wird. Leider besteht diese Option aber nur selten, es sei denn, im Voraus wurden entsprechende Vorkehrungen getroffen. Die Untersuchung von Mutmaßungen über mögliche unerwünschte Therapieeffekte wäre eine weniger große Herausforderung, wenn die Kontaktdaten der Teilnehmer an randomisierten Studien routinemäßig erfasst würden. Dann könnten die Betroffenen erneut kontaktiert und nach weiteren Angaben zu ihrem Gesundheitszustand befragt werden.

Die Untersuchung mutmaßlicher unerwünschter Wirkungen von Behandlungen wird leichter, wenn die unter Verdacht stehenden unerwünschten Wirkungen ein ganz anderes gesundheitliches Problem betreffen als das, wofür die Therapie verordnet wurde.¹⁵ Als Dr. Spock beispielsweise empfahl, Babys zum Schlafen auf den Bauch zu legen, betraf seine Empfehlung *alle* Babys und nicht diejenigen, für die ein überdurchschnittlich hohes Risiko für plötzlichen Kindstod angenommen wurde (s. Kap. 2, S. 45f.). Der fehlende Zusammenhang zwischen der ärztlichen Empfehlung («Babys zum Schlafen auf den Bauch legen») und der mutmaßlichen Folge dieser Empfehlung (plötzlicher Kindstod) trug dazu bei, die Schlussfolge-

nung zu untermauern, dass sich in dem beobachteten Zusammenhang (Assoziation) zwischen dem ärztlichen Rat und dem plötzlichen Kindstod Ursache und Wirkung widerspiegeln.

Eine deutlich größere Herausforderung stellt im Gegensatz dazu die Untersuchung der Vermutung dar, dass wegen Depression verordnete Medikamente zu einem vermehrten Auftreten von suizidalen Gedanken führen, die eine Depression gelegentlich begleiten können. Ohne randomisierte Studien, in denen die verdächtigsten Medikamente mit anderen Therapien für Depressionen verglichen werden, ist es kaum gerechtfertigt anzunehmen, dass sich die Personen, welche diese Medikamente eingenommen haben, und diejenigen, die sie nicht eingenommen haben, hinreichend ähnlich sind, um sie zuverlässig miteinander vergleichen zu können.¹⁶

- Wir brauchen faire Therapietests, weil wir andernfalls gelegentlich zu dem Schluss kämen, dass Therapien wirksam sind, wenn eigentlich das Gegenteil der Fall ist, und umgekehrt.
- Zu den wesentlichen Grundlagen aller fairen Therapietests gehören Vergleiche.
- Wenn Therapien miteinander (oder eine Behandlung mit keiner Behandlung) verglichen werden, dann kommt dabei dem Grundsatz «Gleiches mit Gleichem vergleichen» eine wichtige Rolle zu.
- Wir müssen versuchen, systematische Fehler (Bias) bei der Beurteilung von Behandlungsergebnissen einzugrenzen.

7

Den Zufallsfaktor berücksichtigen

Der Zufallsfaktor und das Gesetz der großen Zahl

Die Glaubwürdigkeit der Erkenntnisse (Evidenz) über Therapieeffekte beruht auf der Vermeidung von systematischen Fehlern – und dem richtigen Umgang mit den Fehlern, die nicht verhindert werden konnten. Wenn diese Eigenschaften fairer Tests nicht gewährleistet sind, dann kann auch keinerlei angemessene Bearbeitung der Forschungsdaten die verbleibenden Probleme mit ihren gefährlichen – und gelegentlich auch tödlichen – Folgen lösen (s. Kap. 1 und 2). Und selbst wenn die zur Minderung von Bias unternommenen Schritte erfolgreich sind, besteht immer noch die Möglichkeit, dass der Zufall uns in die Irre führt.

Wie wir alle wissen, kommt es beim Münzenwerfen nicht selten vor, dass man nacheinander fünfmal oder noch öfter ganze Serien von Kopf oder Zahl wirft. Und wie wir ebenfalls wissen: Je öfter man eine Münze wirft, umso wahrscheinlicher ist es, dass am Ende gleich oft Kopf und Zahl geworfen wurden.

Wenn zwei Therapien verglichen werden, dann können etwaige Unterschiede in den Ergebnissen auch einfach auf einem solchen Zufall beruhen. Nehmen wir einmal an, 40 % der Patienten versterben, nachdem sie Therapie A erhalten haben, während in der Vergleichsgruppe nach Anwendung von Therapie B 60% versterben. **Tabelle 1** zeigt, was man erwarten würde, wenn jeweils zehn Patienten eine der beiden Therapien erhielten. Der Unterschied in der Anzahl der Todesfälle zwischen den beiden Therapien wird als «Risikoverhältnis» (engl. *risk ratio*) ausgedrückt. Das Risikoverhältnis in diesem Beispiel liegt bei 0,67.

Tabelle 1: Liefert diese kleine Studie einen zuverlässigen Schätzwert für den Unterschied zwischen Therapie A und Therapie B?

	Therapie A	Therapie B	Risikoverhältnis (A : B =)
Anzahl der Verstorbenen	4	6	(4 : 6 = 0,67)
von (insgesamt)	10	10	

Wäre es sinnvoll, auf der Grundlage dieser niedrigen Teilnehmerzahlen den Schluss zu ziehen, dass Therapie A besser ist als Therapie B? Wahrscheinlich nicht. Es könnte purer Zufall sein, dass es einigen Patienten in der einen Gruppe besser ging als in der anderen Gruppe. Wenn man den Vergleich in anderen kleinen Gruppen von Patienten wiederholen würde, könnten sich die Zahlen der in jeder Gruppe verstorbenen Personen umkehren (6 zu 4) oder gleich groß sein (5 zu 5). Rein zufällig könnte sich aber auch ein ganz anderes Verhältnis ergeben.

Was aber würden Sie erwarten, wenn in jeder Behandlungsgruppe genau derselbe Anteil von Patienten (40 % und 60 %) verstorben wäre, nachdem jeweils 100 Patienten eine der beiden Therapien erhalten hätten (**Tab. 2**)? Auch wenn das Unterschiedsmaß (das Risikoverhältnis) exakt dasselbe ist (nämlich 0,67) wie bei dem Vergleich in Tabelle 1, so ist der Unterschied zwischen 40 Todesfällen auf der einen und 60 Todesfällen auf der anderen Seite doch eindrucksvoller als der Unterschied zwischen 4 und 6 Todesfällen. Damit ist auch die Wahrscheinlichkeit geringer, dass hier nur Zufall im Spiel ist. Um also zu vermeiden, dass wir uns in Therapievergleichen vom Zufallsfaktor täuschen lassen, müssen wir unsere Schlussfolgerungen auf die Untersuchung einer ausreichend großen Anzahl von

Tabelle 2: Liefert diese mittelgroße Studie einen zuverlässigen Schätzwert für den Unterschied zwischen Therapie A und Therapie B?

	Therapie A	Therapie B	Risikoverhältnis (A : B =)
Anzahl der Verstorbenen	40	60	(40 : 60 = 0,67)
von (insgesamt)	100	100	

Patienten stützen, die versterben können oder deren Zustand sich verschlechtern oder bessern oder unverändert bleiben kann. In diesem Fall spricht man manchmal vom «Gesetz der großen Zahl».

Wie man beurteilt, welche Rolle der Zufall in fairen Tests gespielt haben könnte

Bei der Auswertung der Ergebnisse fairer Therapievergleiche kann uns die Rolle, die der Zufall dabei spielt, dazu verleiten, zwei Arten von Fehlern zu begehen: Entweder wir kommen fälschlich zu dem Schluss, dass zwischen den Behandlungsergebnissen echte Unterschiede bestehen, wobei dies in Wirklichkeit gar nicht der Fall ist; oder wir folgern, dass sie sich nicht voneinander unterscheiden, obwohl es tatsächlich aber Unterschiede gibt. Je größer die Anzahl der beobachteten Behandlungsergebnisse, die uns interessieren, ist, desto geringer ist die Wahrscheinlichkeit, dass wir uns in der beschriebenen Art und Weise irren.

Da wir in Therapievergleichen nicht alle Menschen berücksichtigen können, die an der behandelten Krankheit leiden oder noch daran erkranken werden, lassen sich die «wahren Unterschiede» zwischen den untersuchten Therapien auch niemals endgültig feststellen. Stattdessen müssen Studien möglichst gute Schätzungen darüber abgeben, worin die wahren Unterschiede wahrscheinlich bestehen.

Die Zuverlässigkeit dieser geschätzten Unterschiede wird häufig als «Konfidenz- oder Vertrauensintervall» (engl. *confidence interval*, CI) angegeben. Diese Konfidenzintervalle geben den Bereich an, innerhalb dessen die wahren Unterschiede wahrscheinlich liegen. Den meisten Lesern ist das Konzept «Konfidenzintervall» bereits geläufig, wenn auch nicht unbedingt unter diesem Namen. Wir alle kennen Meinungsumfragen im Vorfeld von politischen Wahlen. Eine solche Umfrage könnte beispielsweise ergeben haben, dass Partei A 10 Prozentpunkte vor Partei B liegt; häufig wird in einer solchen Meldung dann darauf hingewiesen, dass der Unterschied zwischen den Parteien mindestens 5, höchstens aber bis zu 15 Prozentpunkte betragen könnte. Dieses «Konfidenzintervall» gibt an, dass der wahre Unterschied zwischen den Parteien wahrscheinlich

irgendwo zwischen 5 und 15 Prozentpunkten liegt. Je größer die Anzahl der befragten Wähler, umso geringer ist die Unsicherheit bezüglich der Ergebnisse, und umso schmaler ist das zu dem geschätzten Unterschied gehörige Konfidenzintervall (**Abb. 11**).

Genauso wie man den Grad der Unsicherheit im Hinblick auf den geschätzten Unterschied in den Anteilen der Wähler bewerten kann, die zwei politische Parteien unterstützen, so lässt sich auch der Grad der Unsicherheit in Bezug auf den geschätzten Unterschied in den Anteilen der Patienten abschätzen, deren Zustand sich nach den zwei unterschiedlichen Therapien bessert oder verschlechtert. Auch hier heißt es: Je größer die Anzahl der untersuchten Behandlungsergebnisse – z.B. die Genesung nach einem Herzinfarkt – beim Vergleich zweier Therapien ist, umso schmaler werden auch die Konfidenzintervalle um die Schätzer der Therapieunterschiede sein. Für Konfidenzintervalle gilt: «Je schmaler, desto besser.»

Normalerweise wird bei einem Konfidenzintervall angegeben, wie sicher wir sein können, dass der wahre Wert innerhalb des angegebenen Schätzwertbereichs liegt. Ein «95%-Konfidenzintervall» bedeutet beispielsweise, dass wir zu 95 % sicher sein können, dass der wahre Wert dessen, was geschätzt wird, innerhalb des durch das Konfidenzintervall eingegrenzten Bereichs liegt. Das heißt, es besteht eine Chance von 5 zu 100 (5 %), dass der «wahre» Wert tatsächlich außerhalb dieses Bereichs liegt.

Was bedeutet ein «signifikanter Unterschied» zwischen den Therapien?

Dies ist eine schwierige Frage, denn der Begriff «signifikanter Unterschied» kann mehrere Bedeutungen haben. Zunächst einmal kann er einen für den Patienten tatsächlich wichtigen Unterschied bezeichnen. Wenn jedoch die Autoren wissenschaftlicher Forschungsberichte von einem «signifikanten Unterschied» sprechen, beziehen sie sich häufig auf «statistische Signifikanz». Und «statistisch signifikante Unterschiede» sind nicht unbedingt «signifikant», d.h. «bedeutsam», im alltäglichen Sinn des Wortes. Denn ein Unterschied zwischen Therapien, der sehr wahrscheinlich nicht zufallsbedingt ist

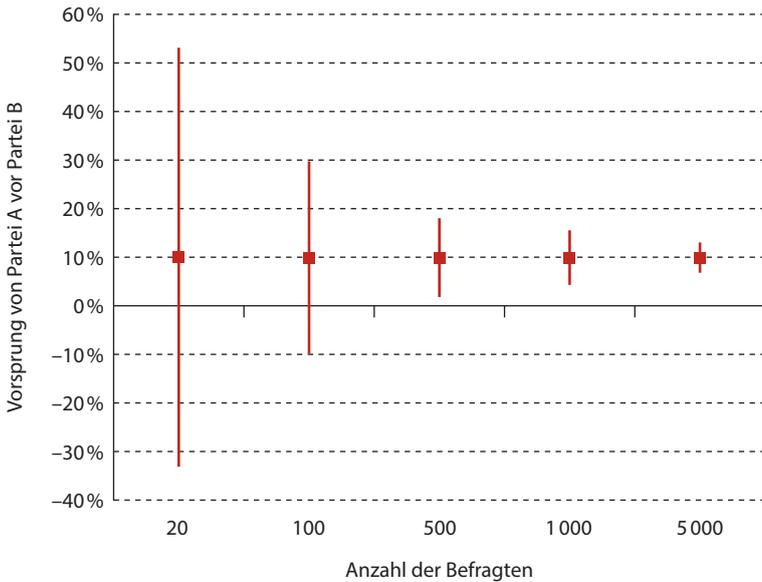


Abbildung 11: Das 95%-Konfidenzintervall (CI) für den Unterschied zwischen Partei A und Partei B wird mit zunehmender Anzahl der befragten Wähler schmaler.

– also «ein statistisch signifikanter Unterschied» – hat unter Umständen nur wenig oder gar keine praktische Relevanz.

Sehen wir uns dazu das Beispiel eines systematischen Reviews von randomisierten Studien an, in denen die Erfahrungen zehntausender gesunder Männer, die täglich eine Aspirin-Tablette eingenommen haben, mit den Erfahrungen zehntausender anderer gesunder Männer verglichen wurden, die kein Aspirin einnahmen. In diesem Review war bei den Männern in der Aspirin-Gruppe eine niedrigere Herzinfarktquote zu beobachten, und der Unterschied war «statistisch signifikant», d. h. er war eher nicht durch reinen Zufall zu erklären. Aber das bedeutet nicht, dass er auch praktische Relevanz hat. Wenn bei einem gesunden Mann die Wahrscheinlichkeit, dass er einen Herzinfarkt erleidet, ohnehin nur sehr gering ist, dann ist die Einnahme eines Medikaments, mit dem dieses Risiko noch stärker gesenkt werden soll, unter Umständen nicht gerechtfertigt, vor allem nicht, wenn man bedenkt, dass Aspirin Nebenwirkungen hat, von denen einige (z. B. Blutungen) tödlich verlaufen können.¹ Auf

der Grundlage der Erkenntnisse aus dem erwähnten systematischen Review können wir davon ausgehen, dass von 1 000 Männern, die zehn Jahre lang täglich 100 mg Aspirin einnehmen, bei fünf in diesem Zeitraum ein Herzinfarkt verhindert würde, drei von ihnen aber eine stärkere Blutung entwickeln würden.

Was bedeutet «statistisch signifikant»?

«Um ehrlich zu sein: Das ist ein kniffliger Begriff. Er kann uns sagen, ob beispielsweise der Unterschied zwischen einem Medikament und einem Placebo oder zwischen den Lebenserwartungen zweier Personengruppen lediglich zufallsbedingt sein könnte ... Er bedeutet, dass es unwahrscheinlich ist, dass ein Unterschied, der so groß ist wie der, den man beobachtet hat, allein auf Zufall beruht. Um den Grad dieser «Unwahrscheinlichkeit» anzugeben, benutzen Statistiker Standardwerte. Üblicherweise nehmen sie ein Signifikanzniveau von 5 % an (manchmal als $p = 0,05$ geschrieben). In diesem Fall bedeutet es, dass ein Unterschied als «signifikant» gilt, weil eine Wahrscheinlichkeit von weniger als 1 zu 20 besteht, dass das, was passiert ist, auf Zufall beruht.»

Spiegelhalter D, zitiert in: *Making Sense of Statistics*. 2010.
www.senseaboutscience.org

Wie man ausreichend große Teilnehmerzahlen für faire Therapietests gewinnt

Manchmal gelingt es in Therapietests, eine ausreichend große Anzahl von Forschungsdaten zu gewinnen, die aus einem oder zwei Studienzentren stammen. Um zuverlässige wissenschaftliche Daten zu erhalten und damit die Auswirkungen von Behandlungen auf seltene Behandlungsergebnisse (Outcomes) wie z.B. den Todesfall beurteilen zu können, muss man normalerweise Patienten in vielen Zentren und oftmals auch in vielen Ländern zur Teilnahme an einer Studie einladen. So ergab sich beispielsweise in einer Studie mit 10 000 Patienten aus 13 Ländern, dass bei schweren Hirnverletzungen die Gabe von Steroiden – die mehr als drei Jahrzehnte lang praktiziert wurde – tödlich war.² In einem weiteren, von demselben Forscherteam organisierten fairen Test erbrachte die Teilnahme von 20 000 Patienten aus 40 Ländern den Nachweis, dass ein kostengünstiges Medikament namens Tranexamsäure die Anzahl der blutungsbedingten Todesfälle nach Verletzungen senkt.³ Da diese

Studien darauf ausgelegt waren, systematische Fehler (Bias) sowie zufallsbedingte Unsicherheiten zu reduzieren, stellen sie vorbildliche Beispiele für faire Tests. Sie lieferten hochwertige Evidenz, die für die Gesundheitsversorgung weltweit von großer Bedeutung ist. Die zweite dieser randomisierten Studien wurde dementsprechend in einer vom *British Medical Journal* (BMJ) durchgeführten Umfrage zur wichtigsten Studie des Jahres 2010 gewählt.

Abbildung 12 beruht auf Daten, die uns freundlicherweise von dem preisgekrönten Forscherteam zur Verfügung gestellt wurden. Sie veranschaulichen, wie wichtig es ist, unsere Schätzungen von Therapieeffekten auf möglichst viele Informationen zu stützen, um die Gefahr zu verringern, dass wir uns vom Zufall täuschen lassen. Die Raute im unteren Abschnitt der Abbildung gibt das Gesamtergebnis der Studie über Tranexamsäure an. Sie zeigt, dass das Medikament die Anzahl der blutungsbedingten Todesfälle um nahezu 30 % reduziert (Risikoverhältnis knapp über 0,7). Dieses Gesamtergebnis stellt die zuverlässigste Schätzung der Wirkung dieses Medikaments dar, auch wenn der Schätzwert der Zentren in Kontinent A

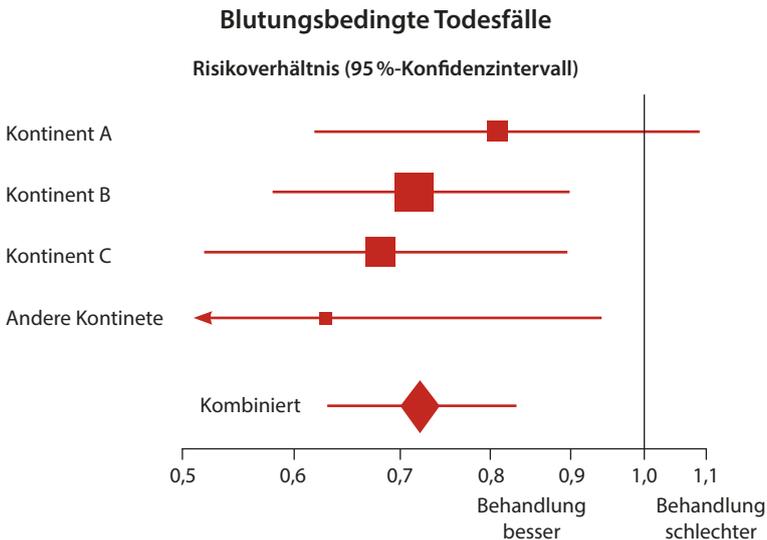


Abbildung 12: Wirkungen von Tranexamsäure auf die Sterblichkeitsrate bei Trauma-Patienten mit signifikanten Blutungen: insgesamt sowie nach dem Kontinent, aus dem die Teilnehmer stammten (unveröffentlichte Daten aus CRASH-2: *Lancet* 2010; 376: 23-32).

einen weniger deutlichen Effekt nahelegt (der statistisch nicht signifikant ist und wahrscheinlich eine Unterschätzung des wahren Effekts darstellt) und der Schätzwert aus Zentren in der Kategorie «andere Kontinente» auf einen *stärkeren* Effekt schließen lässt (wobei es sich wahrscheinlich um eine Überbewertung handelt).

Genauso, wie man den Zufallsfaktor mindern kann, indem man die Daten aus vielen Zentren in einer multinationalen klinischen Studie kombiniert, so können auch die Ergebnisse aus ähnlichen, aber voneinander unabhängigen Studien manchmal statistisch in einer «Meta-Analyse» (s.a. Kap. 8) zusammengefasst werden. Die Methoden für solche Meta-Analysen wurden viele Jahre lang von Statistikern entwickelt. In größerem Umfang wurden sie aber erst in den 1970er-Jahren eingesetzt, und zwar zunächst von Sozialwissenschaftlern in den USA und später dann von Medizinwissenschaftlern. Seit Ende des 20. Jahrhunderts sind Meta-Analysen überall als wichtiger Bestandteil fairer Therapietests anerkannt.

So wurden beispielsweise fünf Studien in fünf verschiedenen Ländern organisiert und unabhängig voneinander finanziert, um einer Frage nachzugehen, die seit 60 Jahren unbeantwortet geblieben war: «Bei welchem Sauerstoffspiegel im Blut ist die Wahrscheinlichkeit am höchsten, dass Frühgeborene ohne größere Behinderungen überleben?» Ist der Sauerstoffspiegel im Blut nämlich zu hoch, können die Säuglinge erblinden; ist er zu niedrig, können sie sterben oder eine Zerebralparese entwickeln. Da die verschiedenen Sauerstoffdosierungen selbst bei diesen winzigen Babys wahrscheinlich nur mäßige Unterschiede bewirken, braucht man eine große Anzahl von Säuglingen, um diese Unterschiede nachweisen zu können. Deshalb erklärten sich die für jede dieser fünf Studien verantwortlichen Forscherteams zu einer Zusammenfassung der Ergebnisse ihrer jeweiligen Studien bereit, um dadurch eine zuverlässigere Schätzung zu ermöglichen, die mit jeder einzelnen ihrer Studien nicht möglich gewesen wäre.⁴

- Bei der Beurteilung der Frage, wie viel Vertrauen man in die Qualität und Quantität der verfügbaren Evidenz setzen kann, muss auch dem «Zufallsfaktor» Rechnung getragen werden.

8

Bewertung der relevanten verlässlichen Evidenz

Kann eine einzige Studie genügen?

Die einfache Antwort lautet: «Wohl kaum». Nur selten wird ein einziger fairer Therapievergleich hinreichend verlässliche Erkenntnisse (Evidenz) liefern, auf deren Grundlage eine Entscheidung über verschiedene Therapieoptionen getroffen werden kann. Gelegentlich kann das aber doch vorkommen. Zu diesen seltenen Einzelstudien gehört eine Studie, in der nachgewiesen wurde, dass die Einnahme von Aspirin während eines Herzinfarkts das Risiko eines vorzeitigen Todes senkt¹, eine zweite, in der deutlich wurde, dass die Gabe von Steroiden bei akuten traumatischen Hirnverletzungen tödlich ist (s. u. und Kap. 7, S. 138), sowie eine dritte Studie, in der Koffein als einziges Medikament identifiziert wurde, mit dem sich bei Frühgeborenen eine Zerebralparese verhindern lässt (s. Kap. 5, S. 99). Normalerweise ist eine Einzelstudie aber nur einer von mehreren Vergleichen, in denen denselben oder ähnlichen Fragen nachgegangen wird. Deshalb sollten die Ergebnisse aus Einzelstudien immer zusammen mit den Ergebnissen aus anderen, ähnlichen Studien ausgewertet werden.

Der britische Statistiker Austin Bradford Hill, einer der Pioniere fairer Therapietests, forderte in den 1960er-Jahren, dass in Forschungsberichten die folgenden vier Fragen beantwortet werden müssten:

- Warum wurde die Untersuchung begonnen?
- Was wurde gemacht?

- Was wurde herausgefunden?
- Und was bedeuten ihre Ergebnisse überhaupt?

Warum wurde die Untersuchung begonnen?

«Nur wenige Prinzipien sind für die wissenschaftliche und ethische Validität medizinischer Forschung wichtiger als der Grundsatz, dass Studien Fragen untersuchen sollten, die dringend beantwortet werden müssen, und dass sie so ausgelegt sein sollten, dass sie eine bedeutsame Antwort auf diese Fragen geben können. Diese beiden Ziele setzen voraus, dass dazu relevante frühere Forschungsarbeiten identifiziert werden. ... Ein unvollständiges Bild der bereits vorhandenen Erkenntnisse stellt eine Verletzung der unausgesprochenen ethischen Grundlage des Vertrags mit den Studienteilnehmern dar, wonach die Informationen, die mit ihrer Hilfe gewonnen werden sollen, notwendig und für andere Menschen von Nutzen sind.»

Robinson KA, Goodman SN. A systematic examination of the citation of prior research in reports of randomized, controlled trials. *Annals of Internal Medicine* 2011; 154: 50–55.

Auch heute haben diese Schlüsselfragen nichts von ihrer Bedeutung verloren, und doch werden sie nur allzu oft unzureichend thematisiert oder gar völlig übersehen. Die Antwort auf die letzte Frage – Was bedeuten die Ergebnisse? – ist besonders wichtig, weil sie sehr wahrscheinlich Einfluss auf Therapieentscheidungen sowie Entscheidungen über zukünftige Forschungsvorhaben hat.

Nehmen wir das Beispiel der kurzzeitigen Verabreichung eines preiswerten steroidhaltigen Medikaments an Frauen mit drohender Frühgeburt. Der erste faire Test zu dieser Therapie, über den 1972 berichtet wurde, ergab, dass die Sterblichkeitswahrscheinlichkeit von Frühgeborenen nach der mütterlichen Einnahme eines solchen steroidhaltigen Medikaments abnimmt. Zehn Jahre später waren dazu weitere Untersuchungen durchgeführt worden: Es handelte sich dabei allerdings um kleine Studien, deren Einzelergebnisse verwirrend waren, weil keine von ihnen früher durchgeführte ähnliche Studien systematisch berücksichtigt hatte. Andernfalls wäre nämlich deutlich geworden, dass man aus ihnen sehr stichhaltige Belege zugunsten eines vorteilhaften Effekts dieser Medikamente hätte ableiten können. Da dies erst 1989 nachgeholt wurde, war den meisten Geburtshelfern, Hebammen, Kinderärzten und Säuglingsschwes-

tern zwischenzeitlich gar nicht bewusst, wie wirksam diese Therapie war. Infolgedessen hatten zehntausende Frühgeborene unnötig leiden und sterben müssen.²

Um die Frage «Was bedeuten die Ergebnisse?» zu beantworten, müssen die aus einem einzelnen fairen Therapievergleich stammenden Erkenntnisse zusammen mit den Ergebnissen anderer, ähnlicher fairer Vergleiche ausgewertet werden. Neue Studienergebnisse zu veröffentlichen, ohne sie im Licht anderer relevanter und im Rahmen systematischer Übersichtsarbeiten zusammengefasster Ergebnisse zu interpretieren, kann zu einer verzögerten Identifizierung sowohl nützlicher als auch schädlicher Therapien und auch zu unnötiger Forschung führen.

Informationen aus der Forschung zusammenfassen

Schon vor mehr als einem Jahrhundert äußerte sich Lord Rayleigh, der damalige Präsident der British Association for the Advancement of Science, über die Notwendigkeit, die Ergebnisse neuer Forschungsarbeiten in den Kontext anderer relevanter Forschungsergebnisse zu stellen: «Bestünde die Wissenschaft, wie zuweilen angenommen wird, lediglich aus dem mühseligen Sammeln von Fakten, würde sie bald zum Stillstand kommen und sozusagen von ihrem eigenen Gewicht erdrückt werden ... Deshalb müssen hier zwei Prozesse gleichzeitig stattfinden: die Aufnahme neuer Ergebnisse sowie die «Verdauung» und Integration bereits vorhandener Erkenntnisse; und da beide Prozesse wichtig sind, können wir uns die Diskussion ihrer jeweiligen Bedeutung an dieser Stelle sparen ... Die Arbeiten, die die meiste Anerkennung verdienen, aber leider nicht immer erhalten, sind diejenigen, bei denen Entdeckung und Erklärung Hand in Hand gehen, in deren Zusammenhang nicht nur die neuen Fakten vorgestellt werden, sondern auch der Zusammenhang mit den schon vorhandenen aufgezeigt wird.»

Rayleigh, Lord. In: Report of the fifty-fourth meeting of the British Association for the Advancement of Science; held at Montreal in August and September 1884. London: John Murray, 1884: pp. 3–23.

Systematische Reviews der relevanten zuverlässigen Evidenz

Zu verlangen, dass wir die Ergebnisse einer bestimmten Studie vor dem Hintergrund anderer relevanter und zuverlässiger Erkenntnisse prüfen sollten, fällt zwar nicht schwer, doch stellt diese Forderung in vielfacher Hinsicht eine Herausforderung dar. Es ist wichtig, dass

Reviews systematisch durchgeführt werden – die Menschen müssen sich auf sie verlassen können. Unsystematische Reviews sind anfälliger für Ergebnisverzerrungen und können uns in die Irre führen.

Die Bedeutung systematischer Reviews

«Systematische Übersichtsarbeiten (Reviews) und Meta-Analysen haben im Gesundheitswesen zunehmend an Bedeutung gewonnen. Ärzte lesen sie, um sich in ihrem Fachgebiet auf dem Laufenden zu halten, und häufig dienen sie auch als Ausgangspunkt für die Erstellung von klinischen Leitlinien. Institutionen für Forschungsförderung verlangen mitunter einen systematischen Review, um sicherzustellen, dass weitere Forschungsvorhaben gerechtfertigt sind, und auch etliche medizinische Fachzeitschriften bewegen sich bereits in diese Richtung. Wie bei allen Forschungsarbeiten hängt auch der Wert eines systematischen Reviews davon ab, was gemacht wurde, was dabei festgestellt wurde und wie übersichtlich und genau darüber berichtet wird. Wie bei anderen Veröffentlichungen variiert die Berichtsqualität auch bei den systematischen Übersichtsarbeiten; und dies schränkt auch die Möglichkeit des Lesers ein, die Stärken und Schwächen solcher Reviews zu beurteilen.»

Moher D, Liberati A, Tetzlaff, Altman DG. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement (www.equator-network.org), 2009.

Auch wenn systematische Übersichtsarbeiten zu bestimmten Therapien anscheinend dieselbe Fragestellung untersuchen, können sie zu unterschiedlichen Schlussfolgerungen gelangen. Manchmal liegt das daran, dass die untersuchten Fragen sich geringfügig unterscheiden oder dass die Forscher unterschiedliche Methoden angewandt haben; und manchmal hat es damit zu tun, dass die Forscher die Ergebnisse nur aus einem für sie interessanten Blickwinkel betrachten. Deshalb ist es wichtig, Reviews zu identifizieren, in der die Therapien mit Fragestellungen untersucht werden, die jenen ähneln, für die wir selbst uns interessieren. Zugleich sollten die Reviews mit hoher Wahrscheinlichkeit so erstellt worden sein, dass sie die Auswirkungen von systematischen Fehlern und Zufallsfaktoren erfolgreich reduziert haben und somit ehrliche Schlussfolgerungen ermöglichen, in denen sich die dargestellten Ergebnisse (Evidenz) auch widerspiegeln.

Wie man Bias in systematischen Reviews reduziert

Genauso wie systematische Fehler die Ergebnisse einzelner Therapietests verzerren und zu falschen Schlussfolgerungen führen können, so können sie auch die Ergebnisse von Übersichtsarbeiten über die vorhandene Evidenz verzerren. Wissenschaftler könnten sich z.B. einfach diejenigen Studien «herauspicken», die ihre eigenen Behauptungen über die betreffende Therapie untermauern.

Um diesen Problemen vorzubeugen, sollten die Arbeitsschritte für systematische Reviews genauso wie bei wissenschaftlichen Studien in Forschungsprotokollen niedergelegt werden. In diesen Protokollen sollten die Wissenschaftler festlegen, welche Maßnahmen sie bei der Erstellung des Reviews zur Verminderung von systematischen Fehlern und Zufallseffekten ergreifen werden. Beispielsweise sollten sie u.a. angeben, mit welchen Fragen zur Therapie sich der Review befassen wird, und Kriterien festlegen, anhand derer über die Eignung von Studien für den Einschluss in den Review entschieden wird, wie potenziell geeignete Studien ermittelt werden sollen und welche Schritte unternommen werden, um bei der Auswahl der für den Review und die Datenanalyse geeigneten Studien systematische Fehler auf ein Minimum zu begrenzen.

Wie man die gesamte relevante Evidenz für systematische Reviews identifiziert

Die Identifizierung der gesamten relevanten Evidenz für einen systematischen Review – unabhängig von der Sprache oder vom Format der relevanten Studienberichte – stellt jedes Mal eine Herausforderung dar. Das liegt nicht zuletzt daran, dass manche relevanten wissenschaftlichen Daten mitunter gar nicht veröffentlicht werden. Dieses sogenannte «Underreporting» («Nichtpublizieren») ist hauptsächlich dadurch bedingt, dass Wissenschaftler ihre Ergebnisse nicht aufschreiben oder ihre Forschungsberichte nicht zur Veröffentlichung einreichen, wenn sie von den Ergebnissen enttäuscht sind. Und pharmazeutische Unternehmen unterdrücken Studien, in denen ihre Produkte keine günstige Beurteilung erhalten. Auch Fachzeitschriften zeigen sich dann als voreingenommen, wenn sie

eingereichte Berichte ablehnen, weil sie deren Ergebnisse für nicht «aufsehenerregend» genug halten.³

Das auf Voreingenommenheit beruhende «Nichtpublizieren» von Forschungsergebnissen ist unwissenschaftlich und unethisch und inzwischen auch allgemein als ernst zu nehmendes Problem erkannt worden. Menschen, die versuchen, sich für eine von mehreren Therapiemöglichkeiten zu entscheiden, können dadurch, dass Studien mit «enttäuschenden» oder «negativen» Ergebnissen seltener zur Veröffentlichung kommen, in die Irre geführt werden. Dagegen ist bei Studien mit spektakulären Ergebnissen eher ein «Overreporting» («Mehrfachpublikation») zu beobachten.

Das Ausmaß von Underreporting ist erstaunlich: Mindestens die Hälfte aller klinischen Studien wird niemals oder unvollständig veröffentlicht. Dieses Nichtpublizieren von Forschungsergebnissen stellt eine Verzerrung dar und betrifft sowohl große als auch kleine klinische Studien. Um dieses Problem in den Griff zu bekommen, wurde u. a. vereinbart, dass Studien schon bei ihrem Beginn registriert werden und dass Wissenschaftler die Protokolle ihrer Studien veröffentlichen sollten.³

Ein durch Voreingenommenheit bedingtes Underreporting von Forschungsergebnissen kann sogar tödliche Folgen haben. Es ist einer Gruppe von britischen Forschern hoch anzurechnen, dass sie noch 1993 die Ergebnisse einer klinischen Studie veröffentlichten, die sie bereits 13 Jahre zuvor durchgeführt hatten. In dieser Studie war es um ein neues Medikament zur Verringerung von Herzrhythmusstörungen bei Herzinfarkt-Patienten gegangen. Neun Patienten waren nach der Einnahme des Medikaments verstorben, während in der Vergleichsgruppe nur ein Todesfall aufgetreten war. «Als wir unsere Studie 1980 durchführten», so schrieben die Wissenschaftler,

dachten wir, die erhöhte Sterblichkeitsrate in der Medikamentengruppe sei zufallsbedingt gewesen. ... Die Entwicklung des Medikaments [Lorcainid] wurde aus wirtschaftlichen Gründen aufgegeben, weshalb diese Studie niemals veröffentlicht wurde; heutzutage stellt sie ein gutes Beispiel für einen «Publikationsbias» dar. Die hier beschriebenen Ergebnisse ... hätten uns vielleicht frühzeitig vor den sich abzeichnenden Problemen warnen können.⁴

Mit den «sich abzeichnenden Problemen» meinten sie, dass ähnliche Medikamente wie das von ihnen getestete Präparat auf dem Höhepunkt ihrer Anwendung allein in den USA jedes Jahr zehntausende von vorzeitigen Todesfällen verursachten (s. Kap. 2, S. 46 f.).⁵

Marketingbasierte Medizin

«Interne Dokumente aus der pharmazeutischen Industrie lassen darauf schließen, dass die öffentlich zugängliche Evidenzbasis die ihren Produkten zugrunde liegenden Daten womöglich nicht exakt widerspiegelt. Der Industrie und den mit ihr zusammenarbeitenden medizinischen Kommunikationsberatern zufolge dienen Veröffentlichungen in der medizinischen Literatur in erster Linie Marketinginteressen. Die Unterdrückung und Änderung negativer Daten sowie das Ghostwriting [s. Kap. 10, S. 179] sind zu einem Mittel geworden, mit dem man Publikationen in medizinischen Fachzeitschriften optimal auf den Absatz der eigenen Produkte ausrichten kann, während das Erfinden von Krankheiten und die Marktsegmentierung hinsichtlich der Ärzteschaft zur effizienten Gewinnmaximierung eingesetzt werden. Unserer Meinung nach ist die evidenzbasierte Medizin zwar ein nobles Ideal, Realität ist derzeit aber die marketingbasierte Medizin.»

Spielmans GI, Parry PI. Aus: Evidence-based Medicine to Marketing-based Medicine: Evidence from Internal Industry Documents. *Journal of Bioethical Inquiry* 2010; 7(1): 13–29. Zugänglich unter <http://tinyurl.com/Spielmans>.

Wie man in systematischen Reviews den Einfluss des Zufalls verringert

In Kapitel 7 (S. 140) haben wir dargelegt, wie der Einfluss des Zufalls verringert werden kann, indem man Daten aus ähnlichen, aber eigenständigen Studien in einer sogenannten «Meta-Analyse» zusammenfasst. Wir haben dazu das Beispiel von fünf Studien, die unabhängig voneinander in fünf verschiedenen Ländern organisiert und finanziert wurden, gewählt, um einem 60 Jahre alten Dilemma auf die Spur zu kommen. Es ging dabei um die Frage, wie hoch der Sauerstoffspiegel im Blut von Frühgeborenen sein muss, um die Wahrscheinlichkeit ihres Überlebens ohne größere Behinderungen zu maximieren. Dieses Beispiel zeigt auf, wie die Verringerung des Zufallsfaktors bereits vor der Verfügbarkeit der Studienergebnisse geplant werden konnte. Ein solches Vorgehen ist aber auch noch nach dem Abschluss einer Reihe von ähnlichen Studien möglich.

So erstellte beispielsweise 1974 ein schwedischer Arzt eine systematische Übersichtsarbeit von Studien, in denen die Ergebnisse von

Brustkrebsoperationen mit bzw. ohne Strahlentherapie miteinander verglichen wurden.⁶ Dabei stellte er fest, dass die strahlentherapeutisch behandelten Gruppen von Frauen in all diesen Studien eine höhere Sterblichkeitswahrscheinlichkeit aufwiesen. Als er die Gesamtheit der Ergebnisse mithilfe einer Meta-Analyse statistisch zusammenfasste, wurde klar, dass die übermäßige Sterblichkeit kaum zufallsbedingt sein konnte. Nachfolgende detailliertere Analysen, die auf den Daten von einzelnen Patientinnen beruhte, bestätigten, dass die in der damaligen Zeit angewandte Strahlentherapie die Sterblichkeit tatsächlich erhöhte.⁷ Auf der Basis dieser Erkenntnis konnten sicherere Vorgehensweisen entwickelt werden.

Wie man in systematischen Reviews das Vorhandensein von persönlichen Interessen und Datenmanipulation erkennt

Was passiert, wenn Reviewautoren noch andere Interessen verfolgen, welche die Durchführung oder Auswertung ihres Reviews beeinflussen? Es kann z.B. vorkommen, dass die Reviewer von dem Unternehmen, das die Prüfung der neuen Therapie in Auftrag gegeben hat, finanziell unterstützt werden. Bei der Bewertung der Evidenz für die Wirksamkeit von Nachtkerzenöl bei Ekzemen fiel die Einschätzung der Reviewer, die Verbindungen zum Hersteller unterhielten, sehr viel enthusiastischer aus als bei den Reviewern ohne entsprechende kommerzielle Interessen (s. Kap. 2, S. 51 f.). Aber nicht nur kommerzielle Interessen führen zu voreingenommenen Reviews. Wir alle haben Vorurteile, die uns voreingenommen machen – und zwar Forscher, Ärzte und Patienten gleichermaßen.

Diese Voreingenommenheit kann Wissenschaftler– häufig aus kommerziellen Gründen – dazu verleiten, vorhandene Erkenntnisse absichtlich zu ignorieren. Sie planen, analysieren und publizieren Forschungsarbeiten, um ihre eigenen Ergebnisse zu einer bestimmten Therapie in einem günstigen Licht erscheinen zu lassen.⁸ Genau das passierte in den 1990er-Jahren, als der Hersteller des Antidepressivums Seroxat (Paroxetin) wichtige Daten zurückhielt, die darauf hinwiesen, dass das Medikament bei Jugendlichen die depressiven Symptome tatsächlich verstärkte, und einiger dieser jungen

Patienten in der Folge einen Suizid als Ausweg aus ihrer Depression in Erwägung zogen.⁹

Ein weiteres Problem stellt das Overreporting dar. Bei dem als «Salamitaktik» bezeichneten Phänomen nehmen Wissenschaftler die Ergebnisse einer einzelnen Studie (der «Salami») und teilen die Ergebnisse in mehrere Berichte auf, ohne klarzustellen, dass es sich bei den einzelnen Berichten nicht um voneinander unabhängige Studien handelt. Auf diese Weise kann eine einzelne «positive» Studie in Fachzeitschriften und unterschiedlichen Artikeln mehrfach veröffentlicht und als mehrere Studien wahrgenommen werden, wodurch ein Bias entsteht.¹⁰ Die Verwirrung, die ein solches Vorgehen stiften kann, lässt sich dadurch verringern, dass Studien bereits bei Studienbeginn in ein Register eingetragen werden und dabei eine Identifikationsnummer zugewiesen bekommen.

Was passieren kann, wenn nicht alle relevante verlässliche Evidenz ausgewertet wird

Zu einer fairen Therapiestudie gehört auch die systematische Prüfung der gesamten relevanten und verlässlichen Evidenz. Nur so kann festgestellt werden, ob die Erkenntnisse aus Tier- oder anderen Laborstudien stammen oder von gesunden Probanden (Freiwilligen), an denen manchmal neue Therapien getestet werden, oder aus früheren Forschungsarbeiten, an denen Patienten beteiligt waren. Wird dieser Schritt ausgelassen oder nicht sorgfältig durchgeführt, können sich daraus ernsthafte Folgen ergeben: So kann es passieren, dass Patienten im Allgemeinen sowie Teilnehmer an klinischen Studien im Besonderen Schaden nehmen (und manchmal auch unnötigerweise sterben) und dass in der Gesundheitsversorgung und -forschung wertvolle Ressourcen verschwendet werden.

Vermeidbare Nachteile für Patienten

Die Therapieempfehlungen bei Herzinfarkt, die über einen Zeitraum von 30 Jahren in verschiedenen Lehrbüchern erschienen sind, wurden mit der Evidenz verglichen, die die Autoren durch eine sys-

tematische Auswertung der Ergebnisse fairer Therapietests aus dem besagten Zeitraum hätten gewinnen und berücksichtigen können.¹¹ Der Vergleich zeigte, dass die Lehrbuchempfehlungen häufig falsch waren, weil die Autoren die relevante Evidenz nicht systematisch geprüft hatten. Das hatte katastrophale Folgen. In einigen Fällen kam es dazu, dass Patienten mit Herzinfarkt deshalb lebensrettende Therapien (z.B. gerinnselauflösende Medikamente) vorenthalten blieben. Und es gab andere Fälle, in denen Ärzte Therapien auch dann noch empfahlen, nachdem faire Tests längst gezeigt hatten, dass diese tödlich verlaufen können (z.B. die Gabe von Medikamenten, die bei Herzinfarkt-Patienten Herzrhythmusstörungen mildern sollen; s. o. und Kap. 2, S. 46 f.).

Die Wissenschaft ist kumulativ, aber Wissenschaftler akkumulieren die Evidenz nicht wissenschaftlich

«Wissenschaftler reden schon seit etwa 25 Jahren von der sogenannten «kumulativen Meta-Analyse»: Dazu lässt man parallel zur voranschreitenden Forschung eine Meta-Analyse zu einer bestimmten Intervention laufen. Jedes Mal, wenn eine Studie abgeschlossen ist, gibt man die entsprechenden Daten ein und erhält so ein aktualisiertes gepooltes Ergebnis. Auf diese Weise entwickelt man ein Gefühl dafür, in welche Richtung sich die Ergebnisse bewegen. Besonders nützlich dabei ist, dass auf diese Weise eine gute Chance besteht, eine statistisch signifikante Antwort zu erkennen, sobald sich Anhaltspunkte dafür ergeben, ohne dass man das Leben von Menschen in weiteren unnötigen wissenschaftlichen Untersuchungen gefährden muss.»

Goldacre B. Bad Science: How pools of blood trials could save lives.
The Guardian, 10. Mai 2008, S. 16.

Noch immer erleiden Patienten Schaden, weil es nicht gelingt, die Ergebnisse von Studien in systematischen Reviews zusammenzufassen, sobald neue Erkenntnisse verfügbar werden. Blutersatzstoffe z.B., bei denen auf Kühlung und Kreuzproben verzichtet werden kann, stellen bei der Versorgung von Blutungen offensichtlich eine attraktive Alternative zu echtem Blut dar. Leider erhöhen diese Produkte auch das Risiko, einen Herzinfarkt zu erleiden und zu versterben. Wie eine systematische Übersichtsarbeit zu den seit den späten 1990er-Jahren veröffentlichten randomisierten Studien ganz klar zeigte, hätte man die mit diesen Blutersatzstoffen verbundenen Ge-

fahren bereits mehrere Jahre vorher erkennen können und auch erkennen müssen.¹

Vermeidbare Nachteile für Studienteilnehmer

Wenn versäumt wird, die gesamte relevante und zuverlässige Evidenz zu bewerten, kann dies auch bei den Teilnehmern an wissenschaftlichen Untersuchungen zu Schädigungen führen, die eigentlich vermeidbar wären. Noch immer erhalten Forscher den Auftrag und die Genehmigung, Studien durchzuführen, bei denen sie den Teilnehmern bekanntermaßen wirksame Therapien vorenthalten. So gab es beispielsweise zuverlässige Hinweise darauf, dass die Gabe von Antibiotika bei Patienten, die sich einer Darmoperation unterzogen, die Wahrscheinlichkeit reduzierte, dass sie an Operationskomplikationen verstarben. Trotzdem führten Forscher, die es versäumt hatten, bereits vorhandene Erkenntnisse systematisch auszuwerten, weiterhin Vergleichsstudien durch, in denen die Hälfte der Teilnehmer an kontrollierten Studien keine Antibiotika erhielt. Offensichtlich hatten die Einrichtungen, die diese Forschung förderten, und die wissenschaftlichen Ethikkommissionen, welche die Studienprotokolle geprüft hatten, diesen schwerwiegenden Fehler übersehen und es versäumt, entsprechenden Druck auf die Forscher auszuüben.

Wenn Forscher das, was bereits über die Wirkungen von Therapien, die sie einsetzen, bekannt ist, nicht systematisch auswerten, gefährden sie damit nicht nur behandlungsbedürftige Patienten, sondern auch gesunde Probanden. An der ersten Phase einer Therapiestudie nimmt oftmals nur eine sehr geringe Anzahl gesunder Freiwilliger teil. 2006 beispielsweise wurden sechs jungen männlichen Probanden an einer privaten Forschungseinrichtung in West London Infusionen eines Medikaments gegeben, das zuvor noch nicht beim Menschen angewendet worden war. Sie alle erlitten lebensbedrohliche Komplikationen, die ihre Gesundheit langfristig beeinträchtigten: Einer von ihnen verlor Finger und Zehen. Diese Tragödie hätte sehr wahrscheinlich verhindert werden können,¹³ wenn erstens ein Bericht über eine schwerwiegende Reaktion auf ein ähnliches Medikament zur Publikation eingereicht worden wäre¹⁴

und wenn zweitens die Forscher systematisch ausgewertet hätten, was man über die Wirkungen solcher Medikamente bereits wusste.¹⁵ Sie hätten in diesem Fall ihre Untersuchung vielleicht überhaupt nicht fortgesetzt oder andernfalls den Probanden das Medikament vielleicht nacheinander anstatt gleichzeitig infundiert. Zudem hätten sie die gesunden jungen Probanden vor den möglichen Gefahren warnen können (und auch müssen).¹⁶

Verschwendung von Ressourcen im Gesundheitswesen und in der Forschung

Auch wenn Patienten bzw. Studienteilnehmer keinen unmittelbaren Schaden nehmen, hat ein Verzicht auf die Durchführung von systematischen Reviews der relevanten zuverlässigen Belege aus wissenschaftlichen Studien negative Folgen. Denn es kann dadurch zur Vergeudung von Ressourcen im Gesundheitswesen und in der Gesundheitsforschung kommen. Während der 1980er- und 1990er-Jahre nahmen beispielsweise insgesamt mehr als 8 000 Patienten an mehreren Studien zu einem neuen Schlaganfallmedikament teil. Niederländische Forscher, welche die Ergebnisse dieser Medikamentenstudien systematisch auswerteten, konnten keinerlei vorteilhafte Effekte nachweisen (s. Kap. 10, S. 168).¹⁷ Deshalb beschlossen sie, auch die Ergebnisse von Medikamentenstudien auszuwerten, die zuvor an Tieren durchgeführt worden waren, und auch hier gelang ihnen der Nachweis günstiger Effekte nicht.¹⁸ Hätten die Wissenschaftler, die die Tierversuche durchgeführt hatten, und die klinischen Forscher die Ergebnisse der Tierstudien zum damaligen Zeitpunkt systematisch ausgewertet, wären wahrscheinlich nicht Tausende von Patienten zur Teilnahme an den klinischen Studien eingeladen worden. Tatsächlich hätten die Ressourcen dann besser für die Behandlung von Schlaganfall-Patienten und für Studien aufgewendet werden können, die für die Identifizierung von Möglichkeiten zur Verbesserung von Schlaganfalltherapien relevant gewesen wären. Und dies ist bei weitem kein Einzelfall.¹⁹

Hätte sich ein Todesfall durch vorherige Prüfung der Evidenz verhindern lassen?

«Unter tragischen Umständen, die hätten verhindert werden können, verstarb im Juni [2001] die zuvor gesunde 24-jährige Ellen Roche, weil eine Chemikalie, die sie im Rahmen einer Asthmastudie der Johns Hopkins University inhalieren sollte, zu progredientem Lungen- und Nierenversagen führte. Bei der Aufarbeitung dieses Todesfalls wurde deutlich, dass der Forscher, der das Experiment durchgeführt, und die Ethikkommission, die es genehmigt hatte, anscheinend zahlreiche Hinweise auf die Gefahren der Chemikalie Hexamethonium, die Ellen Roche inhalierten musste, übersehen hatten. Besondere Brisanz erhielt der Fall zudem dadurch, dass man die Belege für die Gefährlichkeit der Chemikalie leicht in der veröffentlichten Literatur hätte finden können. Die Tageszeitung Baltimore Sun gelangte zu dem Fazit, dass der behandelnde Arzt Dr. Alkis Togias zwar «aufrichtig bemüht» gewesen war, die unerwünschten Wirkungen des Medikaments zu recherchieren, dass seine Recherche sich aber anscheinend auf eine begrenzte Anzahl von Quellen gestützt habe, u. a. auf PubMed, dessen durchsuchbares Archiv aber nur bis zum Jahr 1966 zurückreichte. Die Warnungen vor den mit Hexamethonium einhergehenden Lungenschäden stammten jedoch aus früheren Artikeln, die bereits in den 1950er-Jahren veröffentlicht und in späteren Publikationen zitiert worden waren.»

Perkins E. Johns Hopkins Tragedy. *Information Today* 2001; 18: 51–4.

Am Anfang und am Ende neuer Forschung sollte ein systematischer Review stehen

Der Bericht über eine Studie²⁰, in der die Wirkungen von Steroiden bei akuten traumatischen Hirnverletzungen untersucht wurden, ist beispielhaft dafür, wie die vier obigen, von Bradford Hill formulierten Fragen adressiert werden können: Als Grund für die Durchführung der Studie gaben die Autoren an, dass ihr systematischer Review der vorhandenen Belege sowie der Hinweis auf Unterschiede in der klinischen Anwendung dieser Therapie relevante Unsicherheiten über die Wirkungen dieser weit verbreiteten Therapie ergeben hatte. Ihren Angaben zufolge wurde das Studienprotokoll zu Beginn der Studie registriert und veröffentlicht.

Des Weiteren beschrieben sie, dass sie, um systematische Fehler zu minimieren und Zufallsfaktoren zu kontrollieren, eine hinreichend große Anzahl von Patienten untersucht hatten. Ihren Angaben zufolge erbrachte die Studie den Nachweis, dass die Verabreichung von Steroiden an Patienten mit schweren Hirnverletzungen die Sterbewahrscheinlichkeit bei diesen Patienten erhöhte.

**Autorenhinweise der Herausgeber der medizinischen Fachzeitschrift
«The Lancet» zur Einordnung von Forschungsergebnissen in den jeweiligen
Kontext**

Systematischer Review: In diesem Abschnitt sollten die Autoren ihre Suche nach der gesamten vorhandenen Evidenz beschreiben. Sie sollten auch angeben, wie sie die Qualität dieser Evidenz bewertet haben – d. h., wie sie die Evidenz ausgewählt und zusammengefasst haben.

Interpretation: Hier sollten die Autoren angeben, inwieweit ihre Studie die vorhandene Evidenz bereichert, wenn man sie im Lichte schon durchgeführter Forschungsarbeiten betrachtet.

«In allen ab dem 1. August eingereichten Forschungsberichten – randomisiert oder nicht – ... sind die Ergebnisse im Abschnitt «Diskussion» in den Gesamtzusammenhang der verfügbaren Evidenz einzuordnen.»

Clark S, Horton R. Putting research in context – revisited. *Lancet* 2010; 376: 10–11.

Und schließlich machten sie dem Leser, was besonders wichtig ist, sämtliche Ergebnisse zugänglich. Indem sie ihre ursprüngliche systematische Übersichtsarbeit über frühere Studien mit den neuen Erkenntnissen ihrer eigenen Studie aktualisierten, schafften sie die Grundlage dafür, Tausende der durch diese weit verbreitete Therapie verursachten Todesfälle zu verhindern.

- Eine einzelne Studie liefert nur selten genügend Belege, um bei der Entscheidung über Behandlungsoptionen im Gesundheitswesen eine Orientierungshilfe geben zu können.
- Bewertungen der jeweiligen Vorzüge verschiedener Behandlungsoptionen sollten sich auf systematische Übersichtsarbeiten aller relevanten und zuverlässigen Erkenntnisse stützen.
- Wie auch in einzelnen Studien, in denen Therapien getestet werden, müssen für einen systematischen Review entsprechende Schritte unternommen werden, um den irreführenden Einfluss von systematischen Fehlern und Zufallsfaktoren einzugrenzen.
- Die Nichtberücksichtigung der Ergebnisse systematischer Reviews führt zu eigentlich vermeidbaren Schäden bei Patienten und im Gesundheitswesen und in der Gesundheitsforschung zur Vergeudung von Ressourcen.

9

Reglementierung von Therapietests: hilfreich oder hinderlich?

Inzwischen wissen Sie, dass Therapien leider nur allzu oft nicht sorgfältig bewertet werden und dass unnötigerweise auch weiterhin Unsicherheit über die Wirksamkeit mancher Therapien herrscht. Paradoxerweise halten etliche Vorurteile, wie bereits in Kapitel 5 angemerkt, Ärzte davon ab, mit Patienten zusammenzuarbeiten, um auf diese Weise mehr über Therapieeffekte zu erfahren. So merkwürdig es auch anmuten mag, in den meisten Ländern tragen auch die für die Reglementierung der medizinischen Forschung verantwortlichen Institutionen zu diesem Problem noch bei, indem sie eine künstliche Aufteilung zwischen Forschung und Therapie erzwingen. Forschung gilt als hochriskante Tätigkeit, die nach einer strengen Aufsicht verlangt. Routinetherapien dagegen werden als deutlich weniger problematisch gesehen – und das, obwohl Patienten, wie wir gesehen haben, bei der Verabreichung von nicht oder nur unzureichend bewerteten Therapien außerhalb wissenschaftlicher Studien Gefahren ausgesetzt sein können.

Wer sagt, medizinische Forschung sei schlecht für Ihre Gesundheit?

«Die meisten Diskussionen über Ethik in der medizinischen Forschung betreffen die Frage, wie die Forschung reguliert werden soll. Tatsächlich ist die medizinische Forschung in vielerlei Hinsicht sehr viel strenger reglementiert als die medizinische Praxis. Und wenn Sie sich die unzähligen Richtlinien zur medizinischen Forschung einmal durchlesen, dann kann man Ihnen den Gedanken, dass medizinische Forschung genauso wie das Rauchen schlecht für Ihre Gesundheit sein muss, kaum verübeln.»

Hope T. *Medical ethics: a very short introduction*. Oxford: Oxford University Press, 2004, S. 99.

Warum gilt Forschung als so riskant und reglementierungsbedürftig, nicht aber die alltäglichen Therapien, von denen sehr viel mehr Patienten betroffen sind? Natürlich darf man nicht vergessen, dass Wissenschaftler in der Vergangenheit auch Missbrauch getrieben haben, u. a. in Experimenten, in denen Patienten ausgenutzt und als Mittel zum Zweck benutzt wurden. Und von Zeit zu Zeit geht in der Forschung auch tatsächlich etwas schief, sodass es also durchaus die eine oder andere Horrorgeschichte gibt. Zudem besteht auch immer die Sorge, dass Ärzte die persönlichen Interessen der Menschen, wenn diese erst einmal zu Probanden der Forschung geworden sind, den Gesamtinteressen der Forschung unterordnen.

Noch komplizierter wird die Situation durch die sehr unterschiedlichen Motive von Forschern: Während einige Forscher Studien in erster Linie zum Wohl der Allgemeinheit durchführen, stehen bei anderen als Motivation eindeutig Geld oder die Karriere im Vordergrund. Und manchmal lassen sich die Motive der Wissenschaftler auch gar nicht so genau beurteilen. Aus diesen Gründen kann Forschung auf die Patienten und die Öffentlichkeit durchaus benächtigend wirken. Das ist teilweise auch der Grund, warum Forschung im Gesundheitswesen so starken Reglementierungen unterliegt.

Unabhängige Gremien in der Forschung, die allgemein als Ethik-Kommissionen (engl. *Research Ethics Committees*, RECs, oder in den USA z. B. *Institutional Review Boards*, IRBs, in Deutschland die Ethik-Kommissionen an den medizinischen Fakultäten sowie in jedem Bundesland) bezeichnet werden, tragen dazu bei, Menschen vor Missbrauch im Namen der Wissenschaft zu schützen. Sie prüfen jedes Forschungsvorhaben und nehmen dazu Stellung, ob ein Projekt durchgeführt werden kann oder nicht. Zudem spielen sie eine wichtige Rolle im Rahmen der Forschungsaufsicht und geben der Öffentlichkeit die Gewissheit, dass ihre Interessen bei der Planung der zugelassenen Studien Berücksichtigung finden.

Ethik-Kommissionen bestehen häufig aus unbezahlten freiwilligen Experten, u. a. auch aus Laien. Sie prüfen verschiedenste Arten von Studienprotokollen (was die Wissenschaftler in ihren Forschungsvorhaben zu tun beabsichtigen, s. Kap. 8) und darüber hinaus auch alle Informationen, die den potenziellen Teilnehmern der

Studie ausgehändigt werden. Ethik-Kommissionen können von den Wissenschaftlern Änderungen an ihren Studienprotokollen oder an den Informationen für Studienteilnehmer verlangen. Eine Studie darf ohne Zustimmung der Ethik-Kommission nicht durchgeführt werden. Ethik-Kommissionen sorgen also dafür, dass Studienteilnehmer keinen unnötigen Risiken ausgesetzt werden. Zudem geben sie sowohl den Teilnehmern als auch der Öffentlichkeit die Gewissheit, dass die Wissenschaftler nicht einfach tun können, was sie wollen.

Darüber hinaus unterliegt die Forschung auch noch vielen anderen Formen der Reglementierung. Die meisten Länder verfügen über forschungsspezifische gesetzliche Regelungen. So müssen sich beispielsweise alle Länder der Europäischen Union an die *Clinical Trials Directive* (Richtlinie 2001/20/EC) halten, in denen die Anforderungen an die Durchführung sogenannter «klinischer Prüfungen von Medizinprodukten» niedergelegt sind – im Wesentlichen geht es hierbei um Medikamentenstudien. In mehreren Ländern sind zudem Gesetzes- und Regelwerke in Kraft, die auch auf alle oder wenigstens die meisten Formen medizinischer Forschung Anwendung finden. Potenziell können sich auch viele andere Gesetze auf die Forschung auswirken, selbst wenn sie vorrangig nicht im Zusammenhang mit der Forschung erlassen wurden. So haben z. B. in vielen Ländern die Datenschutzgesetze, welche die Vertraulichkeit personenbezogener Daten schützen sollen, auch für die medizinische Forschung Gültigkeit. In den meisten Ländern sind üblicherweise auch verschiedene Behörden in die Forschungsreglementierung eingebunden.

Zudem wird die Durchführung von Forschungsprojekten auch durch das jeweilige Landes- oder Berufsrecht sowie durch internationale Vereinbarungen reguliert. Die Ärzteschaft und das Pflegepersonal sind beispielsweise an die Landes- oder Berufsordnungen ihrer Berufsverbände gebunden: Wenn sie diese Ordnungen verletzen, können sie ihre Zulassung verlieren oder müssen mit anderen Sanktionen rechnen. Bei der Festlegung von Standards sind oft auch internationale Erklärungen wie die Deklaration des Weltärztebundes (engl. *World Medical Association*) von Helsinki maßgeblich, selbst wenn sie keine Gesetzeskraft haben. In deutscher Übersetzung

ist die Deklaration von Helsinki auf den Webseiten der Bundesärztekammer zu finden (www.bundesaerztekammer.de/downloads/deklHelsinki2008.pdf).

Sind Reglementierungssysteme das Richtige für Therapiestudien?

Ein solches Ausmaß an Reglementierung kann zwar beruhigend wirken, doch erwachsen aus den derzeitigen Reglementierungssystemen vor allem denjenigen erhebliche Belastungen, die eine unzulänglich bewertete Therapie testen wollen, anstatt sie ihren Patienten einfach im normalen Praxisalltag anzubieten. In vielen Ländern ist allein die Komplexität dieses Systems – Gesetze, Behörden, Berufsordnungen usw. – erdrückend und zeitaufwändig. Wissenschaftler müssen sich unter Umständen mehrere Genehmigungen von unterschiedlichen Stellen beschaffen und sehen sich gelegentlich auch mit widersprüchlichen Anforderungen konfrontiert.

Darüber hinaus kann das System in seiner Gesamtheit auch die Erhebung der Informationen, welche die allgemeine Gesundheitsversorgung sicherer machen würden, ernsthaft beeinträchtigen und verlangsamen. Auch wenn sie in bester Absicht erlassen wurden, machen es beispielsweise die Datenschutzgesetze und Verfahrensregeln für den Umgang mit vertraulichen Daten den Forschern extrem schwer, Routinedaten aus Krankenakten zu erheben, die zur punktgenauen Bestimmung von Therapienebenwirkungen beitragen könnten. Bei der Planung klinischer Studien kann es von der Forschungsidee bis zur Rekrutierung des ersten Patienten mehrere Jahre dauern, und selbst dann kann die Rekrutierung von Studienteilnehmern durch die Bedingungen der Zulassungsbehörden noch verzögert werden. Und während die Wissenschaftler versuchen, ihre Studien durch das System zu manövrieren, müssen Menschen unnötig leiden und sterben.

In der Praxis bedeutet dies, dass Ärzte einem Patienten, wenn er seine Zustimmung dazu gibt, ungeprüfte Therapien verabreichen können, solange diese Therapien im Rahmen des «routinemäßigen» ärztlichen Alltags erfolgen. Dagegen würde die Durchführung einer

In einer idealen Welt

«In einer idealen Welt, soweit eine solche möglich erscheint, könnten wir anonymisierte Ergebnisdaten erheben und sie mit der Medikamentenanamnese vergleichen; Ausnahmen würden nur bei denjenigen gemacht, die ihre Bedenken hinsichtlich des Persönlichkeitsschutzes über das Leben anderer Menschen stellen ... In einer idealen Welt würden Patienten, wenn sie sich in Behandlung begeben und ernsthafte Unsicherheiten hinsichtlich der für sie besten Therapie bestünden, einfach und effizient zu einer Therapie randomisiert und der Verlauf beobachtet werden. In einer idealen Welt wären diese Vorgänge einfach Teil unseres Verständnisses von Gesundheitsversorgung, sodass kein Patient sich darüber Gedanken machen müsste.»

Goldacre B. Pharmaco-epidemiology would be fascinating enough even if society didn't manage it really really badly. *The Guardian*, 17. Juli 2010. Online zugänglich unter: www.badscience.net/2010/07/pharmaco-epidemiology-would-be-fascinating-enough-even-if-society-didnt-manage-it-really-really-badly

Studie, in der dieselben Therapien sachgemäß beurteilt würden, bedeuten, dass man den langwierigen Reglementierungsprozess durchlaufen müsste. Deshalb schrecken Ärzte vor einer fairen Bewertung medizinischer Behandlungen zurück. Stattdessen verordnen sie weiterhin Therapien und sind nicht verpflichtet, sich mit etwaigen Unsicherheiten über diese Therapien auseinanderzusetzen (s. Kap. 5).

Aus Sorge um die Gefährdung potenzieller Studienteilnehmer und deren Schutz sind die Systeme zur Reglementierung der Forschung überprotektiv geworden und übersehen dabei die Tatsache, dass Patienten und Öffentlichkeit zunehmend als Partner am Forschungsprozess mitwirken (s. Kap. 11). Es ist allerdings ermutigend, dass die Reglementierungsbehörden langsam erkennen, dass ein «Einheitskonzept» für die Prüfung der ethischen Aspekte von Forschungsvorhaben unnötig hinderlich ist.¹ In Großbritannien beispielsweise werden inzwischen Verfahrensweisen für eine «angemessene Prüfung» evaluiert, um herauszufinden, ob ein vereinfachter und schnellerer Prüfvorgang sich bei wissenschaftlichen Studien, die keine grundlegenden ethischen Probleme aufwerfen, unbedenklich anwenden lässt.

Die Voreingenommenheit der Ethik

«Wenn ein Arzt eine neue Therapie ausprobiert, mit der Absicht, sie sorgfältig zu beobachten, ihre Auswirkungen zu bewerten und die Ergebnisse zu veröffentlichen, dann betreibt er Forschung. Die Objekte dieser Forschung (d. h. die Versuchspersonen/Probanden), so die Idee, bedürfen eines besonderen Schutzes. Sein Studienprotokoll muss von einer Ethik-Kommission geprüft werden. Auch die Einwilligungserklärung der Teilnehmer nach erfolgter Aufklärung (engl. informed consent) wird sorgfältig geprüft, und unter Umständen wird das Forschungsvorhaben untersagt. Andererseits kann ein Arzt diese neue Therapie [in seiner Praxis] ausprobieren, ohne sie genauer untersuchen zu wollen, bloß weil er davon überzeugt ist, dass seine Patienten davon profitieren. In diesem Fall stellt sein Ausprobieren der neuen Therapie keine Forschung dar: Für den <Therapieversuch> benötigt er keine Genehmigung seitens der Ethik-Kommission, und das Einholen der Zustimmung seiner Patienten wird lediglich vom Risiko einer Arzthaftungsklage kontrolliert.

Es scheint, dass die Patienten im zweiten Szenario (ohne Forschung) ein deutlich höheres Risiko tragen als die Patienten im ersten Szenario (als Teilnehmer an einer offiziellen klinischen Studie). Zudem scheint der Arzt im ersten Szenario vom ethischen Standpunkt her mehr Anerkennung zu verdienen. Er führt eine Bewertung der Therapie durch, wohingegen der Arzt im zweiten Szenario die Therapie auf der Grundlage seiner unvollkommenen Intuition anwendet. Da im Mittelpunkt ethischer Verhaltensrichtlinien, die den Schutz des Patienten im Sinn haben, aber das Ziel der Schaffung verallgemeinerbarer Erkenntnisse steht, reglementieren sie den verantwortlichen Studienarzt, nicht aber den verantwortungslosen Abenteurer.»

Lantos J. Ethical issues – how can we distinguish clinical research from innovative therapy? *American Journal of Pediatric Hematology/Oncology* 1994; 16: 72–75.

Information und Einverständnis

Auch die Anforderungen seitens der Reglementierungssysteme an die Bereitstellung von Patienteninformationen und die Zustimmung zur Studienteilnahme dienen eher der Abschreckung als dem Anreiz, sich mit Unsicherheiten bezüglich medizinischer Therapien auseinanderzusetzen. Es ist wichtig – und ethisch unabdingbar – die Interessen all derjenigen zu berücksichtigen, die sich derzeit in Behandlung befinden, und nicht nur der wenigen Personen, die an kontrollierten Studien teilnehmen.² Deshalb sollte der Standard für die Einwilligung in eine Therapie nach erfolgter Aufklärung («informierte Einwilligung») für alle Personen gleich sein, egal, ob sie die Therapie innerhalb oder außerhalb offizieller Therapiestudien erhalten. Um eine Entscheidung treffen zu können, die mit ihren

Wertvorstellungen und Präferenzen in Einklang steht, sollten den Patienten die nötigen Informationen in dem von ihnen gewünschten Umfang und zu dem von ihnen gewünschten Zeitpunkt zugänglich gemacht werden.

Wenn im Praxisalltag eine Behandlung angeboten oder verordnet wird, wird davon ausgegangen, dass die einzelnen Patienten jeweils unterschiedliche Vorlieben und Bedürfnisse haben, die sich im Laufe der Zeit zudem auch ändern können. Ferner gilt als anerkannt, dass die Menschen nicht nur unterschiedliche Ansprüche an die Menge oder die Art der gewünschten Informationen haben, sondern dass sie auch unterschiedlich befähigt sind, all diese Informationen in der verfügbaren Zeit zu verarbeiten, und sich auch in der Ausprägung ihrer Sorgen und Ängste unterscheiden. Ärzte sind aufgefordert, ihre Patienten bei Behandlungsentscheidungen einfühlsam zu unterstützen und jederzeit auf die jeweiligen Bedürfnisse des Einzelnen einzugehen.

Die Einwilligungserklärung noch einmal überdenken

«[Manche] sind zu der Auffassung gelangt, dass die «informierte Zustimmung» [Einwilligungserklärung] für eine gute biomedizinische Praxis gar nicht wesentlich ist, und ... alle Versuche, sie dazu zu machen, weder notwendig noch durchführbar sind. Wir hoffen, dass der im Laufe der letzten 50 Jahre geschaffene Moloch aus Einwilligungserfordernissen nun in einem deutlich kürzeren Zeitraum reformiert und verkleinert werden wird.»

Manson NC, O'Neill O. *Rethinking informed consent in bioethics*. Cambridge: Cambridge University Press, 2007, S. 200.

In der Forschung wird die Aufklärung potenzieller Studienteilnehmer von den Reglementierungsbehörden beaufsichtigt, die oftmals darauf bestehen, dass schon zum Zeitpunkt der Einladung zur Studienteilnahme eine möglichst umfassende Weitergabe aller potenziell relevanten Informationen erfolgt. Bei Patienten, die es vorziehen, die Entscheidung ihrem Arzt zu überlassen, könnte dies unnötigerweise zu Aufregung, Frustration oder Angst führen bzw. unnötige Bedenken wecken.³

Die klinische Studie zur Verabreichung von Koffein an Frühgeborenen, die wir in Kapitel 5 (S. 98) erwähnt haben, ist ein anschauliches Beispiel dafür, welcher Schaden angerichtet werden kann,

wenn man auf einer möglichst umfassenden Aufklärung der Kandidaten für wissenschaftliche Studien beharrt. Für die Koffeinstudie wurden weltweit mehr als 2 000 Frühgeborene rekrutiert, aber dieser Vorgang dauerte ein Jahr länger als erwartet, weil die Rekrutierung für die Studie nur schleppend verlief. Besonders langsam erfolgte sie in Großbritannien, wo mehrere Studienzentren sich wegen regulatorischer Verzögerungen im Genehmigungsprozess aus der Studie zurückzogen. Zu allem Überfluss bestand die Ethik-Kommission auch noch darauf, die Eltern darüber aufzuklären, dass Koffein bei den Säuglingen Krampfanfälle auslösen könne – obwohl diese Komplikation erst nach einer zehnfachen Überdosierung auftrat. Also wurden die Eltern mit offensichtlich beängstigenden Informationen konfrontiert, die sie wahrscheinlich gar nicht gebraucht und vermutlich auch nicht erhalten hätten, wenn das Koffein im Rahmen einer Routinebehandlung verabreicht worden wäre.

Eine vernünftige Einstellung zur Patienteneinwilligung im Rahmen guter medizinischer Praxis

«Was in der Debatte um die Patienteneinwilligung fehlt, ist die Frage, was Patienten bereits wissen, welche Informationen sie haben möchten und wie man mit Patienten umgehen soll, die nur das Nötigste wissen wollen. Es gibt nur wenige Arbeiten, in denen untersucht wird, wie Patienten die Informationen, die sie erhalten, verstehen. Ärzten fällt es oftmals schwer zu beurteilen, in welchem Umfang ihre Patienten oder deren Angehörige die Informationen richtig verstehen. Wie viel verstanden wird, richtet sich nach dem Informationsgeber, nach seinen Erläuterungen und nach der Zeit oder der Umgebung, die für die Informationsaufnahme erforderlich ist. Ein paternalistisches Vorgehen ist in der medizinischen Praxis nicht mehr akzeptabel; gute medizinische Praxis setzt eine vernunftorientierte Vorgehensweise voraus, bei welcher der Sachverhalt verständlich erklärt wird, das Gesagte sich daran orientiert, was der Patient anscheinend möchte, und überprüft wird, ob er es verstanden hat.»

Gill R. How to seek consent and gain understanding. *BMJ* 2010; 341: c4000.

Es gibt kaum Belege dafür, dass die gemeinhin propagierten Formen der Forschungsreglementierung mehr nutzen als schaden.⁴ Das Wenige, was uns an Evidenz vorliegt, ist freilich verstörend. So kann beispielsweise in Studien, in denen die Wirkungen von Therapien untersucht werden, die akut verabreicht werden müssen (z. B. in Notfallsituationen), das «Ritual» der obligatorischen Einholung der

schriftlichen Patienteneinwilligung nicht nur zu Todesfällen führen, die vermeidbar wären, sondern auch zu einer Unterschätzung der Therapieeffekte.⁵

Das Einholen der Patienteneinwilligung stellt eine Public-Health-Intervention dar, die mitunter mehr Schaden anrichtet als Nutzen bringt. Wie bei anderen gut gemeinten Interventionen sollten ihre Wirkungen einer strengen Bewertung unterzogen werden. Die tödlichen Folgen, die wir an anderer Stelle beschrieben haben, hätten vielleicht schon vor Jahrzehnten erkannt werden können, wenn auch die Ethik-Kommissionen verpflichtet wären, solide Belege dafür zu erbringen, dass ihre «Verordnungen» mehr nützen als schaden.

Flexibilität bei der Aufklärung potenzieller Studienteilnehmer, die anerkennt, dass das Vertrauen zwischen Arzt und Patient die Grundlage für eine zufriedenstellende Konsultation darstellt, ist besser als ein starres, standardisiertes Vorgehen. Aber aufgrund der Art und Weise, wie die Reglementierungssysteme in die Forschung eingreifen, können Ärzte derzeit nicht frei entscheiden, wie sie ihre Patienten über eine wissenschaftliche Studie aufklären. Zudem fällt es ihnen oft schwer, über die der Forschung innewohnenden Unsicherheiten zu sprechen. Wie wir in Kapitel 5 erwähnt haben, fühlen sich Ärzte bei der Rekrutierung von Patienten für klinische Studien

Akademischer Schnickschnack oder sinnvolle Entscheidung?

«Vor zwölf Jahren überschritt ich die Arzt-Patient-Grenze, als bei mir im Alter von 33 Jahren Brustkrebs festgestellt wurde. Damals saß ich gerade an meiner Doktorarbeit, in der es um die Schwierigkeiten der Anwendung von randomisierten kontrollierten Studien (RCTs) bei der Bewertung der Wirksamkeit von Therapien in meinem eigenen Fach, der Kieferorthopädie, ging. Im Rahmen meiner Forschungsarbeiten hatte ich erkannt, welche Vorteile die Teilnahme an klinischen Studien hatte, und wusste ironischerweise auch um die Unsicherheiten hinsichtlich der Behandlung von jüngeren Frauen mit Brustkrebs im Frühstadium. Zum Zeitpunkt der Diagnosestellung fragte ich deshalb meinen Arzt, ob es irgendwelche RCTs gäbe, an denen ich teilnehmen könnte. Seine Antwort schockierte mich. Er meinte, ich dürfe nicht zulassen, dass die für mich am besten geeignete Therapie durch wissenschaftliche Spitzfindigkeiten torpediert würde. Doch was war die beste Therapie? Ich jedenfalls wusste es nicht und musste zudem feststellen, dass auch in der Ärzteschaft Zweifel an der optimalen Therapie für Frauen unter 50 Jahren mit Brustkrebs im Frühstadium herrschten. Was sollte ich nur tun?»

Harrison J. Testing times for clinical research. *Lancet* 2006; 368: 909–910.

oftmals unbehaglich, wenn sie sagen müssen: «Ich weiß nicht, welche Behandlung die beste ist», und die Patienten wollen das oft auch gar nicht hören. Sowohl Ärzte als auch Patienten müssen Unsicherheiten besser einschätzen und verstehen können, warum Forschung wichtig ist (s. Kap. 11).

Was Reglementierungssysteme nicht tun

Die Reglementierungssysteme für die Forschung verlangen von Wissenschaftlern, dass sie vor Beginn einer Studie hohe Auflagen erfüllen müssen, und doch gibt es auffallend viele Dinge, die sie nicht oder nicht zufriedenstellend regeln. Zahlreiche Systeme stellen nicht ausreichend sicher, dass die geplanten Studien auch tatsächlich benötigt werden – beispielsweise verlangen sie von den Wissenschaftlern keinen Nachweis darüber, dass sie die bereits vorhandenen Erkenntnisse gründlich ausgewertet haben, bevor sie eine neue Studie in Angriff nehmen (zur Frage, warum systematische Reviews so wichtig sind, s. Kap. 8).

Zudem befinden sich die meisten Maßnahmen zur Reglementierung der Forschung noch im Anfangsstadium, wobei das Hauptaugenmerk auf der Kontrolle der Aufnahme von Teilnehmern in die Studien liegt. Erstaunlich wenige Anstrengungen gelten der Überwachung laufender Studien und der Gewährleistung, dass die Forscher ihre Studienberichte am Ende ihrer Forschungsarbeiten umgehend (bzw. überhaupt) veröffentlichen und angeben, inwieweit ihre Ergebnisse zur Verringerung von Unsicherheiten beitragen.

Was Forschungsreglementierung leisten sollte

«Wenn Ethiker und andere Kritik an klinischen Studien üben wollen, dann sollten sie ihr Augenmerk auf wissenschaftlich unzulängliche Arbeiten, auf die Neuerfindung des Rades und vor allem auf grundlose Ausschlüsse sowie die ungerechtfertigte und unsinnige Inanspruchnahme von Ressourcen richten. Ein Schwachpunkt der aktuellen Debatte ist, dass sie dem Zweck von Studien keine Beachtung schenkt – nämlich zu gewährleisten, dass die in Gebrauch befindlichen Therapien sicher sind und auch wirklich das leisten, was sie besser können als ihre Alternativen. In der Ethik gibt es keinen Königsweg – ebenso wenig wie in Studien.»

Ashcroft R. Giving medicine a fair trial. *BMJ* 2000; 320: 1686.

Menschen, die zur Teilnahme an einer Studie über Therapieeffekte eingeladen werden, brauchen Gewissheit, dass die Studien diese Mühe wert sind und dass ihr persönlicher Beitrag anderen nutzt. Reglementierungssysteme müssen noch mehr tun, um ihnen in beiderlei Hinsicht Gewissheit zu geben. Darüber hinaus müssen sie unnötige Hindernisse für solide Forschung abbauen, deren Ziel die Beantwortung patientenrelevanter Forschungsfragen ist. Zunehmend wächst die Einsicht, dass die Überprüfung der Wirksamkeit von Therapien uns alle angeht. Da Patienten und Öffentlichkeit die sich nun bietenden Gelegenheiten zur Mitwirkung an der Planung und Durchführung von Forschungsvorhaben aufgreifen (s. Kap. 11), werden sie wahrscheinlich auch ein zunehmendes Mitspracherecht haben, wenn es um die Auseinandersetzung mit regulatorischen Hindernissen geht.

- Die Forschungsreglementierung ist unnötig kompliziert.
- Die derzeitigen Forschungsreglementierungssysteme behindern die Durchführung von fairen Therapietests, die eine bessere Gesundheitsversorgung gewährleisten würden.
- Trotz der lästigen Reglementierungsanforderungen, denen Wissenschaftler unterliegen, geben die Reglementierungssysteme aber kaum eine Gewähr, dass die vorgeschlagenen Studien auch wirklich gebraucht werden.
- Die Forschungsreglementierung trägt nur wenig zur Überwachung und Nachbeobachtung genehmigter Forschungsvorhaben bei.

10

Gute, schlechte und überflüssige klinische Forschung

In den vorangegangenen Kapiteln haben wir Fragen angesprochen, die für Patienten und Öffentlichkeit eine wichtige Rolle spielen, und wir haben dargelegt, warum Therapietests richtig geplant werden müssen. Wenn dies gelingt, dann können wir alle mit uns und den Ergebnissen zufrieden sein, selbst wenn sich die erhofften Vorteile nicht bestätigen, weil auf diese Weise wichtige Einsichten gewonnen und Unsicherheiten abgebaut werden konnten.

Obschon ein Großteil der Gesundheitsforschung von hoher Qualität ist und sich zudem ständig verbessert, solange Planungs- und Berichtsstandards¹ erfüllt werden –, so wird aus verschiedenen Gründen aber auch weiterhin schlechte und überflüssige Forschung durchgeführt und veröffentlicht. Und was die gebetsmühlenartig vorgetragene Forderung «es bedarf weiterer Forschungsarbeiten» angeht, so wäre man manchmal besser beraten, weniger zu tun, dabei aber den Schwerpunkt der Forschung auf die Patientenbedürfnisse zu legen, damit sichergestellt wird, dass aus den richtigen Gründen geforscht wird. Mit diesen Aspekten wollen wir uns im vorliegenden Kapitel beschäftigen.

Gute Forschung

Schlaganfall

Schlaganfälle gehören zu den führenden Ursachen für Tod und Langzeitbehinderung. Die Sterblichkeit liegt zwischen 1 : 6 und 2 : 6

beim ersten Schlaganfall und steigt auf 4 : 6 bei weiteren Schlaganfällen an. Eine Ursache, die dem Schlaganfall zugrunde liegt, ist eine Verengung (Stenose) der Halsschlagader (Arteria carotis), die das Gehirn mit Blut versorgt. Zuweilen löst sich das Fettgewebe, das die Innenseiten der Halsschlagader auskleidet, verstopft kleinere arterielle Zuflüsse und verursacht dadurch einen Schlaganfall. Um diese Fettablagerungen zu entfernen, führten Gefäßchirurgen in den 1950er-Jahren erstmals eine Operation, die sogenannte Endarteriektomie, durch. Man hoffte, das Schlaganfallrisiko durch diesen Eingriff zu verringern. Wie bei jeder Operation besteht aber auch hier ein durch den Eingriff selbst bedingtes Komplikationsrisiko.

Trotz zunehmender Popularität der Karotis-Endarteriektomie wurden erst in den 1980er-Jahren randomisierte Studien begonnen, um Vorteile und Risiken der Operation zu bewerten. Man nahm an, dass die dabei gewonnenen Erkenntnisse sowohl für die Patienten als auch ihre Ärzte von entscheidender Bedeutung sein würden. Um den chirurgischen Eingriff mit der besten verfügbaren nichtchirurgischen Therapie zu vergleichen, wurden mit Patienten, die bereits Symptome einer Karotisstenose aufwiesen (leichter Schlaganfall oder flüchtige schlaganfallähnliche Symptome), zwei gut geplante Studien durchgeführt – eine in Europa, die andere in Nordamerika. An diesen Langzeitstudien nahmen mehrere tausend Patienten teil. Die Ergebnisse, die in den 1990er-Jahren veröffentlicht wurden, zeigten, dass die Operation das Schlaganfall- bzw. das Mortalitätsrisiko senken kann, dass dieser Vorteil aber vom Grad der Stenose in der Halsschlagader abhängig ist. Für Patienten mit einer vergleichsweise geringfügigen Verengung war die Operation, die selbst einen Schlaganfall auslösen kann, alles in allem eher von Nachteil. Diese wichtigen Erkenntnisse hatten unmittelbare Auswirkungen auf die klinische Praxis.^{2, 3}

Präeklampsie in der Schwangerschaft

Ein weiteres hervorragendes Beispiel für gute Forschung betrifft schwangere Frauen. Weltweit versterben jährlich ca. 600 000 Frauen an schwangerschaftsbedingten Komplikationen. Die meisten dieser Todesfälle ereignen sich in Entwicklungsländern, und viele davon

stehen mit schwangerschaftsbezogenen Krampfanfällen in Zusammenhang, einer Erkrankung, die man als Eklampsie bezeichnet. Die Eklampsie ist eine verheerende Krankheit, an der sowohl Mutter als auch Kind versterben können. Frauen mit der dafür prädisponierenden Erkrankung, der sogenannten Prä-Eklampsie (auch als schwangerschaftsbedingte Toxämie bezeichnet), weisen Bluthochdruck und Eiweiß im Urin (Proteinurie) auf.

1995 zeigten Forschungsergebnisse, dass sich das *Wiederauftreten* (Rezidive) von Krampfanfällen bei Frauen mit Eklampsie durch die Injektion von Magnesiumsulfat, einem einfachen und preiswerten Medikament, verhindern ließ. Dieselbe Studie ergab zudem, dass Magnesiumsulfat zur Beendigung von Krampfanfällen besser war als andere krampf lösende Medikamente (Antikonvulsiva), zu denen auch ein deutlich teureres Medikament gehörte. Deshalb war den Wissenschaftlern klar, wie wichtig es war herauszufinden, ob Magnesiumsulfat auch das *Auftreten* von Krampfanfällen bei Frauen mit Prä-Eklampsie verhindern konnte.

Die Magpie-Studie, mit der diese Frage beantwortet werden sollte, war ein größeres Unterfangen, an dem in 33 Ländern auf der ganzen Welt mehr als 10 000 Schwangere mit Prä-Eklampsie teilnahmen. Zusätzlich zur normalen medizinischen Versorgung erhielt eine Hälfte der Frauen eine Injektion mit Magnesiumsulfat und die andere Hälfte eine Injektion mit einem Placebo (einem Scheinmedikament). Die Magpie-Studie lieferte eindeutige und überzeugende Ergebnisse. Sie zeigte, dass Magnesiumsulfat das

Meine Erfahrungen mit der Magpie-Studie

«Ich war wirklich froh, an einer so wichtigen Studie teilnehmen zu können. In der 32. Schwangerschaftswoche begannen bei mir Schwellungen durch Wassereinlagerungen aufzutreten, die immer schlimmer wurden, bis bei mir schließlich eine Prä-Eklampsie diagnostiziert und ich in der 38. Woche ins Krankenhaus eingeliefert wurde. Mein Baby wurde durch Kaiserschnitt geboren, und glücklicherweise haben wir uns beide vollständig erholt. Die Prä-Eklampsie ist eine beängstigende Krankheit, und ich hoffe von ganzem Herzen, dass Frauen wie ich von den Ergebnissen der Studie profitieren werden.»

Clair Giles, Teilnehmerin der Magpie-Studie

MRC News Release. Magnesium sulphate halves risk of eclampsia and can save lives of pregnant women. London: MRC, 31. Mai 2002

Risiko für das Auftreten von Krampfanfällen um mehr als die Hälfte verringerte. Darüber hinaus gab es Anhaltspunkte dafür, dass diese Therapie zwar nicht das Sterberisiko des Kindes, wohl aber das der Mutter senken kann. Und abgesehen von geringfügigen Nebenwirkungen schien Magnesiumsulfat weder Mutter noch Baby zu schädigen.^{4,5}

HIV-Infektion bei Kindern

Die Ergebnisse guter Forschung haben auch das Leben von Kindern die mit HIV (dem für die Immunschwächekrankheit AIDS verantwortlichen humanen Immundefizienz-Virus) infiziert sind, entscheidend verbessert. Ende 2009 zeigten die im Rahmen von UNAIDS (dem gemeinsamen HIV/AIDS-Programm der Vereinten Nationen) erhobenen Zahlen, dass weltweit schätzungsweise 2,5 Millionen Kinder mit HIV leben, 2,3 Millionen davon in den südlich der Sahara liegenden afrikanischen Ländern. Damals starben stündlich ca. 30 Kinder an den Folgen von AIDS.⁶ Eine häufige Todesursache sind bakterielle Infektionen wie Lungenentzündungen (Pneumonien), die mit dem geschwächten Immunsystem der Kinder in Zusammenhang stehen. Zur Behandlung von Kindern und Erwachsenen mit nicht durch AIDS bedingten Atemwegsinfektionen wird seit vielen Jahren Co-trimoxazol eingesetzt, ein überall erhältliches preiswertes Antibiotikum. Studien mit HIV-infizierten Erwachsenen hatten zudem ergeben, dass das Medikament auch andere durch bakterielle Infektionen bedingte Komplikationen lindern kann.⁷

Als vorläufige Ergebnisse darauf hindeuteten, dass das Medikament auch die Infektionsrate bei HIV-infizierten Kindern senkte, arbeitete eine Gruppe britischer Wissenschaftler mit Kollegen in Sambia zusammen, um die Wirkungen von Co-trimoxazol als mögliches Prophylaxemedikament in einer großangelegten Untersuchung zu bewerten. In dieser Studie, die 2001 begonnen wurde und ca. zwei Jahre lief, wurde das Antibiotikum bei mehr als 500 Kindern mit Placebo verglichen. Die Ergebnisse zeichneten sich früher als erwartet ab: Es konnte nachgewiesen werden, dass das Medikament sowohl die AIDS-bedingten Todesfälle um 43 % senkte (74 Todes-

fälle in der Co-trimoxazol-Gruppe im Vergleich zu 112 Todesfällen in der Placebo-Gruppe) als auch die Anzahl der Krankenhauseinweisungen reduzierte. Zu diesem Zeitpunkt empfahl das unabhängige Gremium, das die Ergebnisse prüfte, die vorzeitige Beendigung der Studie (weil die Ergebnisse so eindeutig waren und es aus ethischen Gründen nicht vertretbar gewesen wäre, den Kindern der Placebo-Gruppe das Medikament weiter vorzuenthalten).

Die Studie hatte unmittelbar zur Folge, dass alle Kinder, die an der Studie teilgenommen hatten, im Rahmen einer Initiative der sambischen Regierung Co-trimoxazol erhielten. Eine noch weitreichendere Folge war, dass die Weltgesundheitsorganisation und UNICEF ihre Arzneimittelempfehlungen für HIV-infizierte Kinder unverzüglich abänderten.^{8,9}

Von beiden Organisationen wird Co-trimoxazol auch weiterhin als kostengünstige, lebensrettende und sichere Behandlung für HIV-infizierte Kinder empfohlen.¹⁰

Schlechte Forschung

Psychiatrische Störungen

Bedauerlicherweise ist die Durchführung von Forschungsprojekten nicht immer zufriedenstellend bzw. ist die Forschung nicht immer relevant. Sehen wir uns dazu als Beispiel die sogenannte tardive oder Spätdyskinesie an, ein sehr belastendes Beschwerdebild. Dabei handelt es sich um eine schwerwiegende Nebenwirkung, die mit der Langzeitanwendung von Neuroleptika (Antipsychotika) einhergeht. Diese Medikamente werden bei psychiatrischen Erkrankungen, insbesondere bei Schizophrenie, verordnet. Das auffälligste Merkmal von Spätdyskinesien sind ständig wiederholte, unwillkürliche Bewegungen im Mund- und Gesichtsbereich wie Grimassieren, Schmatzen, häufiges Vorstrecken der Zunge und Einziehen oder Aufblasen der Wangen. Gelegentlich sind sie von Zuckungen im Bereich der Hände und Füße begleitet. Diese Nebenwirkungen treten bei einem von fünf Patienten auf, die ein Antipsychotikum länger als drei Monate einnehmen.

In den 1990er-Jahren begann eine Gruppe von Wissenschaftlern systematisch zu erforschen, wie Spätdyskinesien im Laufe der letzten 30 Jahre behandelt worden waren. Verwundert stellten sie 1996 zeigten in einem Artikel fest, dass sie ungefähr 500 randomisierte Studien identifiziert hatten, in denen 90 verschiedene medikamentöse Therapien untersucht wurden. Doch hatte keine dieser Studien irgendwelche brauchbaren Daten ergeben. In manchen Studien war die Anzahl der eingeschlossenen Patienten zu gering, um verlässliche Ergebnisse liefern zu können; in anderen waren die Medikamente so kurz verabreicht worden, dass die Studie keine Aussagekraft hatte.¹¹

Mitglieder derselben Forschergruppe veröffentlichten außerdem eine umfassende Übersicht über Inhalt und Qualität von randomisierten Studien, die ganz allgemein für die Behandlung von Schizophrenien relevant waren. Sie sichteten 2000 Studien: Die Ergebnisse waren enttäuschend. Sicherlich haben die Medikamente die Prognose von Schizophrenie-Patienten im Laufe der Jahre in mancherlei Hinsicht verbessert. So können beispielsweise die Patienten mittlerweile meistens zu Hause oder in einer Gemeinschaft leben. Doch selbst in den 1990er-Jahren (und auch heute noch) wurden (bzw. werden) die meisten Medikamente an stationär behandelten Patienten getestet, sodass bezüglich ihrer Relevanz für die ambulante Therapie Unsicherheit herrscht. Überraschend war darüber hinaus auch die uneinheitliche Bewertung der Behandlungsergebnisse. Die Wissenschaftler stellten fest, dass in den Studien mehr als 600 Therapien – hauptsächlich Medikamente, z.B. aber auch Psychotherapie – untersucht, zur Bewertung der Ergebnisse aber 640 verschiedene Bewertungsskalen angewendet worden waren; davon waren 369 lediglich ein einziges Mal benutzt worden. Es war daher kaum möglich, die Ergebnisse der verschiedenen Studien zu vergleichen; zudem waren die Ergebnisse für Ärzte bzw. Patienten kaum interpretierbar. Neben vielen anderen Problemen stießen die Wissenschaftler auf zahlreiche Studien, die zu klein oder zu kurz angelegt waren, um brauchbare Ergebnisse liefern zu können. Zudem wurden neue medikamentöse Therapien häufig mit unangemessen hohen Dosierungen eines für seine Nebenwirkungen bekannten Medikaments verglichen, obwohl besser verträgliche Therapien verfügbar waren – ein ganz offensichtlich unfairer Test. Die Autoren dieser Übersichtsarbeit kamen zu dem Schluss, dass

nach einem halben Jahrhundert von Studien unzureichender Qualität und Dauer sowie eingeschränktem klinischem Nutzen noch immer genügend Spielraum für gut geplante, korrekt durchgeführte und kompetent publizierte Studien blieb.¹²

Epiduralanalgesie bei Wehenschmerzen

Wie wichtig die Bewertung von Behandlungsergebnissen ist, die für Patienten wichtig sind, zeigt sich anschaulich – auf eine sehr negative Weise – an den frühen Studien über die Epiduralanalgesie zur Schmerzlinderung bei der Geburt.

In den 1990er-Jahren überprüften Wissenschaftler die Erfahrungen mit Studien zur Epidural- im Vergleich zur Nicht-Epiduralanalgesie. Obwohl in den vorangegangenen 20 Jahren schätzungsweise Millionen von Frauen eine den Schmerz blockierende Epiduralanalgesie angeboten worden war, nahmen anscheinend noch nicht einmal 600 Frauen an einigermaßen unverzerrten Vergleichen mit anderen Formen der Schmerzlinderung teil. Die Wissenschaftler identifizierten neun Vergleichsstudien, die zuverlässig ausgewertet werden konnten. Gemeinhin wurden bei diesen Vergleichen die Spiegel von Hormonen und anderen Substanzen gemessen, von denen man annimmt, dass sie die Belastungen während der Wehen anzeigen. Zudem galt das Interesse auch den kindlichen Outcomes. In allen bis auf zwei Studien fehlte jedoch ein Vergleich der von den Frauen selbst angegebenen Schmerzen. Mit anderen Worten: Von den Studienärzten war eine mit Sicherheit äußerst wichtige Zielgröße übersehen worden: und zwar der Grad der Schmerzlinderung bei den betroffenen Frauen.¹³

Überflüssige Forschung

Atemnotsyndrom bei Frühgeborenen

Manche Forschungsarbeiten sind weder gut noch schlecht, sondern schlicht überflüssig. Ein Beispiel dafür stammt aus der Frühgeborenenforschung. Wenn Babys zu früh geboren werden, sind ihre Lungen unter Umständen noch unreif, was mit einem Risiko für lebens-

bedrohliche Komplikationen wie dem Atemnotsyndrom einhergeht. Anfang der 1980er-Jahre lagen überwältigende Belege vor, dass sich die Häufigkeit von Atemnotsyndrom und Sterblichkeit bei Neugeborenen verringern ließ, wenn man den frühgeburtsgefährdeten Schwangeren ein steroidhaltiges Medikament verabreichte. Trotzdem wurden während der folgenden zehn Jahre weiterhin Studien durchgeführt, in denen die Gabe von Steroiden mit Placebo oder Nichtbehandlung verglichen wurde. Hätte man die Ergebnisse früherer Studien systematisch ausgewertet und mittels Meta-Analyse zusammengefasst (s. Kap. 7 und 8), wären viele der späteren Studien wahrscheinlich gar nicht erst begonnen worden – die kollektiven Ergebnisse hätten nämlich gezeigt, dass dafür schlicht kein Bedarf bestand. In diesen unnötigen Studien wurde somit der Hälfte der Studienteilnehmerinnen eine wirksame Therapie vorenthalten.

Schlaganfall

Ein anderes Beispiel für überflüssige Forschung – auch hier wieder, weil die Ergebnisse früherer Studien nicht zusammengefasst ausgewertet wurden – betrifft die Behandlung von Schlaganfällen mit einem Medikament namens Nimodipin (das zur Medikamentenklasse der Kalziumantagonisten gehört). Ausgangspunkt war folgende Überlegung: Wenn es gelänge, das Ausmaß der Hirnschädigung bei Schlaganfall-Patienten zu verringern, dann sollte auch das Behinderungsrisiko der Betroffenen abnehmen. Mit Beginn der 1980er-Jahre wurde, nachdem etliche Tierstudien ermutigende Ergebnisse erbracht hatten, Nimodipin deshalb bei Schlaganfall-Patienten untersucht. Obwohl eine 1988 veröffentlichte klinische Studie mit Schlaganfall-Patienten auf einen vorteilhaften Effekt schließen ließ, waren die Resultate verschiedener weiterer klinischer Studien über Nimodipin und andere Kalziumantagonisten aber widersprüchlich. Als man die akkumulierte Evidenz aus diesen klinischen Studien, die nahezu 8 000 Patienten umfassten, 1999 systematisch auswerte- te, waren keine vorteilhaften Effekte dieser Medikamente nachweisbar (s. Kap. 8, S. 152).¹⁴ Doch die Anwendung von Nimodipin beruhte ja anscheinend auf soliden wissenschaftlichen Belegen. Wie war dies also zu erklären?

Im Lichte der Ergebnisse der Patientenstudien wurden die Resultate der tierexperimentellen Studien erstmals richtig geprüft. Erst als man die Tierstudien in einem systematischen Review zusammenfasste, wurde erkennbar, dass die Versuchsplanung in den tierexperimentellen Studien generell unzulänglich war und die Ergebnisse mit systematischen Fehlern (Bias) behaftet und somit unzuverlässig waren. Anders gesagt: Es hatten von Anfang an gar keine überzeugenden Gründe für die Durchführung von Studien mit Schlaganfall-Patienten vorgelegen.¹⁵

Aprotinin: Wirkung auf intra- und postoperative Blutungen

An diesem Fall überflüssiger Forschung haben alle ihren Anteil: Forschungssponsoren, akademische Einrichtungen, Wissenschaftler, Ethik-Kommissionen wie auch wissenschaftliche Fachzeitschriften (s. Kap. 9). Wie wir in Kapitel 8 dargelegt haben und wie auch die ersten beiden Beispiele unnötiger Forschung zeigen, sollten neue Forschungsvorhaben nicht geplant oder umgesetzt werden, ohne die aus früheren Forschungsarbeiten gewonnenen Erkenntnisse systematisch auszuwerten.

Im Mittelpunkt einer 2005 publizierten schockierenden Analyse standen kontrollierte Studien über Aprotinin, ein Medikament zur Eindämmung von Blutungen während (intraoperativ) und nach einem chirurgischen Eingriff (postoperativ). Aprotinin wirkt. Schockierend daran war die Erkenntnis, dass auch noch lange Zeit, nachdem fundierte Belege vorlagen und zeigten, dass das Medikament die Notwendigkeit von Bluttransfusionen erheblich verringerte, weitere kontrollierte Studien zu diesem Thema durchgeführt wurden.¹⁶

Zum Zeitpunkt der Analyse waren die Berichte über 64 Studien veröffentlicht worden. Zwischen 1987 und 2002 fiel der Anteil der relevanten früheren Studienberichte, die in den späteren Berichten über Aprotinin-Studien zitiert wurden, von ursprünglich 33 % auf nur noch 10 % in den neusten Studienberichten. Nur 7 von 44 späteren Studienberichten nahmen Bezug auf den Bericht zur größten Studie (sie war 28-mal größer als die Studien mittlerer Größe); und in keinem der Berichte wurden die systematischen Übersichtsarbeiten

ten zu diesen Studien erwähnt, die 1994 und 1997 publiziert worden waren.

Wie die Autoren der Analyse betonen, sollte Wissenschaft eigentlich kumulativ sein, doch von vielen Wissenschaftlern werden die Forschungsergebnisse nicht wissenschaftlich akkumuliert. Bei der Planung der neusten Studien wird häufig nicht nur versäumt, die systematischen Übersichtsarbeiten zur vorhandenen Evidenz zu berücksichtigen, sondern in den Studienberichten wird oftmals auch vergessen, die neue Evidenz in den Kontext dieser aktualisierten Übersichtsarbeiten zu stellen (s. Kap. 8).

Verkehrte Forschungsprioritäten

Die meisten Organisationen, die biomedizinische Forschung fördern, und die meisten Wissenschaftler, die sie durchführen, verfolgen ein hehres Ziel: nämlich Erkenntnisse beizusteuern, die einen Beitrag zur Verbesserung der menschlichen Gesundheit leisten. Doch wie viele der Millionen von alljährlich veröffentlichten biomedizinischen Forschungsberichten leisten wirklich einen nützlichen Beitrag zur Verwirklichung dieses noblen Ziels?

Für Patienten relevante Fragen

Wissenschaftler aus Bristol beschlossen, einer grundsätzlichen Frage nachzugehen: «In welchem Umfang finden Fragen, die für Patienten mit Kniegelenkarthrose und ihre behandelnden Ärzte relevant sind, in der Forschung zu diesem Krankheitsbild tatsächlich Berücksichtigung?»¹⁷ Dazu stellten sie zunächst vier Fokusgruppen zusammen: Patienten, Rheumatologen, Physiotherapeuten und Allgemeinmediziner. Diese Gruppen waren einhellig der Meinung, dass sie keine weiteren von pharmazeutischen Unternehmen gesponserten Studien wollten, in denen schon wieder ein neues nichtsteroidales Antiphlogistikum (Medikamentenklasse, zu der z.B. auch Ibuprofen gehört) mit einem Placebo verglichen werden würde. Anstelle von Medikamentenstudien wünschten sich die Patienten eine gründliche Bewertung von Physiotherapie und Opera-

tion sowie von Bewegungsschulungen und Bewältigungsstrategien, die Patienten im Umgang mit dieser chronischen, behindernden und oft auch schmerzhaften Krankheit besser unterstützen können.

Diese Formen der Behandlung und Krankheitsbewältigung bieten natürlich einen deutlich geringeren Spielraum für die kommerzielle Nutzung als Medikamente, weshalb man sie häufig einfach ignoriert.

In wie vielen anderen Bereichen der Therapieforschung würde wohl ein ähnliches Ungleichgewicht hinsichtlich der für Patienten und Ärzte relevanten Fragen zu Therapieeffekten und den Fragen, mit denen sich die Wissenschaftler tatsächlich beschäftigen, zu Tage treten, wenn man sie wie oben beschrieben bewerten würde? Leider scheint eine solche Diskrepanz eher die Regel als die Ausnahme zu sein.^{18–21}

Geringfügige Änderungen in der Arzneimittelformulierung haben nur selten zur Folge, dass die Medikamente vollkommen neue und noch vorteilhaftere Wirkungen zeigen. Und doch herrscht in der Therapieforschung diese Art von Studien vor, nicht nur in Bezug auf Arthritis, sondern auch hinsichtlich anderer chronischer Krankheiten. Was für eine Vergeudung von Ressourcen!

Wer entscheidet, was untersucht wird?

Diese Situation ist alles andere als befriedigend. Wie konnte es dazu kommen? Ein Grund ist, dass das, was von Wissenschaftlern untersucht wird, durch äußere externe Faktoren verzerrt wird.²² So betreibt z.B. die pharmazeutische Industrie Forschung vorrangig, um ihre Verantwortung gegenüber den Anteilseignern zu erfüllen, d.h. um Gewinne zu erzielen. Die Verantwortung gegenüber Patienten und Ärzten kommt erst an zweiter Stelle. Unternehmen unterliegen dem Einfluss großer Märkte – etwa dem Markt der Frauen, die vor der Frage stehen, ob sie sich einer Hormonersatztherapie unterziehen sollen, oder dem Markt aller depressiven, ängstlichen oder unglücklichen Menschen oder dem der Schmerzpatienten. Doch hat diese kommerziell ausgerichtete Vorgehensweise in den letzten Jahrzehnten nur selten wichtige neue Therapien hervorgebracht – noch nicht einmal für die sogenannten «Volkskrankheiten». Stattdessen werden

von der Industrie meist viele sehr ähnliche Verbindungen aus einer Medikamentenklasse auf den Markt gebracht, die sogenannten «Me-too»- oder Analogpräparate. Das lässt an die Zeiten denken, als es in [britischen] Supermärkten als einzige Brotsorte nur geschnittenes weißes Toastbrot zu kaufen gab – das aber in unendlich vielen Varianten. Daher ist es kaum verwunderlich, dass die pharmazeutische Industrie mehr Geld für das Marketing als für die Forschung ausgibt.

Aber wie schafft die Industrie es, die verordnenden Ärzte davon zu überzeugen, dass diese neuen Produkte besser sind als die schon vorhandenen preiswerteren Alternativen? Eine gebräuchliche Strategie ist die Vergabe zahlreicher kleinerer Forschungsprojekte, mit denen man nachweist, dass es besser ist, die neuen Medikamente zu verabreichen als gänzlich auf Medikamente zu verzichten. Über die Frage, ob die neuen Medikamente besser sind als die bereits vorhandenen wird dagegen überhaupt nicht geforscht. Leider fällt es der Industrie nicht besonders schwer, Ärzte zu finden, die bereit sind, ihre Patienten in ein solches nutzloses Unterfangen aufzunehmen. Und diese Ärzte sind es dann auch, die auf diese Weise beforschten Medikamente letztes Endes verschreiben.²³ Zu allem Überfluss

Auswirkungen von Analogpräparaten in Kanada

«In British Columbia wurde ein Großteil (80 %) des Anstiegs bei den Arzneimittelkosten zwischen 1996 und 2003 mit der Anwendung neuer patentierter Medikamente erklärt, die gegenüber den vor 1990 erhältlichen preiswerteren Alternativen aber keine wesentlichen Verbesserungen aufwiesen. Die steigenden Kosten dieser Me-too-Präparate, deren Preise diejenigen von seit langem bewährten Konkurrenzprodukten deutlich übersteigen, bedürfen einer sorgfältigen Prüfung. Konzepte für die Preisgestaltung bei Arzneimitteln wie in Neuseeland ermöglichen unter Umständen Einsparungen, die für die Finanzierung anderer Erfordernisse im Gesundheitswesen nutzbar gemacht werden könnten. So hätten in British Columbia \$ 350 Millionen (26 % der Gesamtausgaben für verschreibungspflichtige Medikamente) eingespart werden können, wenn die Hälfte der 2003 verbrauchten Me-too-Präparate sich preislich an den älteren Alternativen orientiert hätte. Mit diesen Einsparungen könnte man die Honorare von mehr als tausend neuen Ärzten bezahlen.

Geht man davon aus, dass die Liste der 20 weltweit umsatzstärksten Medikamente auch neu patentierte Versionen von Medikamenten aus lange bekannten Medikamentenklassen enthält ... so werden die Ausgabentrends in den meisten Industrienationen wahrscheinlich von den Me-too-Präparaten beherrscht.»

Morgan SG, Bassett KL, Wright JM, et al. «Breakthrough» drugs and growth in expenditure on prescription drugs in Canada. *BMJ* 2005; 331: 815–816.

verschärfen die Arzneimittelzulassungsbehörden das Problem häufig noch dadurch, dass sie darauf bestehen, dass neue Medikamente gegen Placebo und nicht mit den schon vorhandenen wirksamen Medikamenten verglichen werden.

Eine weitere Strategie ist das Ghostwriting. Dabei wird ein Text von einem Auftragschreiber verfasst, als Autor offiziell aber eine andere Person angegeben. Die meisten von uns haben bestimmt schon einmal die Autobiographie einer berühmten Persönlichkeit in der Hand gehabt, die eindeutig aus der Feder eines solchen Ghostwriters stammte. Solche fremderstellten Texte tauchen aber auch in wissenschaftlichen Veröffentlichungen auf – und zwar mit potenziell beunruhigenden Folgen. Manchmal beauftragt die pharmazeutische Industrie ein Kommunikationsunternehmen mit der Erstellung von Artikeln, die das Produkt des betreffenden Herstellers erwartungsgemäß in einem günstigen Licht präsentieren. Nach Fertigstellung des Artikels wird ein Akademiker vertraglich, gegen «Honorar», verpflichtet, als dessen «Autor herzuhalten». Anschließend wird der Beitrag zur Publikation eingereicht. Ganz besonders beliebt ist in dieser Hinsicht die Zeitschriftenrubrik «Kommentare». Ein weiteres Zielobjekt der Industrie sind Zeitschriftensupplemente – das sind separate gebundene Sonderhefte, die zwar den Namen der Trägerzeitschrift führen, häufig aber von der Industrie gesponsert sind und meist einem weniger strengen Peer-Review-Verfahren unterliegen als die Trägerzeitschrift selbst.²⁴ Auf diese Weise generierte und geförderte Werbebotschaften führen dazu, dass die Vor-

Ärzte und die Pharmaindustrie

«Niemand kennt die genaue Summe, welche die Pharmaindustrie an Ärzte zahlt, aber aus den Jahresberichten der neun führenden US-amerikanischen Pharmaunternehmen würde ich schätzen, dass sie sich pro Jahr auf mehrere zehn Milliarden Dollar beläuft. Damit hat die pharmazeutische Industrie eine enorme Kontrolle darüber, wie ihre Produkte von Ärzten beurteilt und verordnet werden. Mit ihren weitreichenden Verbindungen zu Ärzten, vor allem zu leitenden Wissenschaftlern an angesehenen medizinischen Hochschulen, nehmen sie Einfluss auf die Ergebnisse der Forschung, auf die Art und Weise, wie Medizin praktiziert wird, und sogar auf das Verständnis dessen, was eine Krankheit ausmacht.»

Angell M. Drug companies & doctors: a story of corruption.
New York Review of Books, 15. Januar 2009.

züge der jeweiligen Produkte zu stark angepriesen und ihre Nachteile bagatellisiert werden (s. auch Kap. 8, S. 147).

Arzneimittelhersteller platzieren gern auch Werbeanzeigen für ihre Produkte in medizinischen Fachzeitschriften. Üblicherweise enthalten diese Anzeigen Hinweise auf wissenschaftliche Evidenzquellen, um die darin aufgestellten Behauptungen zu untermauern. Auf den ersten Blick mögen sie ja überzeugend sein, doch bei unabhängiger Prüfung der Evidenz ergibt sich ein ganz anderes Bild. Selbst wenn die Ergebnisse aus randomisierten Studien stammen – was die Leser dieser Anzeigen vermutlich als zuverlässige Bewertung einschätzen – so ist doch nicht alles so, wie es scheint. Denn als Wissenschaftler die Werbeanzeigen in führenden medizinischen Fachzeitschriften analysierten, um herauszufinden, ob die Erkenntnisse aus diesen randomisierten Studien plausibel waren, stellten sie Folgendes fest: Überhaupt nur 17% der zitierten Studien waren von hochwertiger Qualität, stützten die bezüglich des fraglichen Medikaments aufgestellten Behauptungen und waren nicht vom Pharmaunternehmen selbst gesponsert worden. Und wie man weiß, besteht eine erhöhte Wahrscheinlichkeit, dass derart gesponserte Forschungsarbeiten zu Ergebnissen gelangen, die für das Produkt des betreffenden Unternehmens günstig ausfallen.^{25, 26}

Zwielichtig, undurchsichtig und hinters Licht geführt?

In einem humoristischen Beitrag zur Weihnachtsausgabe des *British Medical Journal* stellten zwei Wissenschaftler ein erfundenes Unternehmen aus der zweitältesten Branche der Welt mit dem Namen «HARLOT plc» vor, das Studiensponsoren verschiedenste Dienstleistungen offerierte, beispielsweise:

«Garantiert positive Ergebnisse versprechen wir allen Herstellern von dubiosen Medikamenten und Medizinprodukten, die eine Erhöhung ihrer Marktanteile anstreben, allen ärztlichen Berufsgruppen, welche die Nachfrage nach überflüssigen diagnostischen und therapeutischen Leistungen mehrten möchten, sowie allen regionalen und nationalen Gesundheitsbehörden, welche die Umsetzung unvernünftiger und eigennütziger gesundheitspolitischer Maßnahmen umzusetzen bestrebt sind ... und bei zwielichtigen «Me too»-Medikamenten verschafft Ihnen unser Team «Me-Too-Prüfpläne leicht gemacht» garantiert positive Studienergebnisse.»

Zu Ihrer großen Verwunderung gingen bei den Autoren doch etliche anscheinend ernst gemeinte Anfragen zu diesem erstaunlichen Portfolio der Fa. HARLOT plc ein.

Sackett DL, Oxman AD. HARLOT plc: an amalgamation of the world's two oldest professions. *BMJ* 2003; 327: 1442–1445.

In renommierten medizinischen Fachzeitschriften wie *The Lancet*²⁷ wurde in verschiedenen Kommentaren auf die abnormen Anreize, von denen sich manch einer, der in der klinischen Forschung tätig ist, motiviert fühlt, sowie die zunehmend zweifelhaften Beziehungen zwischen Universitäten und der Industrie eingegangen. Eine ehemalige Herausgeberin des *New England Journal of Medicine* stellte die unverblünte Frage: «Ist die Hochschulmedizin käuflich?»²⁸

Kommerzielle Prioritäten stellen aber nicht die einzigen abnormen Einflüsse auf Verhaltensmuster in der biomedizinischen Forschung dar, welche die Interessen der Patienten missachten. Viele Menschen an Universitäten und in Einrichtungen zur Forschungsförderung gehen davon aus, dass sich Verbesserungen der Gesundheit sehr wahrscheinlich auf Versuche zur Entschlüsselung der grundlegenden Mechanismen von Krankheit zurückführen lassen. Deshalb führen sie ihre Untersuchungen in Laboratorien und an Tieren durch. Diese Art von Grundlagenforschung ist unbestreitbar nötig. Allerdings gibt es nur herzlich wenig Belege, mit denen sich rechtfertigen ließe, warum auf sie ein erheblich größerer Anteil an Fördermitteln entfällt als auf die unter Mitwirkung von Patienten durchgeführte Forschung.^{29, 30} Trotzdem haben wir es mit einer riesigen Flut von Laborexperimenten zu tun, deren Relevanz für den Patienten nicht hinreichend evaluiert wird.

Man braucht nur das richtige Gen zu finden

«Es steht ... zu hoffen, dass die genetische Revolution eines Tages alle Probleme des Menschen lösen wird. Wir werden imstande sein, diejenigen Gene zu lokalisieren und zu replizieren, die uns dazu prädisponieren, bessere Häuser zu bauen, der Umweltverschmutzung Herr zu werden, Krebserkrankungen tapferer zu ertragen, die Mittel für Kinderbetreuungseinrichtungen bereitzustellen, die für jedermann zugänglich sind, und uns über Standort und Bauweise eines nationalen Sportstadions einig zu werden. Bald wird jeder Säugling unter genetisch ausgeglichenen Bedingungen geboren. Das Gen, das – sagen wir einmal – bewirkt, dass Mädchen im Abitur besser abschneiden als Jungen, wird identifiziert und entfernt. Es gibt endlose genetische Möglichkeiten. ... Also ja, wir betreten eine unsichere Welt, aber eine, die in gewisser Weise auch Hoffnung macht. Denn welche schwerwiegenden moralischen Dilemmata die genetischen Fragen auch aufwerfen mögen, eines Tages wird es gelingen, das Gen zu isolieren, das sie lösen wird.»

Iannucci A. *The Audacity of Hype*. London: Little, Brown, 2009, S. 270–271.

Ein Grund für diese Ungleichheit ist der Rummel um die klinischen Fortschritte, die man sich aus der Grundlagenforschung, insbesondere der Genetik, erhofft (zu genetischen Untersuchungen s. Kap. 4, S. 81 f.). Sir David Weatherall, ein bedeutender Mediziner und Genetiker, drückte dies 2011 wie folgt aus:

Unsere häufigsten Todesursachen spiegeln den Einfluss einer großen Anzahl von Genen mit jeweils kleinen Effekten in Verbindung mit wichtigen Beiträgen physikalischer und sozialer Einflussfaktoren wider. Diese Arbeiten liefern wertvolle Informationen über etliche Krankheitsprozesse, unterstreichen aber auch die Individualität und Variabilität der zugrunde liegenden Krankheitsmechanismen. Das Zeitalter einer personalisierten Medizin, die auf unserer genetischen Ausstattung beruht, liegt mit Sicherheit noch in weiter Ferne.³¹

Heute, 50 Jahre nach der Entdeckung der DNA-Struktur, scheint diese Kakophonie von Behauptungen über die schon bald zu erwartenden Vorteile der «genetischen Revolution» für die Gesundheitsversorgung allmählich zu verklingen. Anscheinend bricht langsam die Wirklichkeit über uns herein. In einem Vortrag über das Potenzial der Genetik für die Entwicklung neuer Medikamente kommentierte ein Wissenschaftler dies wie folgt:

Wir sind im Zeitalter des Realismus angekommen. ... Genetische Aspekte müssen im Zusammenhang mit anderen Faktoren wie z. B. der Umwelt und dem klinischen Einsatz von Medikamenten gesehen werden. Nur weil ein Medikament bei einem Patienten nicht wirkt, bedeutet dies nicht, dass die Ursache dafür in einer genetischen Variation des Therapieansprechens zu suchen ist.³²

Im Editorial einer Ausgabe der Wissenschaftszeitschrift *Nature* zur Feier des zehnten Jahrestages der Entschlüsselung des menschlichen Genoms hieß es dazu:

... gewisse Fortschritte hat es in Form von Medikamenten gegeben, die gegen spezifische genetische Defekte gerichtet sind, wie sie z. B. bei einigen Krebsarten und auch bei einigen seltenen erblichen Krankheiten

identifiziert wurden. Doch die Komplexität der postgenomischen Biologie hat die anfänglichen Hoffnungen, dass aus dem Rinnsal an Therapien eine wahre Flut werden könnte, zunichte gemacht.³³

Psoriasis-Patienten – Stiefkinder der Forschung

«Nur wenige Studien beinhalteten einen Vergleich der verschiedenen Behandlungsoptionen oder untersuchten Fragen zur Langzeittherapie. Angesichts einer Krankheit mit potenziell lebenslangem (chronischem) Verlauf sind auch die kurzen Studiendauern nicht überzeugend. Sicher wissen wir anscheinend nur, dass unsere Therapien besser sind als gar nichts. Bezeichnenderweise haben die Wissenschaftler die Erfahrungen, Auffassungen, Präferenzen oder die Zufriedenheit der Patienten vollkommen außer Acht gelassen.»

R. Jobling, Vorsitzender der *Psoriasis Association*

Jobling R. Therapeutic research into psoriasis: patients' perspectives, priorities and interests. In: Rawlins M, Littlejohns P, eds. *Delivering quality in the NHS 2005*. Abingdon: Radcliffe Publishing Ltd, S. 53–56.

Wenn wir verantwortungsvoll handeln wollen, dann führt einfach kein Weg daran vorbei, dass wir gut geplante Studien unter der Mitwirkung von Patienten brauchen, um die therapeutischen Konzepte zu prüfen, die sich aus der Grundlagenforschung ergeben. Und nur allzu oft werden solche Theorien nie bis zum Ende verfolgt, um herauszufinden, ob sie für die Patienten auch relevant sind. Mehr als zwei Jahrzehnte, nachdem Wissenschaftler den Gendefekt entdeckt haben, der die Mukoviszidose (zystische Fibrose) verursacht, stellen Mukoviszidose-Kranke noch immer eine entscheidende Frage: «Wann endlich wird sich diese Entdeckung für unsere Gesundheit bezahlt machen?»

Aber auch dann, wenn die Forschung patientenrelevant zu sein scheint, sehen Wissenschaftler bei der Planung ihrer Studien noch oft genug über die Interessen der Patienten hinweg. Sehr aufschlussreich ist ein Fall, in dem Lungenonkologen aufgefordert wurden, sich in die Lage der Patienten zu versetzen und zu überlegen, ob sie sich zur Teilnahme an den sechs Lungenkrebsstudien bereit erklären würden, für die sie als Patienten vielleicht infrage kämen. Zwischen 36 und 89 % antworteten darauf, dass sie selber *nicht* an solchen Studien teilnehmen würden.³⁴

Auch in klinischen Studien über Psoriasis – eine chronische und belastende Hautkrankheit, von der weltweit ca. 125 Millionen Men-

schen betroffen sind – finden die Patienteninteressen kaum Berücksichtigung.^{35, 36} So hat beispielsweise die *Psoriasis Association* in Großbritannien herausgefunden, dass Wissenschaftler in vielen Studien zur Beurteilung der Effekte der verschiedenen Therapien auch weiterhin ein weitgehend in Verruf geratenes Bewertungssystem anwenden. Zu den Schwachstellen dieses Systems gehört, dass im Mittelpunkt Zielgrößen wie die Gesamtfläche der betroffenen Haut und die Dicke der Hautveränderungen stehen, während die Patienten, was kaum verwunderlich sein dürfte, vor allem unter Problemstellen im Gesicht, an den Handflächen und Fußsohlen sowie im Genitalbereich leiden.³⁷

- Unnötige Forschung bedeutet Verschwendung von Zeit, Arbeit, Geld und anderen Ressourcen; zudem ist sie ethisch nicht vertretbar und für den Patienten potenziell schädlich.
- Neue Forschungsvorhaben sollten nur in Angriff genommen werden, wenn eine aktualisierte Übersicht über die bisherige Forschung den Bedarf dafür nachweist und nachdem ihre Eintragung in ein Register erfolgt ist.
- Die wissenschaftlichen Daten aus neuen Forschungsarbeiten sollten dazu benutzt werden, um die bisherigen Übersichtsarbeiten über alle relevanten Erkenntnisse zu aktualisieren.
- Ein Großteil der Forschung ist von minderwertiger Qualität und wird aus fragwürdigen Gründen betrieben.
- Die Forschungsagenda steht unter widersinnigen Einflüssen seitens der Industrie wie auch der Wissenschaft.
- Für den Patienten relevante Fragen werden oft gar nicht thematisiert.

11

Richtige Forschung geht uns alle an

Wie wir im vorigen Kapitel dargelegt haben, wird im Rahmen schlechter oder überflüssiger Forschung – welche die Fragen, auf die es den Patienten ankommt, weder jetzt noch in Zukunft beantwortet – viel Zeit, Geld und Arbeit verschwendet. Wir haben Sie hoffentlich davon überzeugen können, dass aus der produktiven Partnerschaft zwischen Patienten, Ärzten, Öffentlichkeit und Wissenschaft in Zukunft auch bessere Therapietests resultieren sollten.

Welchen Beitrag können Patienten und Öffentlichkeit zur Verbesserung der Forschung leisten?

Die einstmals geschlossene Welt der Medizin ist dabei, ihre Pforten immer weiter zu öffnen, um neuen Ideen und früheren «Außenseitern» Zutritt zu gewähren. Der Paternalismus ist im Schwinden begriffen. Infolgedessen tragen nun in zunehmendem Maße die Patienten und die Öffentlichkeit zur Gesundheitsforschung bei – und zwar im Hinblick darauf, was erforscht wird und wie Studien durchgeführt werden.¹ Weltweit wird die Zusammenarbeit mit Patienten als Partnern im Forschungsprozess zunehmend befürwortet, und inzwischen stehen für Ärzte, die Patienten und Öffentlichkeit in ihre Arbeit einbeziehen möchten, auch entsprechende Orientierungshilfen zur Verfügung.^{2, 3, 4}

Patienten verfügen über Erfahrungen, die Abwägungen und Entscheidungen erleichtern und neue Einsichten eröffnen können. Ihr am eigenen Leib erfahrenes Wissen ermöglicht wertvolle Einblicke in die Art und Weise, wie Menschen auf Krankheiten reagieren und

welchen Einfluss dies auf ihre Therapieentscheidungen hat. Wie die inzwischen angesammelte Evidenz aus Umfragen⁵, systematischen Übersichtsarbeiten über Forschungsberichte¹, Berichten zu einzelnen Studien⁶ und Folgenabschätzungen⁷ zeigt, kann die Einbindung von Patienten und Öffentlichkeit zur Verbesserung von Therapiestudien beitragen.

Die Entscheidung des Patienten: David und Goliath

«Wer hat die Macht, dafür zu sorgen, dass sich die Forschung tatsächlich mit den Fragen befasst, die für die Patienten in all ihrer Not und Vielfalt am dringlichsten sind? Warum werden nicht die Fragen gestellt, die die größte Relevanz haben? Wer stellt derzeit die Fragen, und wer sollte sie eigentlich stellen? Wer sollte die Priorisierung kontrollieren? Patienten sind am ehesten in der Lage, die Gesundheitsthemen zu benennen, die für sie am wichtigsten sind und die Einfluss auf ihr Wohlbefinden, ihre Versorgung, ihre Lebensqualität und Lebenserwartung haben. Die Patienten sind wie David: Sie bestücken ihre Schleudern gegen die Goliaths der Pharmaunternehmen, welche die Evidenz brauchen, um ihre Waren zu vermarkten und Gewinne zu erzielen, und die Studienärzte, die getrieben sind von Neugier, der Notwendigkeit, sich Forschungsgelder zu sichern, dem Wunsch nach fachlicher Anerkennung und beruflicher Weiterentwicklung. Profite, wissenschaftlicher Forschungsdrang, Fördermittel und wissenschaftliche Publikationen sind nur dann akzeptabel, wenn ihre hauptsächliche Motivation das Wohl des Patienten ist. Unabhängige Patienten und Organisationen, die sich für hochwertige Forschung einsetzen, sollten ihre Schleudern herausholen, ihre Munition sorgfältig auswählen, genau zielen und ihre Gegner bezwingen.»

Refractor. Patients' choice: David and Goliath. *Lancet* 2001; 358: 768.

Zu den zahlreichen Initiativen in dieser Richtung gehört die *Cochrane Collaboration* (www.cochrane.org). Der Name steht für ein internationales Netzwerk von Wissenschaftlern, welche die beste verfügbare Evidenz zu medizinischen Behandlungen systematisch zusammenfassen und die Erfahrungen von Patienten schon seit ihren Anfängen im Jahr 1993 berücksichtigen. Die 2004 gegründete *James Lind Alliance* (www.lindalliance.org) bringt Patienten, Pflegekräfte und Ärzte an einen Tisch, um die nach ihrer übereinstimmenden Meinung wichtigsten unbeantworteten Fragen über die Effekte medizinischer Therapien zu identifizieren und nach ihrer Wichtigkeit zu ordnen (d.h. zu priorisieren). Diese Informationen über Unsicherheiten in der Therapie tragen dazu bei, dass die Förderer von Gesundheitsforschung wissen, was für Patienten und Ärzte am wichtigsten ist.⁸ Seit 2008 finanziert die Europäische Kom-

mission ein Projekt, mit dem die Rolle von Patientenorganisationen in klinischen Studien gestärkt werden soll. Ziel ist die Bündelung von Erfahrungen innerhalb der europäischen Länder durch Workshops, Berichte und andere Formen des Austauschs.⁹ Auch in anderen Ländern ist die Öffentlichkeit im Allgemeinen aktiv in Forschungsvorhaben eingebunden.

Eine wichtige Partnerschaft

«Forschung im *National Health Service*, in deren Mittelpunkt der Mensch steht, ist ohne die Beteiligung der Patienten und der Öffentlichkeit undenkbar. Ganz gleich, wie kompliziert die Forschung oder wie brillant der Forscher auch sein mag: Stets können Patienten und Öffentlichkeit ganz besonders wertvolle Erkenntnisse liefern. Bei der Planung, Umsetzung und Bewertung von Forschungsvorhaben sorgen ihre Anregungen dafür, dass Studien effizienter, glaubwürdiger und oftmals auch kosteneffektiver werden.»

Professor Dame Sally Davies. Vorwort zu Staley K.
Exploring impact: public involvement in NHS, public health and social care research.
Eastleigh: INVOLVE, 2009. Zugänglich unter www.invo.org.uk.

Diese Rollen entwickeln sich auf vielfältige Weise immer weiter¹⁰ und ermöglichen so den Patienten und der Öffentlichkeit die Zusammenarbeit mit Ärzten und anderen Gesundheitsberufen, und es werden dafür ständig neue Methoden entwickelt (s. u. den Abschnitt «Die Kluft zwischen Patienten und Wissenschaftlern überwinden» sowie Kap. 13, Punkt 2: Richtige Planung und Durchführung von Forschungsvorhaben).¹¹ Dies gilt für das gesamte Spektrum der Forschungsaktivitäten:

- Formulierung der zu untersuchenden Fragen
- Planung von Projekten, darunter auch die Auswahl der wichtigsten Zielgrößen
- Projektmanagement
- Erstellung von Informationsbroschüren für Patienten
- Analyse und Interpretation der Ergebnisse
- Verbreitung und Umsetzung der Erkenntnisse zur Unterstützung von Behandlungsentscheidungen.

Mitwirkung von Patienten an der Forschung

Wie ist es überhaupt zur Mitwirkung von Patienten an der Forschung gekommen? In Kapitel 3 haben wir beispielsweise dargestellt, dass die Therapieexzesse, die Brustkrebspatientinnen früher zugemutet wurden, sowohl aufseiten einer neuen Art von «Arzt-Wissenschaftler» als auch aufseiten der Patientinnen zu neuen Denkanstößen und Veränderungen geführt haben. Ärzte und Patienten arbeiteten zusammen, um wissenschaftliche Belege beizubringen, die einerseits strengen wissenschaftlichen Standards genügten und andererseits den Bedürfnissen der Frauen Rechnung trugen. Als Frauen die Praxis der radikalen Mastektomie infrage zu stellen begannen, gaben sie damit zu verstehen, dass es ihnen um mehr als die Bekämpfung ihrer Krebserkrankung ging: Sie verlangten auch ein Mitspracherecht bei den Strategien, mit denen wirksame Wege zur Bewältigung der Krankheit gefunden werden können.

Für diejenigen Patienten und Mitglieder der Öffentlichkeit, die in vollem Umfang als «Co-Wissenschaftler» mitwirken wollen, bieten sich verschiedene Möglichkeiten an. Sie können sich individuell oder als Mitglied einer Selbsthilfegruppe einbringen, oder sie können an einer moderierten Gruppenaktivität wie z.B. einer Fokusgruppe teilnehmen. Unabhängig von der Form ihrer Mitwirkung kann es mit Sicherheit nicht schaden, wenn sie sich mit den praktischen Grundlagen der Forschungsmethodik vertraut machen, damit sie sich selbstbewusst und effektiv in die Partnerschaft mit Ärzten und anderen Berufsgruppen im Gesundheitswesen einbringen können. Und dafür brauchen sie hochwertige Informationen und Schulungen, die auf ihre Rolle zugeschnitten sind. Wir werden in Kapitel 12 erklären, warum es für das richtige Verständnis so überaus wichtig ist, wie diese Informationen – insbesondere, was die Statistik betrifft – dargeboten werden. Zudem gibt es noch zahlreiche andere, weniger herausragende Möglichkeiten, wie Patienten und Öffentlichkeit einen Beitrag zu den Forschungsanstrengungen leisten können – vor allem dann, wenn es gelingt, eine Kultur der Zusammenarbeit zu entwickeln, welche die Erkenntnisse und Beobachtungen aus Sicht des Patienten anerkennt.

Der aktive «Patient-Wissenschaftler» von heute kann dankbar auf die wegweisenden Leistungen der ersten «Patientenpioniere» zurückblicken, die erkannt haben, dass sie für sich eintreten und den Status quo infrage stellen müssen – und dass sie, um dies tun zu können, präzise Informationen benötigten. So machte sich beispielsweise in den USA Anfang der 1970er-Jahre eine kleine Gruppe von Brustkrebs-Patientinnen unter der Leitung von Rose Kushner daran, sich weiterzubilden, um wirksam eingreifen zu können. Anschließend führten sie Schulungen für andere Frauen durch. Kushner war Brustkrebs-Patientin und freie Autorin, und sie stellte Anfang der 1970er-Jahre das traditionell autoritäre Arzt-Patient-Verhältnis und die Notwendigkeit der radikalen Brustoperation infrage.¹² Auf der Grundlage ihrer gründlichen Bewertung der Erkenntnisse über die Wirkungen der radikalen Mastektomie verfasste sie ein Buch. Gegen Ende des Jahrzehnts hatte sie einen solchen Einfluss und eine solche Anerkennung errungen, dass sie in Zusammenarbeit mit dem US-amerikanischen *National Cancer Institute* Vorschläge für neue Forschungsvorhaben prüfte.¹³ Ähnlich veranlasste auch in Großbritannien der Informationsmangel Frauen dazu, das Heft selbst in die Hand zu nehmen. So rief etwa Betty Westgate in den 1970er-Jahren die *Mastectomy Association* ins Leben, und in den 1980er-Jahren gründete Vicky Clement-Jones die Wohltätigkeitsorganisation CancerBACKUP (mittlerweile Teil von *Macmillan Cancer Support*).

Das Thema AIDS mithilfe von Laien neu überdenken

«Glaubwürdigkeitskämpfe in der AIDS-Arena finden an mehreren Fronten statt: Es sind ungewöhnlich viele Parteien beteiligt. Und dass sich Laien in die Verkündung und Beurteilung wissenschaftlicher Behauptungen einmischen, hat das, was wir über AIDS zu wissen glauben, mitgeprägt – genauso, wie es auch dazu beigetragen hat, unser Verständnis darüber, wer als «Lai» und wer als «Experte» anzusehen ist, infrage zu stellen. Stets geht es dabei um die Glaubwürdigkeit der jeweiligen Wissensansprüche oder Wortführer. Im Kern geht es jedoch darum, wie eigentlich Glaubwürdigkeit festgestellt wird: Wie wird über wissenschaftliche Behauptungen entschieden, und wer trifft diese Entscheidungen? [Wie diese Studie zeigt,] sind Debatten innerhalb der Wissenschaft gleichzeitig auch Debatten über die Wissenschaft und wie sie betrieben werden sollte – bzw. wer sie betreiben sollte.»

Epstein S. *Impure science: AIDS, activism and the politics of knowledge*.
London: University of California Press, 1996.

Ende der 1980er-Jahre hatten mit HIV infizierte oder an AIDS erkrankte Menschen in den USA erstaunlich viel Wissen über ihre Krankheit angehäuft. Politisch waren sie darauf ausgerichtet, ihre Interessen gegen das Establishment zu verteidigen, wodurch sie Patienten den Weg zur Mitbestimmung über das Design von Studien ebneten. Diese Mitsprache führte schließlich dazu, dass den Patienten in den Studien verschiedene Behandlungsoptionen und flexible Studiendesigns angeboten wurden, um sie zur Teilnahme an Studien zu ermutigen. Diesem Beispiel folgte man Anfang der 1990er-Jahre in Großbritannien, als man eine AIDS-Patientengruppe an Studien beteiligte, die am *Chelsea and Westminster Hospital* in London durchgeführt wurden: Die Patienten halfen bei der Auswahl des richtigen Studiendesigns.¹⁴

Diese AIDS-Aktivistinnen ließen die Studienleiter aufhorchen: Was manche Wissenschaftler als ein durch organisierte Patientengruppen verursachtes Chaos ansahen, war genau genommen die legitime Infragestellung der Interpretation von Unsicherheit durch die Wissenschaftler. Bis dahin hatten die Wissenschaftler bei ihrem Vorgehen die von Patienten bevorzugten Studienzielgrößen schlicht übersehen. Andererseits lernten die Patienten einzuschätzen, welche Gefahren damit verbunden sind, wenn man vorschnell über die Wirkungen neuer Medikamente urteilt und den Forderungen nach Zulassung eines «vielversprechenden» neuen AIDS-Medikaments nachgibt, bevor es einer strengen Prüfung unterzogen wurde. Die Wissenschaftler mögen eingewendet haben, dass eine solche «mitfühlende Zulassung» (engl. *compassionate release*) neuer Medikamente die Qual der Ungewissheit für gegenwärtige und zukünftige Patienten nur verlängere. Die Patienten hielten jedoch dagegen, dass dadurch sowohl bei Patienten als auch Wissenschaftlern letztendlich die Einsicht beschleunigt wurde, dass wir besonnene, kontrollierte Therapiebewertungen brauchen, die gemeinsam geplant werden und die die Bedürfnisse beider Seiten berücksichtigen.¹⁵

In den 1990er-Jahren hat besonders eine AIDS-Studie sehr eindrücklich gezeigt, wie wichtig es ist, Patienten an der Forschung zu beteiligen. Sie fiel in die Zeit, als das Medikament Zidovudin gerade erst zur Behandlung von AIDS zugelassen worden war. Es gab fundierte Belege für eine positive Wirkung bei Patienten mit fortge-

schrittener Erkrankung. Die naheliegende Frage lautete nun, ob sich die Krankheitsprogression hinauszögern und die Überlebensdauer verlängern ließe, wenn Zidovudin bereits in einem früheren Infektionsstadium verabreicht würde. Um diese Möglichkeit zu prüfen, wurden sowohl in den USA als auch in Europa Studien auf den Weg gebracht. Die US-Studie wurde vorzeitig abgebrochen, als sich ein möglicher, aber noch ungewisser positiver Effekt ergab. Trotz der in den USA erzielten Ergebnisse wurde die europäische Studie unter aktiver Beteiligung und mit Zustimmung der Patientenvertreter bis zum Erreichen eines eindeutigen Endpunkts fortgeführt. Ihre Schlussfolgerungen waren grundverschieden: Die frühzeitige Gabe von Zidovudin im Infektionsverlauf brachte anscheinend keinerlei Nutzen. Die unter diesen Umständen einzigen eindeutigen Effekte des Medikaments waren seine unerwünschten Nebenwirkungen.¹⁶

Wie Patienten faire Therapiestudien gefährden können

Die Beteiligung von Patienten an der Forschung wirkt sich auf faire Therapiestudien nicht immer nur förderlich aus. Eine im Jahr 2001 unter Wissenschaftlern durchgeführte Umfrage zeigte einige ausgesprochen positive Erfahrungen mit der Einbeziehung von Patienten in klinische Studien auf. Gleichzeitig deckte sie aber auch einige sehr reale Probleme auf, die hauptsächlich darauf zurückzuführen waren, dass die Beteiligten keine Erfahrungen mit dieser Art von Zusammenarbeit hatten. Zunächst einmal gab es häufig beträchtliche Verzögerungen beim Beginn solcher Forschungsprojekte. Zudem bestanden Bedenken wegen Interessenkonflikten und der «Repräsentativität» einiger Patienten, die noch nicht erkannt hatten, wie wichtig es ist, in die im Rahmen des Studienmanagements stattfindenden Sitzungen nicht nur ihre eigenen Interessen einzubringen.⁵

Viele dieser Probleme ergaben sich anscheinend aus dem aufseiten der Patienten verständlicherweise bestehenden Mangel an Wissen über den Ablauf und die Finanzierung von Forschungsvorhaben. Verzweifelte Umstände führen manchmal zu dem verzweifelten Bemühen, sich Zugang zu Behandlungen zu verschaffen, die noch nicht ausreichend geprüft wurden und die – selbst bei Patienten, die

im Sterben liegen – möglicherweise mehr schaden als nützen. Wir haben bereits an anderer Stelle darauf hingewiesen, dass diese Art von Lobbyismus durch Patienten und ihre Fürsprecher im Hinblick auf die «mitfühlende» Zulassung «vielversprechender» neuer medikamentöser Therapien für AIDS auch seine Schattenseiten hat: Die Identifizierung von Therapien, die auf patientenrelevante Ergebnisse ausgerichtet sind, wurde dadurch verzögert; so hat sich z. B. in der jüngeren Vergangenheit die Fürsprache durch falsch informierte Einzelpersonen bzw. Patientengruppen kontraproduktiv auf die Verordnung von Medikamenten gegen multiple Sklerose (MS) und Brustkrebs ausgewirkt.

Mitte der 1990er-Jahre wurden Interferone zur Behandlung von Patienten mit schubförmig remittierender MS auf der Grundlage spärlicher Wirksamkeitsnachweise zugelassen. Sehr schnell verlangten alle Patienten mit allen möglichen MS-Formen nach diesen kostspieligen Medikamenten, und ihre Anwendung wurde von den Krankenkassen sogar finanziert. Interferone wurden zu einer anerkannten Standardtherapie für diese kräftezehrende Krankheit. Infolgedessen werden wir niemals erfahren, wie man Interferone bei multipler Sklerose richtig verordnet – darüber wurde nie geforscht, und mittlerweile ist es zu spät, um die Uhr zurückzudrehen. Im Laufe der Zeit ist eines aber ganz deutlich geworden: Interferone haben unangenehme Nebenwirkungen wie z. B. «grippeähnliche» Symptome.

Auch Herceptin (Trastuzumab) ist, wie wir in Kapitel 1, S. 42–44, erläutert haben, kein Wundermedikament, das allen Brustkrebs-Patientinnen hilft. Zunächst einmal hängt seine Wirksamkeit von einer besonderen genetischen Konstellation des Tumors ab, die nur bei einer von fünf Frauen mit Brustkrebs vorliegt. Zudem hat das Medikament potenziell schwerwiegende kardiale Nebenwirkungen. Und doch wurden Politiker von Patientenvertretern durch Schüren des Medienrummels dazu gebracht, mit dem Strom der öffentlichen Meinung zu schwimmen: Die Anwendung von Herceptin wurde offiziell empfohlen, ohne sich groß um die vorhandene Evidenz bzw. die Erkenntnis zu kümmern, dass weitere Belege für ein ausgewogenes Verhältnis von Vor- und Nachteilen noch ausstanden.

«Pester Power» und neue Medikamente

«Bei neuen Medikamenten handelt es sich naturgemäß um unausgereifte Produkte, da die vollständigen Angaben über ihre Sicherheit, Wirksamkeit und Kosten auswirkungen noch nicht vorliegen. Dabei sollte auch bedacht werden, dass die enthusiastische Fürsprache für alles «Neue» nicht die alleinige Domäne der Presse ist, sondern oftmals auch in den anderen Medien sowie in der Medizin und den Naturwissenschaften zu beobachten ist. «Pester Power» ist ein Konzept, das man normalerweise mit kinderorientierten Werbestrategien in Zusammenhang bringt. In diesem Zusammenhang müssen wir uns die Frage stellen, ob wir es hier mit der *Pester Power* von Patienten zu tun haben oder etwa mit direkt an den Patienten als Endverbraucher gerichteter Werbung (sogenannte Direct-to-Consumer- oder DtC-Werbung), durch die das Bewusstsein für neue Produkte geweckt wird und Patienten, Wohltätigkeitsorganisationen und tatsächlich auch Ärzte dann verlangen, dass diese Produkte zugänglich gemacht werden. Wenn dies zutrifft, dann müssen wir mehr über diejenigen in Erfahrung bringen, die hinter dieser Art von Marketing stehen, welche Auswirkungen dieses Marketing auf das Arzt- und das Konsumentenverhalten hat und ob es im Rahmen der geltenden Zulassungsregeln überhaupt statthaft ist.»

Wilson PM, Booth AM, Eastwood A et al. Deconstructing media coverage of trastuzumab (Herceptin): an analysis of national newspaper coverage. *Journal of the Royal Society of Medicine* 2008; 101: 125–32

Sind Patientenverbände unabhängig?

Ein weiterer, weniger bekannter Interessenkonflikt herrscht im Verhältnis zwischen Patientenorganisationen und der pharmazeutischen Industrie. Die meisten Patientenverbände verfügen nur über knappe Finanzmittel, sind auf die Arbeit Freiwilliger angewiesen und beziehen nur geringe unabhängige Fördermittel. Zuwendungen von und gemeinsame Projekte mit pharmazeutischen Unternehmen können ihnen zu Wachstum und mehr Einfluss verhelfen, aber auch die Ziele der Patienten, darunter auch ihrer Forschungsvorstellungen, verzerren und falsch darstellen. Die Größenordnung dieses Problems lässt sich nur schwer ermessen, doch einen faszinierenden Einblick gewährt eine Umfrage, die durchgeführt wurde, um festzustellen, in welchem Maße Patienten- und Verbraucherorganisationen, die mit der Europäischen Arzneimittelbehörde EMA (*European Medicines Agency*) zusammenarbeiten, durch Firmensponsoren gefördert werden. Die EMA koordiniert die Bewertung und Überwachung neuer Medikamente in ganz Europa und lässt Patienten- und

Verbrauchergruppen erfreulicherweise aktiv an ihren Zulassungsaktivitäten teilhaben. Als allerdings zwischen 2006 und 2008 eine Umfrage unter 23 dieser Gruppen durchgeführt wurde, ergab sich dabei, dass 15 teilweise oder sogar in erheblichem Umfang von Arzneimittelherstellern oder Verbänden der pharmazeutischen Industrie gefördert wurden. Darüber hinaus konnte weniger als die Hälfte dieser Gruppen gegenüber der Behörde genaue Angaben über Herkunft oder Höhe der von ihnen bezogenen Fördermittel machen.¹⁷

Die Bürger an der Verbesserung der Gesundheitsversorgung beteiligen

«Die Verschmelzung der Interessen von Patientenvertretungen, denjenigen, welche medizinische Therapien verkaufen, und denen, die sie verordnen, sorgt für eine starke Mixtur aus verschiedenen Einflüssen, welche die Entscheidungsträger fast immer in eine Richtung drängt: noch mehr Tests, noch mehr Eingriffe, noch mehr Betten, noch mehr Tabletten ...

Als jemand, der sich seit mehr als zehn Jahren mit diesem Thema befasst, habe ich das Gefühl, dass in dieser Debatte häufig eine Stimme fehlt, nämlich diejenige, die das öffentliche Interesse nachhaltig vertritt. Gesponserte Patientenvertretungen sind schnell bereit, eine neue Therapie oder Technologie in den Himmel zu heben, zögern aber, ihre begrenzte Wirksamkeit, ihre übermäßigen Kosten oder offenkundigen Gefahren öffentlich zu kritisieren. Und wie viele Journalisten neigen auch Politiker dazu, sich unnötigerweise durch führende Ärzte und leidenschaftliche Fürsprecher einschüchtern zu lassen, die sich nur allzu oft für Marketingkampagnen einspannen lassen, in denen Krankheitsdefinitionen erweitert und die teuersten Lösungen beworben werden.

Vielleicht kann die Gründung neuer Bürgerlobbys im Gesundheitswesen, die sich damit auskennen, wie man wissenschaftliche Evidenz nutzen und missbrauchen kann, bewirken, dass die Debatte über die Ausgabenpriorisierung auf eine fundiertere Grundlage gestellt wird. Solche Bürgergruppen könnten es sich zur Aufgabe machen, irreführendes Marketing in den Medien routinemäßig aufzudecken und sowohl Öffentlichkeit als auch Entscheidungsträgern realistische und differenzierte Bewertungen der Risiken, Vorteile und Kosten einer sehr viel breiteren Palette von Gesundheitsstrategien zugänglich zu machen.»

Moynihan R. Power to the people. *BMJ* 2011; 342: d2002.

In einigen Fällen haben Pharmaunternehmen selbst Patientenorganisationen gegründet, um Einfluss zugunsten ihrer Produkte zu nehmen. So hat eines der Unternehmen, die Interferon herstellen, eine neue Patientengruppe mit dem Namen «Action for Access» ins Leben gerufen, um den britischen *National Health Service* dazu zu bewegen, Interferone zur Behandlung der multiplen Sklerose verfügbar zu machen (siehe oben).^{18, 19} Die Botschaft, die diese Patien-

tengruppen mit ihrer beträchtlichen Öffentlichkeitswirkung verkündeten, lautete, dass Interferone zwar wirksam, aber zu teuer seien. Dabei ging es in Wahrheit um die Frage, ob diese Medikamente überhaupt irgendwelche vorteilhaften Wirkungen haben.

Zusammenarbeit von Patienten und Wissenschaftlern

Die Kluft überbrücken

Wir haben weiter oben auf die Probleme aufmerksam gemacht, die aus der Mitwirkung von Patienten an Therapietests erwachsen können und wie Patienten faire Therapietests unabsichtlich torpedieren können. Wie bei den meisten Dingen gilt auch hier: Gute Absichten sind keine Gewähr dafür, dass mehr Nutzen als Schaden entsteht. Dennoch finden sich auch eindeutige Beispiele für die Vorteile, die sich aus der Zusammenarbeit von Wissenschaftlern und Patienten mit dem Ziel einer Steigerung der Forschungsrelevanz und Forschungsplanung ergeben. Deshalb suchen viele Wissenschaftler aktiv nach Patienten, mit denen sie zusammenarbeiten können.

Wie wertvoll eine vorbereitende Zusammenarbeit sein kann, zeigt folgendes Beispiel, in dem Wissenschaftler gemeinsam mit Patienten und potenziellen Patienten einige schwierige Fragen erörterten, die bei der Prüfung von Notfalltherapien auftreten. Damit Therapien bei einem akuten Schlaganfall Erfolg haben können, müssen sie möglichst rasch nach Auftreten des Schlaganfalls eingeleitet werden. Da die Wissenschaftler sich unsicher waren, wie man am besten verfährt, baten sie Patienten und ihre Betreuungspersonen um Unterstützung. Sie hielten ein Sondierungstreffen mit einer Gruppe von Patienten und Ärzten ab und führten Fokusgruppen durch, an denen ältere Menschen teilnahmen. Dadurch gelang es, klarere Studienpläne auszuarbeiten, und die Patienten unterstützten die Wissenschaftler beim Entwurf und bei der Überarbeitung der Aufklärungsbroschüre für diese Studie.²⁰

Diese sorgfältigen Forschungsvorarbeiten mündeten in Pläne für die Durchführung einer randomisierten Studie, die von der zuständigen Ethik-Kommission ohne Umschweife genehmigt wurde. Die Teilnehmer der Fokusgruppe hatten das ethische Dilemma erkannt,

das mit dem Versuch einhergeht, eine Einwilligungserklärung (engl. *informed consent*) von einer akut erkrankten Person zu erhalten, die unter Umständen verwirrt oder sprachunfähig, wenn nicht gar bewusstlos ist. So konnten sie Lösungen vorschlagen, die zu einem für alle Beteiligten akzeptablen Studiendesign und zu bedeutsamen Verbesserungen in der Informationsbroschüre führten.

Immer häufiger finden sich unter den Mitgliedern von Forschungsteams auch Sozialwissenschaftler, welche die sensiblen Aspekte von Krankheit gemeinsam mit Patienten formal untersuchen und die Durchführung von Studien auf diese Weise verbessern. In einer klinischen Studie bei Männern mit lokalisiertem Prostatakrebs wollten Wissenschaftler drei sehr unterschiedliche Behandlungsmöglichkeiten miteinander vergleichen: Operation, Strahlentherapie und das «beobachtende Abwarten» (engl. *watchful waiting*). Das führte zu Problemen, und zwar sowohl aufseiten der Ärzte, die den Patienten die Studienteilnahme vorschlagen sollten, als auch aufseiten der Patienten, die eine Entscheidung über ihre Teilnahme an dieser Studie treffen mussten. Bei den Ärzten stieß die Beschreibung der Option «beobachtendes Abwarten» auf so wenig Gegenliebe, dass sie sie bis zuletzt aufhoben. Zudem erläuterten sie diese Möglichkeit, ohne dabei allzu zuversichtlich zu wirken, weil sie fälschlicherweise davon ausgingen, dass diese Option für die Männer, denen die Teilnahme an der Studie angeboten werden sollte, inakzeptabel wäre. Deshalb wurden Sozialwissenschaftler gebeten, der Frage der Akzeptanz nachzugehen, um so festzustellen, ob die Studie in dieser Form überhaupt durchführbar war.

Die Ergebnisse, zu denen die Sozialwissenschaftler gelangten, kamen einer Offenbarung gleich.²¹ Sie zeigten, dass eine Studie, in der «beobachtendes Abwarten» als dritte Option angeboten wird, dann annehmbar ist, wenn man sie als «aktive Überwachung» (engl. *active monitoring*) bezeichnet, wenn der Arzt sie nicht als letzte Möglichkeit anbietet und wenn er sie einfühlsam und für den Patienten verständlich erklärt.

Die Untersuchung, die die Kluft zwischen Ärzten und Patienten schließen sollte, hatte die besonderen Aspekte aufgedeckt, die beiden Parteien Schwierigkeiten bereiteten und die sich durch eine bessere Darstellung der Behandlungsoptionen leicht beheben ließen. Ein Er-

gebnis war, dass die Akzeptanzrate bei den Männern, die zur Teilnahme an der Studie eingeladen wurden, im Laufe der Zeit von vier pro zehn auf sieben pro zehn Zustimmungen anstieg. Diese schnellere Rekrutierung bedeutete, dass man über die Wirkung dieser drei Therapien für Männer mit lokalisiertem Prostatakarzinom früher Bescheid wissen würde als ohne diese Vorarbeiten. Und da Prostatakrebs häufiger vorkommt, können davon in Zukunft viele Männer profitieren, und zwar früher, als es sonst möglich gewesen wäre.

Durch Zusammenarbeit ein Zeichen für die Zukunft setzen

Es gibt zahllose Möglichkeiten, wie man Patienten und die Öffentlichkeit in Therapietests einbeziehen kann. Wie wir bereits ausgeführt haben, können sie die Hauptakteure sein – diejenigen, die Verständnislücken aufdecken und darauf aufmerksam machen, dass neue Wege gefunden werden müssen. Sie können von Wissenschaftlern dabei unterstützt werden; sie können in einigen Phasen des Forschungsprojekts mitwirken, in anderen dagegen nicht; sie können von dem Moment an beteiligt sein, in dem eine bestimmte Unsicherheit aufgedeckt wird, mit der man sich auseinandersetzen muss, indem man die Ergebnisse des Projekts verbreitet, umsetzt und in eine aktualisierte systematische Übersichtsarbeit integriert; und sie können auf ganz unterschiedliche Art und Weise in ein und dasselbe Projekt eingebunden werden. Gelegentlich initiieren sie ein Forschungsvorhaben auch selbst. Dafür gibt es keine allgemeingültige Regel: Die Angemessenheit der verschiedenen Strategien und Vorgehensweisen in der jeweiligen Studie diktiert die Wahl der Strategie. Wie die oben beschriebene Studie zum lokalisierten Prostatakarzinom zeigt, entwickeln sich die Methoden mit der Zeit weiter – selbst im Verlauf eines Projekts.

Aus der Zusammenarbeit von Patienten und Wissenschaftlern erwächst ein starkes Bündnis, auf dessen Grundlage es gelingen kann, Therapieunsicherheiten zum Wohle aller Beteiligten abzubauen. Die verschiedenen Methoden, die diese Zusammenarbeit möglich machen und die sich – mit der Billigung und Unterstützung nationaler Forschungseinrichtungen – für Einzelstudien als geeignet erweisen, verheißen Gutes für die Zukunft.

- Wenn Patienten und Wissenschaftler zusammenarbeiten, können sie zur Identifizierung und Verringerung von Therapieunsicherheiten beitragen.
- Anregungen vonseiten der Patienten können zu besserer Forschung führen.
- Manchmal gefährden Patienten faire Therapietests, ohne es zu wollen.
- Die Beziehungen zwischen Patientenorganisationen und der pharmazeutischen Industrie können zur Verzerrung von Informationen über die Wirkungen von Therapien führen.
- Um einen wirksamen Beitrag leisten zu können, müssen Patienten mehr über die Forschung wissen und einen schnelleren Zugang zu unvoreingenommenen Informationen erhalten.
- Es gibt keinen Königsweg für die Realisierung gemeinschaftlicher Mitwirkung in der Forschung.
- Die Mitwirkung von Patienten sollte sich am jeweiligen Forschungszweck orientieren.
- Die Methoden zur Einbindung von Patienten in die Forschung entwickeln sich kontinuierlich.

12

Wodurch zeichnet sich eine bessere Gesundheitsversorgung aus?

In den vorangehenden Kapiteln haben wir anhand zahlreicher Beispiele veranschaulicht, warum Therapien sich auf solide Forschung stützen sollten und können – auf Forschung also, die auf die Untersuchung patientenrelevanter Fragen ausgelegt ist. Ganz gleich, ob wir Teil der Öffentlichkeit oder Patienten oder im Gesundheitswesen beschäftigt sind: Auf die eine oder andere Weise gehen die Wirkungen medizinischer Behandlungen uns alle an. Und dabei spielt robuste Evidenz aus fairen Therapietests eine wirklich wichtige Rolle.

In diesem Kapitel gehen wir der Frage nach, wie solche Evidenz die Praxis der Gesundheitsversorgung mitgestalten kann, damit Ärzte und Patienten gemeinsam über die Behandlung Einzelner entscheiden können. Gute Entscheidungen sollten auf solider Evidenz basieren, die uns über die wahrscheinlichen Konsequenzen verschiedener Behandlungsmöglichkeiten aufklärt. Allerdings haben diese Konsequenzen für verschiedene Menschen jeweils eine unterschiedliche Bedeutung und einen unterschiedlichen Stellenwert. So gesehen kann der eine Patient auf der Grundlage derselben Evidenz zu einer ganz anderen Entscheidung gelangen als ein anderer Patient. Beispielsweise kann für einen Berufsmusiker ein voll funktionsfähiger Finger, für einen Chefkoch ein gut ausgebildeter Geruchssinn und für einen Fotografen ein gutes Sehvermögen sehr viel wichtiger sein als für andere Menschen. Um ein für sie besonders wichtiges Behandlungsergebnis zu erzielen, sind sie unter Um-

ständen bereit, größere Anstrengungen oder größere Risiken in Kauf zu nehmen. Die Schnittstelle zwischen Evidenz und Entscheidungen ist so komplex, dass sich ein Großteil dieses Kapitels einigen der öfter auftretenden Fragen zu diesem Problem widmet.

Zuvor wollen wir uns jedoch näher mit *Shared Decision Making* befassen und darlegen, wie eine solche «partizipative Entscheidungsfindung» in der Praxis aussehen könnte. Bei einer solchen partnerschaftlichen Entscheidungsfindung handelt es sich um einen Mittelweg zwischen ärztlichem Paternalismus auf der einen und Patienten sich selbst zu überlassen auf der anderen Seite. Patienten beklagen sich regelmäßig über mangelhafte Informationen, doch haben sie verständlicherweise unterschiedliche Vorstellungen davon, wie viel Verantwortung sie dabei zu übernehmen bereit sind.^{1,2}

Partizipative Entscheidungsfindung (Shared Decision-Making)

«Partizipative Entscheidungsfindung wird definiert als «der Prozess der Einbeziehung von Patienten in medizinische Entscheidungen». Dem Anspruch nach sind Ärzte darum bemüht (bzw. sollten es sein), Probleme hinreichend deutlich und offen zu beschreiben, damit Patienten die Unsicherheiten, mit denen die meisten Entscheidungen in der Medizin behaftet sind, verstehen und so gesehen auch nachvollziehen können, dass es hier Entscheidungen zwischen konkurrierenden Möglichkeiten zu treffen gilt. Die ärztliche Expertise liegt in der Diagnosestellung und dem Aufzeigen der Behandlungsoptionen nach klinischen Prioritäten; die Aufgabe des Patienten ist es, seine Wertvorstellungen und persönlichen Prioritäten, die von seinen jeweiligen sozialen Umständen geprägt sind, auf wohlinformierter Grundlage zu erkennen und mitzuteilen.»

Aus: Thornton H. Evidence-based healthcare. What roles for patients? In: Edwards A, Elwyn G, eds. *Shared decision-making in health care. Achieving evidence-based patient choice*. Second edition. Oxford: Oxford University Press, 2009, S. 39.

Manche Patienten ziehen es vor, nicht allzu genau über ihre Krankheit und ihre Behandlungsoptionen informiert zu werden, und überlassen die Entscheidungen darüber lieber ihrem beratenden Arzt; aber es gibt auch viele, die gern mehr wissen möchten. Wer sich weitergehend informieren wollen, sollte Zugang zu sorgfältig formulierten Informationsmaterialien erhalten und sich an erfahrene Gesundheitsfachkräfte wenden können, die sie dazu beraten können, wie und wo sie diese Informationen in dem für sie geeigneten Format erhalten können.

Was die «ideale ärztliche Konsultation» ausmacht, kann sich von Patient zu Patient stark unterscheiden. Manche Leute sind damit zufrieden, wenn sie eine untergeordnete Rolle übernehmen können, während andere für sich eine Führungsrolle beanspruchen. Eine stärkere Mitsprache bei der Entscheidungsfindung kann – mit ärztlicher Unterstützung – sehr befriedigend sein. Und wenn ein Patient erst einmal erlebt hat, wie so etwas funktioniert, wird es vielleicht sogar zu seinem bevorzugten Vorgehen. Wie wir unten noch sehen werden, kann schon eine einfache Frage des Patienten das Gespräch eröffnen. Wichtig ist, dass die Patienten sich in ihre medizinische Versorgung eingebunden fühlen, wenn sie, unabhängig vom Grad ihrer Mitwirkung, als ebenbürtige Partner behandelt werden.

Wie könnten die hier gemachten Vorschläge bei Ihnen selbst aussehen?

Auch wenn keine zwei Arztbesuche identisch ablaufen, so sind doch die in diesem Buch dargelegten Grundprinzipien, wie man zur bestmöglichen Entscheidung gelangt, dieselben. Das Ziel ist, dass sich sowohl beim Patienten als auch beim Arzt am Ende der Konsultation Zufriedenheit darüber einstellt, dass sie das Problem vor dem Hintergrund der besten verfügbaren relevanten Evidenz gemeinsam erörtert haben. Patienten wenden sich mit den verschiedensten gesundheitlichen Problemen an ihre Ärzte: manche mit akuten, manche mit chronischen, manche mit lebensbedrohlichen und wieder andere mit bloß «lästigen» Beschwerden. Ihre persönlichen Umstände sind zwar von einer unendlichen Vielfalt, aber sie alle haben Fragen, mit denen man sich auseinandersetzen muss, damit über das weitere Vorgehen entschieden werden kann (**Abb. 13**).

Um diesen Gedanken zu veranschaulichen, wollen wir mit einem Gespräch zwischen Patient und Arzt über ein häufig auftretendes Problem, die Kniegelenkarthrose (Gelenkverschleiß), beginnen. Im Anschluss werden wir uns einigen grundlegenden Fragen über die Anwendung wissenschaftlicher Erkenntnisse in der Praxis widmen

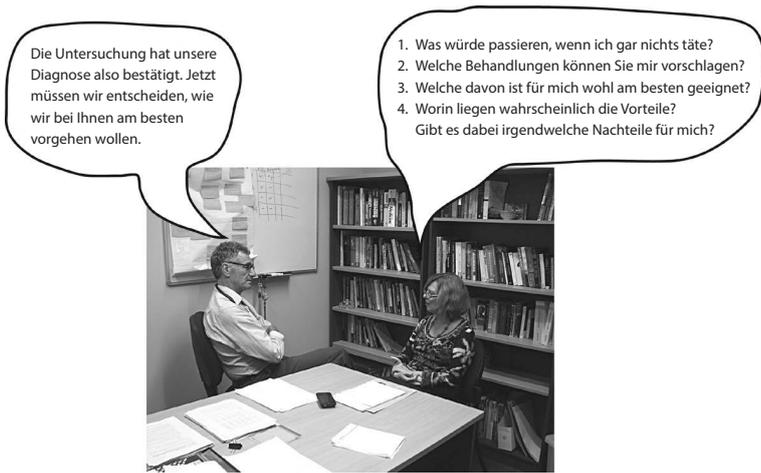


Abbildung 13: Gespräch zwischen Arzt und Patient und einige Fragen, die man stellen sollte.

– Fragen, die Patienten mit den verschiedensten Krankheitsbildern beantwortet wissen wollen, wenn sie einen Arzt aufsuchen, und Fragen, welche die Leser dieses Buchs nach der Lektüre der vorangegangenen Kapitel zu Recht werden stellen wollen.

Partizipative Entscheidungsfindung: Besuch beim Arzt wegen einer häufigen Erkrankung

Arzt: Nun, Sie leiden an einer mittelschweren Kniegelenkarthrose. Das kommt häufiger vor, wenn man älter wird. Man bezeichnet das auch als Gelenkabnutzung. Der normale Krankheitsverlauf ist fluktuierend, d.h. es geht mal besser, und mal geht es schlechter. Dieser Gelenkverschleiß schreitet aber nur langsam im Laufe von Jahren bzw. Jahrzehnten voran. Welche Beschwerden haben Sie derzeit?

Patient: Na ja, wenn ich es übertreibe, schmerzen die Knie doch ziemlich. Das kann dann über Stunden anhalten, sodass ich Probleme mit dem Schlafen habe. Vor kurzem sind die Schmerzen schlimmer geworden. Deshalb habe ich mich gefragt, ob ich wohl einen Kniegelenkersatz brauche.

Arzt: Ein Kniegelenkersatz ist sicher eine Option, aber normalerweise greifen wir auf diese Möglichkeit erst zurück, wenn einfachere Maßnahmen versagen.

Patient: Was können Sie mir also sonst vorschlagen?

Arzt: Nun, einfache Analgetika oder entzündungshemmende Medikamente können die Schmerzen lindern. Abgesehen von Medikamenten können aber auch einige spezielle Übungen zur Stärkung der Muskeln im Bereich des Kniegelenks helfen, die Funktionsfähigkeit des Knies zu erhalten und die Schmerzen zu verringern. Möchten Sie gern mehr darüber wissen?

Patient: Die Medikamente bereiten mir Magenprobleme. Deshalb würde ich gern mehr über die Übungen erfahren.

Arzt: Schön. Ich gebe Ihnen eine Broschüre mit, in der einige Übungen erklärt werden. Aber Sie sollten auch Ihren Physiotherapeuten aufsuchen. In der Zwischenzeit können Sie gegen die Schmerzen gefahrlos regelmäßig Paracetamol einnehmen und sich einfach viel bewegen.

Patient: Das ist gut zu wissen. Aber gibt es sonst keine anderen Behandlungsmöglichkeiten mehr?

Arzt: Für die Behandlung einer schweren Arthrose sind noch weitere Behandlungsmöglichkeiten verfügbar. Aber in diesem Stadium werden Sie vielleicht feststellen, dass Sie eine stetige Besserung erzielen können, wenn Sie Ihre Muskeln mithilfe der Übungen aufbauen. Dann werden Sie auch wieder besser schlafen, weil Sie weniger Schmerzen haben und insgesamt wieder mehr leisten können. Sie könnten auch schwimmen gehen oder öfter mit Ihrem Hund rausgehen. Das stärkt nicht nur Ihre Muskeln, sondern hilft Ihnen auch, sich «wohl zu fühlen». Außerdem können Sie so auch Ihr Gewicht in Schach halten! Ich denke, über drastischere Optionen brauchen wir jetzt noch nicht nachzudenken. Wir warten erst einmal ab, wie Sie mit den Übungen und dem Schmerzmittel zurechtkommen. Aber kommen Sie ruhig wieder vorbei, wenn Sie mit Ihren Fortschritten nicht zufrieden sind.

Fragen zur Umsetzung von wissenschaftlicher Evidenz in die Praxis

Frage 1: Wenn ein Patient an einer lebensbedrohlichen Krankheit leidet, lohnt es sich dann nicht, alles Erdenkliche zu versuchen?

Es kann durchaus verlockend sein, das neueste «Wundermittel» auszuprobieren oder dem Beispiel von Prominenten zu folgen, die in der Boulevardpresse Behauptungen über eine Therapie aufstellen, die sie selber ausprobiert haben wollen, vielleicht sogar ein «alternativmedizinisches» Verfahren, das sich gut verkauft, das aber nie getestet wurde. Schulmedizinische Maßnahmen muten dagegen mitunter sehr viel weniger glanzvoll und vielversprechend an. Die meisten dieser Therapien, die bei lebensbedrohlichen Krankheiten zur Anwendung kommen, sind jedoch gewissenhaften Tests unterzogen worden, um herauszufinden, wie wirksam und unbedenklich sie sind. Sich also zunächst einmal mit der besten Evidenz vertraut zu machen, kann viel Zeit, Kummer und Geld sparen.

Im Großen und Ganzen erkennt die Schulmedizin an, dass hinsichtlich der Wirksamkeit und Unbedenklichkeit der auf dem Markt befindlichen Medikamente gewisse Unsicherheiten herrschen. Sie ist deshalb bestrebt, diese Unsicherheiten durch Tests sowie die ständige und systematische Auswertung der Evidenz auf ein akzeptables Maß zu verringern, um die angebotenen Therapien auf diese Weise zu verbessern. Solche Verbesserungen hängen in ganz entscheidendem Maße von der Mitwirkung von Patienten ab, die erkannt haben, dass dies die einzige Möglichkeit ist, um einen dauerhaften Fortschritt zu erzielen.

Verständlicherweise sind Patienten mit lebensbedrohlichen Krankheiten entschlossen, alles Mögliche auszuprobieren, auch ungeprüfte «Therapien». Für sie wäre es aber deutlich besser, wenn sie die Teilnahme an einer für sie geeigneten klinischen Studie in Betracht ziehen würden, in der eine neue Therapie mit der derzeit besten Behandlung verglichen wird. Ein solcher Vergleich deckt nicht nur auf, welche zusätzlichen Vorteile die neue Behandlung möglicherweise mit sich bringt, sondern auch, welche Schäden sie

hervorrufen könnte. Lebensbedrohliche Krankheiten bedürfen mitunter hoch wirksamer Therapien – und es gibt keine Therapie, die nicht auch irgendwelche Nebenwirkungen hätte. Umso wichtiger ist es, dass eine neue Therapie gründlich und auf faire Weise getestet wird, um die Ergebnisse systematisch erfassen und feststellen zu können, ob sie den Patienten wirklich hilft.

Frage 2: Vielleicht möchten Patienten wissen, ob eine Therapie wirkt. Doch was ist, wenn sie nicht alle Einzelheiten erfahren wollen?

Es ist wichtig, das richtige Maß zwischen Informationsüberladung einerseits und dem Vorenthalten von brauchbaren Informationen andererseits zu finden, damit die Patienten eine ausreichend informierte Entscheidung treffen können. Ebenso wichtig ist es, daran zu denken, dass ein Patient, wenn er auf dem Weg zu einer Entscheidung jeweils das Für und Wider abwägen muss, anfangs nur wenige, im weiteren Verlauf aber mehr Informationen benötigt. Während der Konsultation sollte sich sowohl beim Arzt als auch beim Patienten das Gefühl einstellen, dass der Patient die Menge an Informationen erhalten hat, die er braucht, um im Weiteren gemeinsam mit dem Arzt entscheiden zu können, worin die derzeit beste Behandlungsmaßnahme für ihn besteht. Doch das ist noch nicht alles. Nachdem der Patient eine Weile über alles nachgedacht hat, hat er vielleicht noch weitere Fragen und möchte nähere Einzelheiten wissen. Dann sollte der Arzt ihm helfen herauszufinden, was er möglicherweise noch wissen muss, und ihm alles, was noch unklar ist, erklären.

Bei manchen Entscheidungen fällt die Abwägung schwer. Möglicherweise läuft es darauf hinaus, dass man das kleinere von zwei Übeln wählen muss. In Kapitel 4 haben wir beispielsweise über das Aortenaneurysma gesprochen – eine Aussackung der vom Herzen wegführenden Hauptarterie – bei dem es zur Ausbildung eines tödlich verlaufenden Gefäßblocks kommen kann. Das Problem kann durch einen größeren Eingriff behoben werden. Allerdings versterben ein oder zwei pro 100 Patienten bei der Operation selbst. In diesem Fall muss also die operationsbedingte Frühmortalität gegen

das Risiko einer späteren tödlich verlaufenden Aortenruptur abgewogen werden. Langfristig ist eine Operation die bessere Wahl. Doch entscheiden sich manche Patienten nach eigenem Ermessen gegen diesen Eingriff oder schieben ihn zumindest auf, weil noch ein wichtiges persönliches Ereignis wie z.B. die Hochzeit ihrer Tochter ansteht. Anstatt sich also blindlings in eine Lösung zu verrennen, die als «einzig Hoffnung» anmutet, ist es besser, die jeweiligen Risiken und den richtigen Zeitpunkt abzuwägen.

Frage 3: Statistiken sind verwirrend. Sollten Patienten sich wirklich mit Zahlen befassen?

Wie Zahlen manchmal präsentiert werden, kann schon sehr abschreckend wirken – oder auch geradewegs in die Irre führen. Wenn Sie zwei Behandlungen aber wirklich miteinander vergleichen oder mehr darüber herausfinden möchten, wie sich eine Krankheit, an der Sie erkrankt sind, bei anderen Menschen auswirkt, kommen immer irgendwo Zahlen ins Spiel. Einige Darstellungsformen eignen sich dafür aber besser als andere.

Um die Bedeutung von Zahlen für medizinische Laien (und auch Ärzte!) zu veranschaulichen, eignen sich Häufigkeiten am besten. Das bedeutet, es wird mit ganzen Zahlen gerechnet. So ist beispielsweise die Formulierung «15 von 100 Personen» der Angabe 15% generell vorzuziehen. Zudem hilft es oftmals, die Zahlen nicht nur mit Worten auszudrücken, sondern auch irgendwie graphisch darzustellen – z.B. in Form eines Balkendiagramms (**Abb. 14**), eines Kreis- oder Tortendiagramms, in Form von Strichmännchen/Smileys und traurigen Gesichtern in Kästen etc. oder in Tabellenform. Zahlen mit dieser Art von «Entscheidungshilfe» darzustellen, gewährleistet, dass möglichst viele Menschen die Bedeutung der Daten verstehen.

Um die Wirkung von Blutdruckmedikamenten auf das Risiko für eine Herzkrankheit und einen Schlaganfall bei Patienten mit Bluthochdruck im Laufe von zehn Jahren darzustellen, eignet sich beispielsweise ein Balkendiagramm.³

Von 100 Menschen mit nicht behandeltem Bluthochdruck werden in den nächsten zehn Jahren erwartungsgemäß 13 eine Herz-

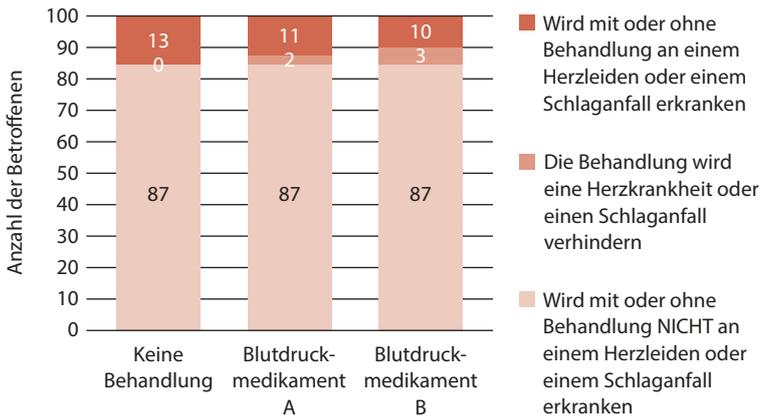


Abbildung 14: Was wird mit 100 Menschen in derselben Situation im Laufe der nächsten 10 Jahre passieren?

krankheit entwickeln oder einen Schlaganfall erleiden. Wenn alle 100 Personen das Blutdruckmedikament A nähmen, würden 11 von ihnen eine Herzkrankheit oder einen Schlaganfall bekommen – und bei zwei von ihnen ließe sich eine Herzerkrankung oder ein Schlaganfall vermeiden. Hätten alle 100 Patienten Blutdruckmedikament B eingenommen, dann käme es bei 10 von ihnen zu einer Herzkrankheit oder einem Schlaganfall, was bei dreien von ihnen vermeidbar wäre. Das ist einfach. Und doch werden diese einfachen Zahlen häufig in Worte gekleidet, die nur Statistiker verstehen können.

Sehen wir uns nun an, wie sich diese Zahlen statt in einem Balkendiagramm in einer Tabelle ausnehmen (**Tab. 3**). In diesem Beispiel konzentrieren wir uns auf die bessere Therapie – in diesem Fall Medikament B: Dazu wollen wir die Zahlen zunächst als natürliche Häufigkeiten (einfache Zahlen) ausdrücken, die wir uns im Anschluss genauer ansehen wollen.

Ohne Behandlung beträgt das Risiko für eine Herzerkrankung oder einen Schlaganfall 13 % (oder 13 von 100), während es bei Behandlung mit Medikament B 10 % (oder 10 von 100) sind – das macht einen Unterschied von 3 % (oder 3 von 100) aus. Da durch Medikament B drei von 13 Fällen einer Herzerkrankung oder eines

Tabelle 3: Herzinfarkt- bzw. Schlaganfallrisiko mit und ohne Behandlung.

	Keine Behandlung	Behandlung mit Medikament B
Herzerkrankung bzw. Schlaganfall (im Laufe von 10 Jahren)	13 von 100 Personen	10 von 100 Personen
Keine Herzerkrankung bzw. kein Schlaganfall	87 von 100 Personen	90 von 100 Personen
Insgesamt	100	100

Schlaganfalls verhindert werden, die sonst aufgetreten wären, beträgt die relative Risikoreduktion $3/13$ oder ca. 23 %. Wir können also sagen, dass mit Behandlung B die *absolute* Risikoreduktion 3 % und die *relative* Risikoreduktion 23 % beträgt. Das sind zwei verschiedene Möglichkeiten, um ein und dasselbe auszudrücken.

Die relative Risikoreduktion gibt immer einen höheren Zahlenwert an – manchmal auch einen sehr viel höheren – und erregt deshalb auch deutlich mehr Aufmerksamkeit. Wenn Sie also irgendwo eine Schlagzeile sehen, in der es heißt: «23 % Schlaganfälle verhindert», dann sagt das gar nicht aus. Denn alle weiteren Angaben zur spezifischen Personengruppe der Betroffenen, zur Zeitspanne bzw., was am wichtigsten ist, zur Höhe des Schlaganfallrisikos ohne eine Behandlung fehlen. Sehr wahrscheinlich handelt es sich in diesem Fall um die relative Risikoreduktion (aber das müssten Sie prüfen).

Manchmal unterscheiden sich die Zahlen beträchtlich voneinander. Sehen wir uns einen Zeitungsbericht über eine Studie zum Prostatakrebs-Screening an: «Könnte die Zahl der Todesfälle um 20 % senken» klingt nach viel. Die Ergebnisse hätten aber auch anders ausgedrückt werden können, und zwar als 1 verhinderter Todesfall pro 1410 gescreente Männer (oder als klitzekleine 0,07 %, d.h. die Verhütung von sieben vorzeitigen Todesfällen pro zehntausend gescreente Männer). Die 20 % geben die relative Risikoreduktion und die 0,07 % die absolute Risikoreduktion an. Letztere ist sehr viel niedriger, weil die Sterblichkeitsrate bei Prostatakrebs niedrig ist – und es wahrscheinlich nicht bis in die Schlagzeilen geschafft hätte. Die Quintessenz lautet: Wenn eine Behauptung in ei-

ner Schlagzeile viel zu optimistisch klingt, dann ist sie es wahrscheinlich auch!⁴

Es kommt also auf die Zahlen an, und richtig präsentiert können sie bei der Entscheidungsfindung sogar helfen. Patienten sollten sich nicht scheuen, ihren Arzt zu bitten, ihnen die Ergebnisse so zu erklären, dass sie sie auch problemlos verstehen können – bei Bedarf und wenn es der Klarheit dient auch mithilfe von bildlichen Darstellungen. Wenn Entscheidungen über eine Behandlung gemeinsam getroffen werden sollen, dann müssen sich Ärzte wie auch Patienten klar darüber sein, was die Zahlen tatsächlich bedeuten.

Lassen Sie sich von aufsehenerregenden Statistiken nicht ins Bockshorn jagen

«Sagen wir, Ihr Risiko, im Alter zwischen 50 und 60 Jahren einen Herzinfarkt zu erleiden, ist um 50% höher, wenn Sie einen hohen Cholesterinspiegel haben. Das hört sich ziemlich schlimm an. Wenn wir sagen, dass Ihr zusätzliches Risiko, einen Herzinfarkt zu bekommen, bei einem hohen Cholesterinspiegel nur 2% beträgt, hört sich das für mich ganz in Ordnung an. Es handelt sich aber um dieselben (hypothetischen) Zahlen. Versuchen wir es einmal damit: Von hundert Männern zwischen 50 und 60 Jahren mit einem normalen Cholesterinspiegel werden vier erwartungsgemäß einen Herzinfarkt erleiden; dagegen ist zu erwarten, dass von hundert Männern mit einem hohen Cholesterinspiegel sechs einen Herzinfarkt erleiden werden. Das macht pro hundert zwei zusätzliche Herzinfarkte.»

Goldacre B. *Bad Science*. London: Fourth Estate 2008, S. 239–240.

Frage 4: Wie kann ein Patient wissen, ob die wissenschaftliche Evidenz auf ihn zutrifft?

Alle Entscheidungen beruhen auf allen möglichen früheren – individuellen oder kollektiven – Erfahrungswerten. Faire Therapietests wie etwa randomisierte Studien sind gut strukturierte Versionen dieser Erfahrungen. Sie sind dazu ausgelegt, systematische Fehler auf ein Minimum zu begrenzen. Aber ob nun gut strukturiert oder nicht, es wird immer eine gewisse Unsicherheit bezüglich der Frage bleiben, inwieweit frühere Erfahrungen in eine Empfehlung für den nächsten Patienten einfließen können. Wenn also die Patienten, die in fairen Therapietests untersucht wurden, eine ähnliche Krankheit in einem ähnlichen Stadium oder mit einem ähnlichen Schweregrad hatten wie der fragliche Patient, dann lautet die vernünftigste Annahme, dass dieser Patient ähnlich auf die Therapie anspricht – es

sei denn, es liegt ein guter Grund für die Annahme vor, dass die Studienpatienten ganz andere Eigenschaften hatten oder sich ihre Krankheit deutlich von der des fraglichen Patienten unterschied.

Aber auch wenn die Evidenz auf ihn anwendbar wäre, könnte der Patient natürlich zu Recht fragen: «Die Menschen sind doch alle verschieden. Also werden sie doch auch unterschiedlich auf Therapien ansprechen?» Ein «fairer Therapietest» sagt uns nur, wie eine Therapie im Durchschnitt wirkt, garantiert aber nur selten, dass sie bei allen Menschen gleich gut wirkt; und sie kann üblicherweise nicht vorhersagen, bei wem unerwünschte Nebenwirkungen auftreten werden. Die wissenschaftliche Evidenz kann Orientierungshilfen geben, welche Behandlung wahrscheinlich am besten ist, damit sie dann beim einzelnen Patienten ausprobiert werden kann. Bei einem Exanthem (ein entzündlicher Hautausschlag) beispielsweise könnte die evidenzbasierte Therapie auf eine Körperregion angewendet werden, während ein anderer Körperbereich als Kontrolle dient (s. Kap. 6, S. 120). Durch den Vergleich des Ansprechens in den beiden Körperregionen können Arzt und Patient erkennen, ob die Therapie wirkt oder ob eine unerwünschte Wirkung auftritt. Wenn dermatologische Therapien – z. B. Aknetherapien im Gesicht – zum ersten Mal bei einem Patienten angewendet werden, ist es sogar üblich, zunächst einmal ein «Testpflaster» aufzubringen.

Meistens haben wir aber nicht das Glück, einen so direkten Vergleich zur Hand zu haben. Bei einigen chronischen und nicht lebensbedrohlichen Problemen wie Schmerzen oder Juckreiz kann man ein Medikament beim selben Patienten in bestimmten Zeiträumen wiederholt geben und absetzen. Ein solches Vorgehen wird auch als N-gleich-1-Studie (engl.: *n-of-1 trial*) bezeichnet; das bedeutet, dass die Anzahl der Teilnehmer (N) in der Studie eins beträgt – also einen einzigen Patienten umfasst. Die Prinzipien, die wir in Kapitel 6 für einen fairen Vergleich erarbeitet haben, gelten auch für solche Studien mit einzelnen Patienten, einschließlich einer unverzerrten oder verblindeten Bewertung des Behandlungsergebnisses usw. Im Idealfall würden wir für die Hauttherapien oder Tabletten dann Placebokontrollen einsetzen, aber das lässt sich oftmals nur schwer organisieren.

Bei vielen Krankheiten haben wir nicht die Möglichkeit, eine Therapie auszuprobieren und abzuwarten, was passiert: Das Behandlungsergebnis ist nicht zeitnah verfügbar oder zu ungewiss. Ob beispielsweise Aspirin bei einem Patienten einen Schlaganfall verhindert, kann man erst erkennen, wenn es zu spät ist. In der Präventivmedizin stellt das in den meisten Fällen ein Problem dar. Dasselbe gilt auch für die Behandlung zahlreicher akuter Erkrankungen wie z.B. Hirnhautentzündung (Meningitis), Lungenentzündung (Pneumonie) oder Schlangenbisse, bei denen wir keine Gelegenheit haben, die Therapie bei jedem einzelnen Patienten zu testen und abzuwarten. Deshalb sind wir darauf angewiesen zu prüfen, ob und wie wir die Evidenz aus den Erfahrungen mit der Untersuchung anderer Patienten anwenden können.

Wenn wir in der Praxis dann Glück haben und die Evidenz sich auf einen Patienten anwenden lässt, muss man sich unbedingt fragen, inwiefern sich der Schweregrad der Krankheit bei diesem Patienten (bzw. bei Gesunden das vorhergesagte Risikoausmaß) mit dem der Studienpatienten vergleichen lässt. Im Allgemeinen profitieren schwerer erkrankte Patienten stärker von einer Therapie als weniger schwer erkrankte. Wenn also der Schweregrad gleich oder höher ist als in den Studien, in denen der Nutzen einer Therapie nachgewiesen wurde, dann können wir in der Regel darauf vertrauen, dass sich die Evidenz auf unseren Patienten übertragen lässt. Wenn ihre Krankheit weniger schwer ist (bzw. die noch Gesunden ein vergleichsweise geringeres vorhergesagtes Risiko aufweisen), lautet die wichtigste Frage, ob auch ein Nutzen, der geringer als der in den Studien beobachtete ausfällt, noch als lohnenswert angesehen werden kann.

Frage 5: Bedeuten genetische Untersuchungen und «personalisierte Medizin» nicht, dass Ärzte für jeden einzelnen Patienten eine spezifische Therapie entwickeln könnten, und dass alles, was wir hier erörtert haben, damit hinfällig wird?

Obwohl der Gedanke, dass wir bei Bedarf für jeden einzelnen Patienten eine spezifische Therapie entwickeln könnten, zweifellos verlockend und für einige wenige Krankheiten vielleicht auch mach-

bar ist, so erscheint es doch sehr unwahrscheinlich, dass sich dieser Ansatz in der Behandlung von Patienten auf breiter Front durchsetzen wird. Wie wir schon im Zusammenhang mit der Besprechung genetischer Untersuchungen erläutert haben (Kap. 4, S. 81 f.), hängt der Ausbruch der meisten Krankheiten nicht nur von komplexen Wechselwirkungen zwischen mehreren Genen, sondern auch von den sogar noch viel komplexeren Wechselwirkungen zwischen Genen und Umweltfaktoren ab.

Die Ergebnisse genetischer Analysen sind wichtig als Entscheidungshilfe für Familien und Einzelpersonen mit erblichen Krankheiten wie Chorea Huntington, Thalassämien (erbliche Blutungsstörungen) sowie einigen anderen (meist seltenen) Krankheiten. Diese genetischen Informationen erweisen sich bei der Beratung von Familien, in denen diese Krankheiten gehäuft auftreten, als großer Segen. Was aber häufiger auftretende Krankheiten, an denen wir alle erkranken können, betrifft, so liefert die genetische Untersuchung zusätzlich zu den Informationen, die aus der Familienanamnese und der klinischen Untersuchung bekannt sind, allerdings nur wenig Neues. Auch wenn sich diese Situation wahrscheinlich noch ändern wird, bedeutet unser derzeit begrenztes Wissen, dass wir uns vor einer Überbewertung der Risiken für die häufigen Krankheiten, die sich auf der Grundlage von Genanalysen vorhersagen lassen, hüten müssen.

Wir sollten an dieser Stelle erklären, dass keiner der Autoren dieses Buchs sein genetisches Profil in Auftrag gegeben hat oder dies zu tun beabsichtigt. Deshalb sollte es Sie auch nicht wundern, dass wir uns im Allgemeinen gegen genetische Untersuchungen aussprechen, es sei denn, ein Patient hat entweder eine Familienanamnese, die Anhaltspunkte gibt für das Vorliegen einer bestimmten bekannten erblichen Krankheit oder eine der wenigen derzeit bekannten Krankheiten, bei denen ein Gen oder mehrere Gene eindeutig vorhersagen, wer von ihnen auf eine Therapie ansprechen wird.

Frage 6: Wie kann jemand, der an einer Krankheit leidet, die gerade in einer laufenden klinischen Studie untersucht wird, und dessen Arzt nichts von dieser Studie weiß, doch etwas darüber herausfinden?

Weniger als einer von 100 Arztbesuchern wird für eine klinische Studie angemeldet. Je nach Krankheit und äußeren Bedingungen unterliegt dieser Anteil erheblichen Schwankungen. Selbst innerhalb von Krebszentren – in denen Studien allgemein anerkannt sind und weithin genutzt werden – herrscht in diesem Punkt eine enorme Bandbreite: Die meisten an Krebs erkrankten Kinder werden in Studien aufgenommen, aber bei den Erwachsenen sind es weniger als einer von zehn. In den meisten Fällen richtet sich die Aufnahme in eine Studie danach, welches Zentrum der Patient zur Behandlung aufsucht: Nimmt das Zentrum nicht an einer Studie teil, dann können sie auch keine Patienten rekrutieren. Deshalb müssen sich Patienten unter Umständen nach einem Zentrum erkundigen, das sich an klinischen Studien beteiligt. Es gibt einige ambulante Studien, für die sich Patienten direkt anmelden können. Solche Studien werden häufiger im Rahmen von Forschungsvorhaben durchgeführt, anhand derer man herausfinden möchte, wie man Menschen mit psychischen Gesundheitsproblemen wie z. B. einer Depression oder Angststörung helfen kann. Für einige andere Studien werden die Teilnehmer neuerdings auch direkt über das Internet rekrutiert. Beispielsweise wurden für eine vor kurzem durchgeführte Studie, in der die Wirkungen von Dehnübungen vor der Aufnahme sportlicher Aktivitäten beurteilt werden sollten, alle Teilnehmer auf diese Weise rekrutiert: Sie haben niemals eine Klinik aufgesucht, sondern sämtliche Anweisungen über das Internet erhalten. Auch die Nachbeobachtung wurde über das Internet durchgeführt.

Wenn Ärzte zögern, ihre Patienten für die Teilnahme an Studien anzumelden, sollten die Patienten versuchen, den Grund dafür herauszufinden. Es könnte beispielsweise sein, dass der Patient nicht wirklich für die Studienteilnahme geeignet ist. Es kann aber auch einfach daran liegen, dass der Arzt die zusätzliche Arbeit in der Form von belastenden regulatorischen Anforderungen scheut (s. Kap. 9). Patienten, die glauben, dass sie sich wahrscheinlich für die Teilnah-

me an einer laufenden Studie eignen, sollten daher nicht locker lassen. Wenn eine passende Studie läuft und ein Patient klarstellt, dass er unbedingt daran teilnehmen möchte, sollte sein Arzt dies unterstützen (s. a. den Abschnitt «Zusätzliche Quellen» ab. S. 247).

Frage 7: Wie lässt sich am besten feststellen, ob die Evidenz (im Netz oder anderswo) zuverlässig ist? Worauf sollte man dabei achten?

Leider gibt es keine hundertprozentig verlässlichen Kennzeichen, an denen sich ablesen lässt, ob eine Information zuverlässig ist. Wenn man sich selber nicht mit der Originalliteratur befassen will, muss man sich in diesem Punkt auf die Einschätzung Dritter verlassen. Deshalb ist es wichtig, die mutmaßliche Kompetenz der betreffenden Person (oder Organisation) zu beurteilen und festzustellen, ob es da irgendwo Interessenkonflikte (oder eine offene Rechnung) gibt. Wenn nicht, sollten Sie sich fragen, ob Sie darauf vertrauen, dass diese Person die besten Forschungserkenntnisse aufgespürt und ausgewertet hat: Wurden die Forschungsergebnisse beschrieben? Sind entsprechende Literaturangaben vorhanden?

Angenommen, jemand möchte z.B. wissen, ob Beta-Carotin (Vorstufe von Vitamin A) das Krebsrisiko erhöht oder es senkt. Eine Suche in Google nach «beta-carotene cancer» liefert mehr als 800 000 Treffer. Sieht man sich die ersten zehn Treffer an, dann finden sich darunter vier Primärstudien und sechs Übersichtsarbeiten oder Meinungen. Unter dreien dieser sechs Links standen Werbeanzeigen für Vitamine oder alternative Medikamente auf derselben Seite: ein beunruhigendes Zeichen.

Auf einer dieser schlechteren Webseiten heißt es:

Frage: Beugt Beta-Carotin Krebs vor? Antwort: Studien haben gezeigt, dass Beta-Carotin helfen kann, das Krebsrisiko zu senken. Beta-Carotin kommt in gelben, roten und dunkelgrünen Gemüsesorten, aber auch in Früchten vor. Man geht allgemein davon aus, dass die Einnahme eines Beta-Carotin-haltigen Nahrungsergänzungsmittels dieselbe Wirkung hat wie der Verzehr von Obst und Gemüse, das Beta-Carotin

enthält. Dies ist aber nicht der Fall. In Studien wurde unter den Studienteilnehmern ein erhöhtes Lungenkrebsrisiko festgestellt.

Abgesehen von den Werbeanzeigen werden zwar «Studien» erwähnt, sie werden aber nicht näher beschrieben bzw. es werden keine Literaturangaben zu veröffentlichten Studien gemacht – das sollte Sie aufhorchen lassen. Es lässt sich unmöglich sagen, ob der Autor diese «Studien» recherchiert und ausgewertet hat oder lediglich über Studien gestolpert ist, deren Schlussfolgerungen ihm gefallen haben.

Stellen wir dieser Website den Eintrag in Wikipedia gegenüber (der ebenfalls unter den ersten zehn Treffern genannt wird):

Eine Übersichtsarbeit über alle randomisierten, kontrollierten Studien im wissenschaftlichen Schrifttum, die von der Cochrane Collaboration erstellt und 2007 in JAMA veröffentlicht wurde, ergab, dass Beta-Carotin die Sterblichkeit um 1 bis 8% erhöht (relatives Risiko 1,05, 95%-Konfidenzintervall 1,01 bis 1,08).^[15] Allerdings umfasste diese Meta-Analyse zwei große Studien über Raucher, sodass nicht klar ist, ob die Ergebnisse auf die Allgemeinbevölkerung übertragbar sind.^[16]

Dieser Eintrag nennt die Art der Evidenz (randomisierte Studien) und gibt die Literaturstellen an (die hochgestellten Zahlen in eckigen Klammern). Die Tatsache, dass keine Werbeanzeigen auf der Seite zu finden sind und spezifische Angaben zur Evidenz gemacht werden, schafft Vertrauen.

Frage 8: Gibt es zuverlässige, empfehlenswerte Informationsquellen?

Es gibt keine Informationsquelle, die über *alle* Krankheiten und Therapien Auskunft gibt. Um die in diesem Buch vorgestellten Prinzipien anzuwenden, möchte vielleicht der eine oder andere Leser seine Fertigkeiten in diesem Bereich noch weiterentwickeln. Ergänzend zu Kapitel 6 bis 8 in diesem Buch gibt das Buch *Smart Health Choices*⁵ etliche Tipps, wie man hochwertige Gesundheitsin-

formationen findet und worauf man dabei achten sollte. Ab S. 247 finden Sie weitere Lektüreempfehlungen.

Nur wenige der vorhandenen Webseiten beruhen vorwiegend auf systematischen Übersichtsarbeiten. Zu diesen gehören die *Cochrane Database of Systematic Reviews* (www.cochrane.org/cochrane-reviews oder summaries.cochrane.org), die vorwiegend englischsprachige Zusammenfassungen für Laien enthält, sowie die Internetseite des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG; www.gesundheitsinformation.de). Zudem gibt es zahlreiche Webseiten in englischer Sprache, die generell gute Informationen bereitstellen, sich aber nicht immer auf systematische Reviews der besten verfügbaren Evidenz stützen – beispielsweise *NHS Choices* (www.nhs.uk) und *PubMed Health* (www.pubmed.gov/health).

Natürlich muss man auf der Hut sein. Achten Sie vor allem auf Interessenkonflikte, etwa auf Internetseiten, denen finanzielle Vorteile daraus erwachsen, wenn die Besucher der Webseite die angebotenen Informationen glauben, oder Seiten, die etwas zu verkaufen versuchen. Manchmal lässt sich das allerdings nur schwer erkennen: Beispielsweise erhalten, wie in Kapitel 11 erwähnt, manche Patientenvertretungen finanzielle Unterstützung von Pharmaunternehmen, die nicht deklariert ist und den dargebotenen Informationen eine gewisse «Färbung» verleiht.

Frage 9: Wie kann man vermeiden, dass man mit einem «Krankheitsetikett» versehen und unnötig behandelt wird?

Die Medizin hat erstaunliche Fortschritte gemacht: z. B. Impfstoffe und Antibiotika zur Verhütung und Behandlung von Infektionen, Gelenkersatz, Kataraktoperation und die Therapie von Krebserkrankungen bei Kindern, um nur einige wenige zu nennen. Doch dieser Erfolg verführt die Medizin auch dazu, ihre Fühler in Bereiche vorzustrecken, in denen sie weniger Nutzen mit sich bringt. Für einen Mann mit einem Hammer sieht die ganze Welt wie ein Nagel aus, und für einen Arzt (oder einen Arzneimittelhersteller!) mit einer neuen Therapie sieht alles nach Krankheit aus. Weil z. B. bessere Therapien gegen Diabetes und Bluthochdruck verfügbar geworden

sind, sind Ärzte versucht, die Anwendung dieser Therapien auch ihren Patienten vorzuschlagen, deren Befunde nur geringfügig von den Normalwerten abweichen. Dadurch steigt die Anzahl der Menschen, die als Diabetiker oder Hypertoniker etikettiert werden, dramatisch an, sodass viele Menschen «medikalisiert» werden, deren Werte früher als normal eingestuft worden wären.

Abgesehen von den unerwünschten Wirkungen der (zuweilen unnötigen) Behandlungen hat diese «Etikettierung» auch psychische und soziale Konsequenzen, die das persönliche Wohlbefinden beeinträchtigen und Probleme am Arbeitsplatz oder beim Versicherungsträger mit sich bringen können. Deshalb ist es für Patienten und Öffentlichkeit wichtig, diese Ereigniskette zu erkennen und erst einmal innezuhalten, um den wahrscheinlichen Schaden gegen den Nutzen abzuwägen, bevor man einer Therapie allzu schnell zustimmt. Wie wir in Kapitel 4 dargelegt haben, sind für diese problematische Etikettierung häufig Screenings verantwortlich, die mit Überdiagnosen und potenzieller Übertherapie einhergehen.

Die erste Verteidigungsstrategie ist Wachsamkeit gegenüber solchen Etiketten und weiteren Untersuchungsangeboten. Die scheinbar flapsige Bemerkung, dass ein normaler Mensch jemand ist, der noch nicht ausreichend untersucht wurde, hat einen sehr ernsten

Wer ist Diabetiker?

Wie wird also entschieden, wer Diabetes hat? Als ich Medizin studierte, galt folgende einfache Regel: Wenn man einen Nüchternblutzucker von mehr als 140 hatte, war man Diabetiker. 1997 wurde diese Krankheit vom *Expert Committee on the Diagnosis and Classification of Diabetes Mellitus* neu definiert. Danach gilt man bereits ab einem Nüchternblutzuckerspiegel von 126 als Diabetiker. Jeder, der einen Blutzuckerspiegel zwischen 126 und 140 hat, wurde früher als normal angesehen, leidet nach dieser Definition nun aber an Diabetes. Diese geringfügige Änderung hat aus mehr als 1,6 Millionen Menschen Patienten gemacht. Stellt das ein Problem dar? Kann sein, kann auch nicht sein. Weil wir die Regeln geändert haben, behandeln wir heute mehr Menschen wegen Diabetes. Das kann bedeuten, dass wir bei einigen dieser neuen Patienten die Wahrscheinlichkeit für die Entwicklung diabetischer Komplikationen verringert haben. Doch da diese Patienten einen leichteren Diabetes haben (d. h. vergleichsweise niedrige Blutzuckerspiegel zwischen 126 und 140), ist ihr Risiko für diese Komplikationen aber zunächst einmal auch vergleichsweise gering.

Welch HG, Schwartz LM, Woloshin S. *Overdiagnosed: making people sick in the pursuit of health*. Boston: Beacon Press, 2011: S. 17–18.

Hintergrund. Deshalb ist es klug, stets nachzufragen, ob eine Krankheit mit einem hohen oder einem niedrigen Risiko einhergeht, und auch zu fragen, was passieren könnte, wenn man erst einmal abwarten würde: Wie könnte die Krankheit überwacht werden, und was wäre ein Zeichen dafür, dass man handeln müsste? Manche Ärzte sind erleichtert, wenn Patienten nicht gleich behandelt werden wollen oder nicht sofort weitere Untersuchungen verlangen. Es gibt aber auch Ärzte, die in die Etikettenfalle – Etikett = Krankheit = obligatorische Behandlung – hineintappen, weil sie nicht erkennen, dass der Patient vielleicht damit zufrieden wäre, erst einmal abzuwarten, um zu sehen, ob seine Beschwerden von allein besser werden oder sich verschlimmern.

Wie geht es weiter?

Die oben angesprochenen Probleme – mit den individuellen Sorgen und Wertvorstellungen, mit dem Verständnis statistischer Zahlen und ihrer Anwendung auf Einzelpatienten und mit den Bedenken gegenüber der Ausweitung wirksamer Therapien auf immer leichtere Krankheitsschweregrade – zeigen alle, dass eine bessere Kommunikation vonnöten ist: und zwar zwischen Arzt und Patient sowie zwischen dem Gesundheitswesen und den Bürgern, denen es dient. Deshalb beenden wir dieses Kapitel mit der Salzburger Erklärung zur partizipativen Entscheidungsfindung, in der an die verschiedenen Gruppen gerichtete Handlungsempfehlungen formuliert sind, die auf eine Verbesserung unserer Zusammenarbeit abzielt.^{6,7}

Die Salzburger Erklärung zur partizipativen Entscheidungsfindung

Wir rufen Ärzte dazu auf,

- anzuerkennen, dass es ihre ethische Pflicht ist, wichtige Entscheidungen zusammen mit ihren Patienten zu treffen;
- einen wechselseitigen Informationsfluss anzuregen und ihre Patienten zu ermutigen, Fragen zu stellen, ihre Lebensumstände zu erklären und ihre persönlichen Präferenzen zum Ausdruck zu bringen;
- genaue Informationen über Behandlungsalternativen, deren Unwägbarkeiten, Nutzen und mögliche Folgeschäden entsprechend den Grundsätzen guter Risikokommunikation zu vermitteln;

- die Informationen auf die individuellen Bedürfnisse von Patienten zuzuschneiden und ihnen genügend Zeit zu lassen, ihre Wahlmöglichkeiten abzuwägen;
- anzuerkennen, dass die meisten Entscheidungen nicht sofort getroffen werden müssen, und Patienten und deren Familien Zeit, Ressourcen und Unterstützung für ihre Entscheidungsfindung zu geben.

Wir rufen Ärzte, Forscher, Verleger, Journalisten und andere Akteure dazu auf,

- sicherzustellen, dass sie Informationen veröffentlichen, die klar, evidenzbasiert und aktuell sind, und dass sie Interessenskonflikte offen legen.

Wir rufen Patienten dazu auf,

- sich zu trauen, Bedenken und Fragen offen zu äußern und deutlich zu machen, was für sie wichtig ist;
- zu erkennen, dass sie ein Recht darauf haben, an ihrer Versorgung gleichberechtigt mitzuwirken;
- qualitativ hochwertige Informationen zu Gesundheitsfragen zu suchen und zu nutzen.

Wir rufen politische Entscheidungsträger dazu auf,

- politische Rahmenbedingungen zu schaffen, die die partizipative Entscheidungsfindung einschließlich deren Messung als Anreiz zur Verbesserung fördern;
- durch Gesetzesänderungen die Entwicklung von Fähigkeiten und Instrumenten zur partizipativen Entscheidungsfindung zu unterstützen.

Warum?

Ein großer Teil der Versorgung wird gemäß den jeweiligen Fähigkeiten und der Bereitschaft der Ärzte erbracht, eben diese Versorgung bereitzustellen. Nicht immer entspricht sie den allgemein anerkannten Standards oder geht auf die Wünsche und Präferenzen der Patienten ein.

Ärzte erkennen häufig nur unzureichend, wie groß der Wunsch von Patienten nach Beteiligung ist, um ihre gesundheitlichen Probleme zu verstehen, für sie zur Verfügung stehende Behandlungsalternativen zu kennen und Entscheidungen auf Grundlage ihrer persönlichen Präferenzen zu treffen.

Viele Patienten und deren Familien erleben es als schwierig, bei Entscheidungen über ihre eigene Gesundheitsversorgung einen aktiven Part zu übernehmen. Manche scheuen sich davor, das Handeln von Ärzten und anderen Gesundheitsfachkräften zu hinterfragen. Viele haben nur begrenzte Kenntnisse über das Thema Gesundheit und gesundheitliche Einflussfaktoren und wissen nicht, wo sie Informationen bekommen können, die klar, vertrauenswürdig und leicht verständlich sind.

13

Aus den richtigen Gründen forschen: ein Entwurf für eine bessere Zukunft

Zweifellos hat die medizinische Forschung ihren Beitrag zu einer besseren Lebensqualität und einer höheren Lebenserwartung geleistet. Und doch haben wir in diesem Buch gezeigt, dass die bestehenden – wirtschaftlichen wie auch akademischen – «Impulsgeber» der Forschung nicht genug tun, um die Prioritäten von Patienten zu identifizieren und sich mit ihnen auseinanderzusetzen.

Für die medizinische Forschung werden riesige Geldsummen ausgegeben: Weltweit sind es mehr als 100 Milliarden Dollar pro Jahr.¹ Der größte Teil dieser Fördermittel wird allerdings in Laborstudien und Tierversuche investiert, anstatt in Studien zu fließen, die für Patienten wahrscheinlich relevantere Erkenntnisse liefern können.

Selbst wenn es um die Entscheidung geht, welche Fragen zu Therapieeffekten untersucht werden sollen, werden die Prioritäten der Patienten weitgehend ignoriert. Aufgrund ihrer Wirtschaftsmacht übt die Pharmaindustrie einen großen Einfluss auf die Entscheidungen über die jeweiligen Forschungsthemen aus. Da die Industrie für jeden Patienten, der für ihre klinischen Studien rekrutiert wird, großzügig bezahlt (Tausende von britischen Pfund/amerikanischen Dollar), beteiligen sich Akademiker – und die Institutionen, für die sie arbeiten – viel zu oft an klinischen Studien, die sich mit Fragen befassen, die eher für die Industrie als für die Patienten von Interesse sind.

Bedauerlicherweise wird ein Großteil des für medizinische Forschung zur Verfügung gestellten Geldes auf verschiedenen Stufen vergeudet – und zwar dadurch, dass die falschen Forschungsfragen gestellt werden, dass Studien durchgeführt werden, die unnötig oder schlecht geplant sind, dass versäumt wird, die Forschungsergebnisse vollständig zu veröffentlichen und zugänglich zu machen, und dadurch, dass verzerrte und unbrauchbare Forschungsberichte erstellt werden (**Abb. 15**). Das sollte uns allen zu denken geben – Forschern, Geldgebern, Ärzten, Steuerzahlern und vor allen Dingen den Patienten.

Bevor wir unseren Entwurf für eine bessere Zukunft vorstellen, wollen wir uns kurz folgenden Fragen widmen: Wenn Forschung besser werden soll, warum ist es dann unerlässlich,

1. die richtigen Forschungsfragen zu stellen
2. Forschungsvorhaben richtig zu planen und auszuführen
3. sämtliche Ergebnisse zu veröffentlichen und zugänglich zu machen
4. unvoreingenommene und brauchbare Forschungsberichte zu erstellen?

Leitsätze für bessere Forschung

1. Die richtigen Forschungsfragen stellen

Manchmal wissen Ärzte nicht, welche Behandlung für ihre Patienten wahrscheinlich am besten ist, weil die verfügbaren Therapiemöglichkeiten nicht richtig untersucht worden sind. Es kann sein, dass Industrie oder Wissenschaft an Studien, die wichtige Auswirkungen auf die Patientenversorgung haben können, nur wenig oder gar kein Interesse haben. Dadurch bleiben wichtige Fragen unbeantwortet. Die Nichtbeantwortung dieser Fragen kann aber ungeheures Leid hervorrufen. Sehen wir uns als Beispiel einfach nur folgende Frage an: Erhöht oder mindert die Gabe von Steroiden die Überlebenschancen von Patienten mit Hirnschäden, die sie sich in-

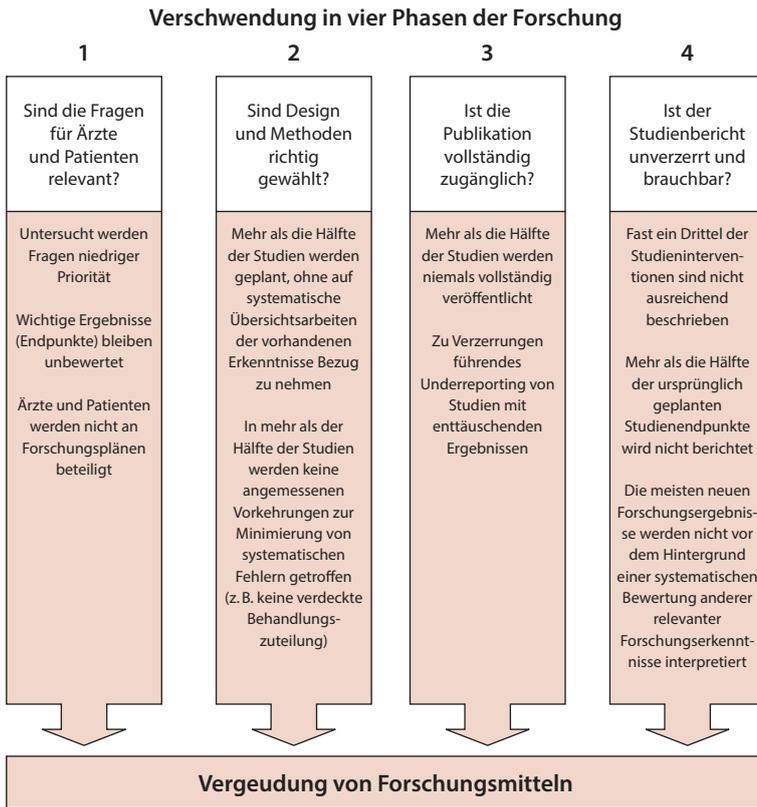


Abbildung 15: Wie die in die medizinische Forschung fließenden Gelder auf den verschiedenen aufeinanderfolgenden Stufen verschwendet werden.¹

folge einer körperlichen Verletzung zugezogen haben? Jahrzehntlang wurden Steroide angewendet, bevor eine gut geplante Studie zeigte, dass diese etablierte Therapie wahrscheinlich Zehntausenden von hirnerkrankten Patienten das Leben gekostet hatte.² Industrie sowie verschiedene an Universitäten tätige Forscher hatten Vorschlägen für eine dementsprechende Studie anfänglich Widerstand entgegengebracht. Der Grund dafür war einfach: Sie waren an kommerziellen Studien beteiligt, in denen sie die Wirkungen teurer neuer Medikamente (sogenannter neuroprotektiver Substanzen) auf Zielgrößen untersuchten, die für die Patienten von fragwürdiger

Relevanz waren. Zudem wollten sie den Wettbewerb um Studienteilnehmer vermeiden.

Ein weiterer Grund dafür, sich dieser bislang unbeantworteten Fragen anzunehmen, besteht darin, dass dies dazu beiträgt, die kostbaren Ressourcen, die uns für die Gesundheitsversorgung zur Verfügung stehen, nicht zu verschwenden. Als in den 1940er-Jahren Humanalbumin-Lösungen, die als intravenöse Infusion verabreicht wurden, zur Wiederbelebung (Reanimation) von Verbrennungsoptionen und anderen kritisch kranken Patienten eingeführt wurden, ging man theoretisch davon aus, dass sie das Sterberisiko dieser Patienten reduzieren würden. Erstaunlicherweise wurde diese Theorie erst in den 1990er-Jahren fairen Tests unterworfen. Zu diesem Zeitpunkt konnten in einer systematischen Übersichtsarbeit keine Belege dafür gefunden werden, dass Humanalbumin-Lösungen das Sterberisiko im Vergleich zu einfachen Salzlösungen, einer anderen im Rahmen der Reanimation verabreichten Flüssigkeit, senkten. Wenn Albumin überhaupt einen Einfluss auf das Sterberisiko hatte, dann wurde dieses Risiko – das zeigte dieser systematische Review – dadurch höchstens erhöht.³ Diese Erkenntnisse veranlassten Ärzte in Australien und Neuseeland, sich zusammenzutun, um die erste ausreichend große faire Vergleichsstudie zu Humanalbumin-Lösungen und Kochsalz (Salzwasser) durchzuführen.⁴ In dieser Studie – die eigentlich ein halbes Jahrhundert früher hätte durchgeführt werden sollen – konnte der Nachweis, dass Albumin besser als Kochsalzlösung war, nicht erbracht werden. Da Albumin ungefähr zwanzigmal teurer ist als Kochsalzlösung, sind somit im Laufe der letzten 50 Jahre weltweit gigantische Geldsummen aus den Gesundheitsetats verschwendet worden.

2. Forschungsvorhaben richtig planen und durchführen

Angeregt durch Umfragen, in denen die schlechte Qualität vieler Berichte zu klinischen Studien beklagt wurde, hat man Berichtsstandards entwickelt, die inzwischen angewendet werden. Solche Standards lassen erkennen, wie viele Patienten zur Teilnahme an einer Studie eingeladen wurden und wie viele von ihnen eine Teilnahme abgelehnt haben. Die Ergebnisse werden den zu Beginn der

Studie festgelegten Behandlungsgruppen entsprechend dargestellt. Wir haben aber noch einen weiten Weg vor uns. Verbesserungsbedürftig sind nach wie vor (a) die Wahl der in einem Forschungsvorhaben zu untersuchenden Fragen, (b) die Art und Weise, wie diese Fragen formuliert werden, um zu gewährleisten, dass es sich bei den für die Bewertung ausgewählten Behandlungsergebnissen auch um diejenigen handelt, welche die Patienten als wichtig erachten, und (c) die Bereitstellung von Informationen für Patienten (s. Kap. 11 und 12).

Um festzustellen, ob eine beantragte Studie durchführbar und akzeptabel ist, können Vorarbeiten unter Beteiligung einer Gruppe von Patienten sinnvoll sein. Sie helfen Schwachstellen in den Entwürfen aufzudecken oder Endpunkte mit höherer Relevanz festzulegen, oder lassen vielleicht sogar erkennen, dass schon das Konzept an sich ein «Rohrkrepierer» ist.^{5,6}

Dadurch könnte eine Menge Zeit, Geld und Frustration gespart werden. Die klinische Studie über Männer mit lokalisiertem Prostatakarzinom, über die wir in Kapitel 11 (S. 196) berichtet haben, hat gezeigt, wie das Studiendesign verbessert werden konnte, indem man die Formulierungen, welche die Ärzte bei der Erläuterung der Studienziele und der Behandlungsoptionen benutzten, sorgfältig auswählte. Die Erforschung der Sichtweise von Patienten half, die Bedenken und Informationsbedürfnisse der zur Studienteilnahme eingeladenen Männer zu identifizieren. Diese konnten dann in den Informationen, die man den potenziellen Teilnehmern gab, berücksichtigt werden, was eine akzeptable Studie möglich machte.⁷

3. Sämtliche Ergebnisse veröffentlichen und zugänglich machen

Die selektive Berichterstattung von Forschungsergebnissen kann zu schwerwiegenden systematischen Fehlern führen. Einige «negative» Studien kommen niemals zur Veröffentlichung, weil ihre Ergebnisse den Erwartungen der Studienleiter oder Sponsoren nicht entsprechen. Ohne die Veröffentlichung eines Studienberichts, der über die Studie Auskunft gibt, verschwinden diese Studien aber von der Bildfläche, ohne die geringste Spur zu hinterlassen.⁸ Zudem kann es

vorkommen, dass selektiv über die Ergebnisse von veröffentlichten Studien berichtet wird, d. h., dass einige Ergebnisse von der Veröffentlichung ausgeschlossen werden, weil sie für die untersuchte Therapie nicht so «positiv» oder günstig ausfallen.⁹ Es mussten schon Patienten leiden und sterben, weil Forschungsberichte über die Wirkungen von Therapien verzerrt wiedergegeben wurden. Deshalb ist eine solche Vorgehensweise sowohl unethisch als auch unwissenschaftlich.

4. Unverzerrte und brauchbare Forschungsberichte erstellen

Aber auch wenn Studien veröffentlicht werden, so werden häufig wichtige Elemente weggelassen, die dem Leser eine echte Bewertung und Anwendung der Studienergebnisse ermöglichen würden. In einer Übersichtsarbeit zu 519 randomisierten Studien, die im Dezember 2000 in angesehenen Fachzeitschriften publiziert worden war, wurde festgestellt, dass 82 % der analysierten Studien keine Informationen zur Geheimhaltung der Behandlungszuteilung enthielten, und in 52 % keine näheren Angaben darüber gemacht worden waren, welche Maßnahmen sie zur Reduzierung des Beobachterbias ergriffen hatten – beides sind, wie wir in Kapitel 6 erläutert haben, wichtige Kennzeichen guter Studien.¹⁰ Die ungenügenden Angaben von Einzelheiten erstrecken sich sogar auf die Beschreibung der untersuchten Therapien. So wurde beispielsweise in einer Studie, in der nachgewiesen wurde, dass das Aushändigen einer bestimmten Broschüre Patienten mit Reizdarmsyndrom (im Vergleich zu Patienten ohne diese Broschüre) geholfen hat, nicht beschrieben, worum es in der Broschüre inhaltlich ging oder wie man sie sich beschaffen kann; diese «Behandlung» konnte deshalb nicht von anderen Patienten oder Ärzten angewendet werden. Dies ist nur ein Beispiel von vielen. Insgesamt wurde in dieser Untersuchung festgestellt, dass etwa ein Drittel der Studien keine Angaben zu derart wichtigen Einzelheiten enthielt.¹¹

Und schließlich versäumen es die meisten veröffentlichten Studien, ihre Ergebnisse in den Kontext ähnlicher, früher durchgeführter Studien zu stellen. Ohne diesen wichtigen Schritt kann man, wie wir in Kapitel 8 dargelegt haben, unmöglich sagen, was

die Ergebnisse tatsächlich bedeuten. Alle vier Jahre durchgeführte Überprüfungen von randomisierten Studien, über die im Laufe von 12 Jahren (von 1997 bis 2009) in fünf größeren medizinischen Fachzeitschriften berichtet worden war, lassen das wahre Ausmaß dieses Problems erkennen. Insgesamt wurde überhaupt nur in 25 von 94 (27%) Studien auf systematische Übersichtsarbeiten zu ähnlichen Studien Bezug genommen. Tatsächlich berücksichtigten lediglich drei von 94 Studienberichten aktualisierte Übersichtsarbeiten, die neue Erkenntnisse beinhalteten und so deutlich machten, inwiefern sich die neuen Ergebnisse von der Gesamtheit der vorhandenen Evidenz unterscheiden. Leider gibt es auch keine Anzeichen dafür, dass sich die Berichterstattung über die Jahre verbessert hätte.¹² Dieses Versäumnis kann dazu führen, dass Ärzte unterschiedliche Therapien anwenden, je nachdem, welche Fachzeitschriften sie zufällig lesen.

Entwurf für eine bessere Zukunft

Sicherlich *könnte* man medizinische Forschung aus den richtigen Gründen betreiben, und auch die Berichterstattung darüber *könnte* angemessen sein. Für sich allein genommen ist keiner der nachfolgenden Vorschläge wirklich neu. Doch in ihrer Gesamtheit und mit Unterstützung von Patienten und Ärzten stellt unser Acht-Punkte-Plan eine Blaupause für eine bessere Zukunft im Hinblick auf die Überprüfung und Anwendung von Therapien dar.

1. Das Allgemeinwissen über die Beurteilung der Glaubwürdigkeit von Therapieeffekten verbessern

Eine Bedingung für den Wandel besteht in einem stärkeren Bewusstsein der Öffentlichkeit dafür, inwieweit die Belege für die Effekte von Therapien durch systematische Fehler und Zufallsfaktoren ernsthaft verzerrt sein können. Eines der wichtigsten Kennzeichen wissenschaftlicher Untersuchungen – und zwar die Erkennung und Minimierung von systematischen Fehlern – gehört heutzutage nicht wirklich zum «Allgemeinwissen». Daher bedarf es

größerer Anstrengungen, um diese wichtigen Verständnislücken zu füllen und diese Begriffe in den allgemeinen Bildungskatalog aufzunehmen, und zwar schon in der Schule.

2. Die Möglichkeiten zur Erstellung, Pflege und Verbreitung von systematischen Übersichtsarbeiten über die wissenschaftliche Evidenz zu Therapieeffekten verbessern

Viele der Fragen zu den Wirkungen von Therapien lassen sich problemlos beantworten, wenn man die vorhandene Evidenz in systematischen Übersichtsarbeiten zusammenfasst, diese Übersichtsarbeiten aktualisiert und die Ergebnisse effizient unter Ärzten und Patienten verbreitet. Doch es wird noch lange dauern, bis die Erkenntnisse aus der vorhandenen Evidenz vollständig in Form von systematischen Übersichtsarbeiten erfasst sind. Diesem Mangel abzuhelpfen, sollte zu den vorrangigen Zielen eines Gesundheitssystems gehören, damit zuverlässige Informationen über die Wirkungen medizinischer Behandlungen zusammengeführt und zugänglich gemacht werden.

3. An die Ehrlichkeit appellieren, wenn über die Wirkungen einer Therapie Unsicherheit herrscht

Oftmals fällt es Ärzten schwer, Unsicherheiten zuzugeben, und manchmal wird dies von den Patienten auch gar nicht so gern gesehen. Deshalb wird ihnen gelegentlich ein falsches Gefühl von Sicherheit vermittelt, und sie werden nicht darüber aufgeklärt, welche Unsicherheiten hinsichtlich der Evidenz bestehen. Wenn Ärzte und Patienten zugunsten einer effizienteren Bewertung von Therapieeffekten erfolgreich zusammenarbeiten sollen, dann müssen beide Parteien bereit sein anzuerkennen, dass unzureichend bewertete Therapien erheblichen Schaden anrichten können, und sie müssen sich mehr damit befassen, mit welchen Methoden man verlässliche Evidenz gewinnen kann. Wir müssen herausfinden, wie sich das am besten bewerkstelligen lässt.

4. Identifizierung und Priorisierung von Forschungsvorhaben, in denen von Patienten und Ärzten als relevant erachtete Fragen untersucht werden

Die Konzepte von Forschungsförderungs- sowie akademischen Einrichtungen sind zum Einen von Grundlagenforschung beherrscht, von der die Patienten in absehbarer Zukunft aber eher nicht profitieren werden, zum anderen von Forschungsprojekten, die Gewinnmaximierung für die Industrie zum Ziel haben. Angewandte Forschung zu Fragen, die kein finanzielles Potenzial bergen, aber auf die es den Patienten ankommt, muss um Fördermittel kämpfen, selbst wenn ihre Durchführung öffentlich unterstützt wird. Wir sollten dafür sorgen, dass mehr getan wird, um herauszufinden, welche Fragen Patienten und Ärzte bezüglich der Wirkungen von Therapien haben, und dass Forschungsförderungseinrichtungen diese Fragen bei der Priorisierung von Forschungsvorhaben berücksichtigen, um auf diese Weise die bestehenden Unsicherheiten abzubauen.

5. Sich gegen Doppelstandards bei der Einwilligung in eine Therapie zur Wehr setzen

Ärzte, die bereit sind, Unsicherheiten bezüglich der Wirkungen von Therapien zuzugeben, und sich dieses Problems in formalen Therapievergleichen anzunehmen, unterliegen im Umgang mit ihren Patienten strengeren Regeln als ihre Kollegen, die das nicht tun. Dieser übertriebene Doppelstandard ist unlogisch und unhaltbar. Wenn Unsicherheit über Therapieeffekte herrscht, dann sollte die Teilnahme an randomisierten Studien oder anderen Formen einer unvoreingenommenen Beurteilung die Norm sein. Wir sollten dafür Sorge tragen, dass sich eine Teilnahme an Forschungsarbeiten über die Wirkungen von Therapien nicht als ein notwendigerweise riskantes Unterfangen darstellt, denn das würde bedeuten, dass das «Standard»-Vorgehen immer wirksam und sicher ist.

6. Gegen Ineffizienz in der Forschungsgemeinschaft vorgehen

Nicht selten wird erstaunt festgestellt, dass von Wissenschaftlern keine systematische Auswertung der vorhandenen Evidenz verlangt

wird, wenn sie sich um die Finanzierung neuer Forschungsprojekte und die Zustimmung von Ethik-Kommissionen bemühen. Denn dies hat unweigerlich zur Folge, dass schlecht geplante und, offen gesagt, auch unnötige Forschung immer noch in einem Umfang durchgeführt wird, der sowohl aus ethischen als auch wissenschaftlichen Gründen nicht hinnehmbar ist. Wir sollten Druck auf Forschungsförderungseinrichtungen und Ethik-Kommissionen ausüben, damit sie sicherstellen, dass Forscher keine neuen Forschungsprojekte in Angriff nehmen, ohne auf die systematischen Übersichtsarbeiten zur vorhandenen relevanten Evidenz einzugehen. Berichte über neue Forschungsarbeiten sollten mit der Bezugnahme auf systematische Übersichtsarbeiten eingeleitet werden, um darzulegen, warum wir das neue Forschungsprojekt brauchen, und auch damit abschließen, um aufzuzeigen, worin sich die neuen Ergebnisse von der Gesamtheit der vorhandenen Evidenz unterscheiden.

7. Zu Verzerrungen führende Publikationspraktiken ächten

Um zu Verzerrungen führende Publikationspraktiken auszumerzen, müssen sowohl zu Beginn als auch bei der Beendigung einer Studie entsprechende Maßnahmen ergriffen werden: Dazu gehört, dass Studien bei Beginn registriert und die Studienprotokolle zur Kontrolle öffentlich zugänglich gemacht werden. Nach Studienabschluss sollten die Ergebnisse aller Studien veröffentlicht und die Rohdaten zur Kontrolle und zur weiteren Auswertung zugänglich gemacht werden.

8. Transparenz der Informationen über wirtschaftliche Interessen und andere Interessenkonflikte fordern

Mittlerweile liegen solide Hinweise darauf vor, dass bei der Planung, Durchführung, Auswertung, Interpretation und Anwendung von Forschungsergebnissen eigennützige finanzielle und andere Interessen manchmal über die Interessen der Patienten gestellt werden. Dies gefährdet das gegenseitige Vertrauen. Dieses Vertrauen ist aber nötig, um zu gewährleisten, dass Forschung sich stärker an Patien-

teninteressen ausrichtet. Deshalb sollte man verlangen, dass alle Beteiligten – angefangen bei den Wirtschaftsunternehmen bis hin zu den Interessenvertretungen der Patienten – ihre über das Wohlergehen der Patienten hinausgehenden persönlichen Interessen offenlegen müssen.

Es ist Zeit zu handeln

Hinsichtlich der Überprüfung von Therapien ist eine völlige Neuausrichtung längst überfällig. Für ein Aktionsbündnis von Ärzten und Patienten sind die im Folgenden empfohlenen Schritte besonders zweckmäßig. Die Zeit ist gekommen, dass Sie, die Leser, sich für einen Wandel einsetzen.

Ein Aktionsplan: Was Sie selber tun können

- Identifizieren Sie Fragen zu Therapieeffekten, die Ihnen wichtig sind.
- Lernen Sie Unsicherheiten erkennen. Sagen Sie Ihre Meinung. Stellen Sie Fragen, und verlangen Sie ehrliche Antworten.
- Scheuen Sie sich nicht, Ihren Arzt danach zu fragen, welche Behandlungsmöglichkeiten es gibt, was passieren könnte, wenn Sie sich für eine bestimmte Therapie entscheiden, UND was passieren könnte, wenn Sie das nicht tun.
- Wenn Sie über verschiedene Behandlungsmöglichkeiten nachdenken, könnten Ihnen die Informationen über Entscheidungshilfen unter www.patient-als-partner.de helfen. Siehe auch «Zusätzliche Quellen»
- Nutzen Sie verlässliche Websites (wie z.B. www.gesundheitsinformation.de). Siehe dazu Kapitel 12 und den Abschnitt «Zusätzliche Quellen».

- Bewahren Sie sich eine gesunde Skepsis gegenüber unbegründeten Behauptungen und Medienberichten über «therapeutische Durchbrüche» und gegenüber der Art und Weise, wie die Medien mit «Zahlen» umgehen – vor allem mit großen Zahlen in den Schlagzeilen!
- Misstrauen Sie Therapien, die Ihnen oder Ihren Angehörigen angeboten werden und die auf Überzeugungen und Dogmen gründen, aber nicht durch verlässliche Evidenz untermauert sind.
- Hüten Sie sich vor unnötiger «Krankheitsetikettierung» und übermäßiger Diagnostik (s. Kap. 2 und 4). Finden Sie heraus, ob die fragliche Krankheit für Sie mit einem hohen oder niedrigen Risiko verbunden ist. Fragen Sie nach, was passieren würde, wenn eine Behandlung nicht sofort eingeleitet würde.
- Erklären Sie sich zur Teilnahme an einer klinischen Studie nur unter der Bedingung bereit, dass 1. die Studie registriert und das Studienprotokoll öffentlich zugänglich gemacht wird, 2. im Studienprotokoll Bezug auf systematische Übersichtsarbeiten zur vorhandenen Evidenz genommen und dadurch nachgewiesen wird, dass die Durchführung der Studie gerechtfertigt ist, und 3. Ihnen schriftlich versichert wird, dass die vollständigen Studienergebnisse veröffentlicht und an alle Teilnehmer verschickt werden, die den entsprechenden Wunsch äußern.
- Unterstützen Sie Ärzte, Wissenschaftler, Forschungssponsoren und andere, die bemüht sind, Forschungsprojekte zu fördern, die sich mit unzureichend beantworteten Fragen zu Therapieeffekten, die Sie selbst als wichtig erachten, auseinandersetzen, und arbeiten Sie mit ihnen zusammen.
- Treten Sie für eine breitere Aufklärung über die Auswirkungen von systematischen Fehlern (Bias) und Zufallsfaktoren ein, und nehmen Sie über Ihren gewählten politischen Vertreter und andere Einfluss, damit dies – und zwar schon ab der Grundschule – in die schulische Bildung aufgenommen wird.

Literatur

Vorwort von Ben Goldacre

1. Lexchin J, Bero LA, Djulbegovic B, *et al.* Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003; 326: 1167–1170.
2. Schwitzer G, Ganapati M, Henry D, *et al.* What are the roles and responsibilities of the media in disseminating health information? *PLoS Med* 2(7): e215.
3. Wilson PM, Booth AM, Eastwood A, *et al.* Deconstructing media coverage of trastuzumab (Herceptin): an analysis of national newspaper coverage. *Journal of the Royal Society of Medicine* 2008; 101: 125–132.
4. Shang A, Huwiler-Müntener K, Nartey L, *et al.* Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. *Lancet* 2005; 366: 726–732.
15. Bjelakovic G, Nikolova D, Gluud LL, *et al.* Antioxidant supplements for prevention of mortality in healthy participants and patients with various diseases. *Cochrane Database of Systematic Reviews* 2008, Issue 2. Art. No.: CD007176.

Einleitung

1. Nolte E, McKee CM. Does health care save lives? Avoidable mortality revisited. London: Nuffield Trust, 2004.
2. Nolte E, McKee CM. Measuring the health of nations: updating an earlier analysis. *Health Affairs* 2008; 27(1): 58–77.
3. Gigerenzer G. Reckoning with risk. London: Penguin Books, 2003. Darin zitiert: Franklin B. Letter to Jean Baptiste Le Roy, 13 November 1789. Writings, vol x.
4. Goldacre B. Bad Science. London: Fourth Estate, 2008, px.
5. Matthews A, Dowswell T, Haas DM, *et al.* Interventions for nausea and vomiting in early pregnancy. *Cochrane Database of Systematic Reviews* 2010, Issue 9. Art. No.: CD007575.
6. Irwig L, Irwig J, Trevena L, *et al.* Smart health choices: making sense of health advice. London: Hammersmith Press, 2008. pdf kostenlos zugänglich unter: www.health.usyd.edu.au/shdg/resources/smart_health_choices.php and from www.jameslindlibrary.org.
7. Woloshin S, Schwartz LM, Welch HG. Know your chances: understanding health statistics. Berkeley: University of California Press, 2008. pdf kostenlos zugänglich unter: www.jameslindlibrary.org.

Kapitel 1

1. Vandenbroucke JP. Thalidomide: an unanticipated adverse event. 2003. Zugänglich unter: www.jameslindlibrary.org.
2. Stephens T, Brynner R. *Dark medicine: the impact of thalidomide and its revival as a vital medicine*. Cambridge, Mass: Perseus Publishing, 2001.
3. Thomson D, Capstick T. How a risk management programme can ensure safety in thalidomide use. *Pharmaceutical Journal* 2004 Feb 14: 194–195.
4. Krumholz HM, Ross JR, Presler AH, *et al*. What have we learnt from Vioxx? *BMJ* 2007; 334: 120–123.
5. Erklärungen der Fa. Merck vom 7. Dezember 2009 zum Vioxx-Gerichtsverfahren um in Ontario, Kanada, sowie vom 4. März 2010 zum Vioxx-Urteil in Australien, zugänglich unter www.merck.com.
6. Cohen D. Rosiglitazone what went wrong? *BMJ* 2010; 341: c4848.
7. Lehman R, Yudkin JS, Krumholz HM. Licensing drugs for diabetes: surrogate end points are not enough, robust evidence of benefits and harms is needed. *BMJ* 2010; 341: c4805.
8. Blackstone EH. Could it happen again? The Björk-Shiley convexo-concave heart valve story. *Circulation* 2005; 111: 2717–2719.
9. Wilson PM, Booth AM, Eastwood A, *et al*. Deconstructing media coverage of trastuzumab (Herceptin): an analysis of national newspaper coverage. *Journal of the Royal Society of Medicine* 2008; 101: 125–132.
10. Timmins N. Drugs watchdog gets harsh treatment. *Financial Times*, 8. Oktober 2005, S. 6.
11. Hawkes N. Wonder drug is «cure» for cancer, say doctors. *The Times*, 20. Oktober 2005.
12. Press MF, Sauter G, Bernstein L, *et al*. Diagnostic evaluation of HER-2 as a molecular target: an assessment of accuracy and reproducibility of laboratory testing in large, prospective, randomized clinical trials. *Clinical Cancer Research* 2005; 11(18): 6598–6607.
13. NICE draft guidance on trastuzumab (Herceptin) for early breast cancer (Pressemitteilung), 9. Juni 2006. www.nice.org.uk/page.aspx?o=328789.
14. Cumming J, Mays N, Daubé J. How New Zealand has contained expenditure on drugs. *BMJ* 2010; 340: 1224–1226.
15. NHS NICE Technology Appraisal TA34. *Guidance on the use of trastuzumab for the treatment of advanced breast cancer*. Erscheinungsdatum März 2002; Revisionsdatum April 2005. www.nice.org.uk/TA34.

Kapitel 2

1. Gilbert R, Salanti G, Harden M, *et al*. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of clinicians' recommendations from 1940–2000. *International Journal of Epidemiology* 2005; 34: 74–87.
2. Furberg CD. Effect of antiarrhythmic drugs on mortality after myocardial infarction. *American Journal of Cardiology* 1983; 52: 32C–36C.

3. Chalmers I. In the dark. Drug companies should be forced to publish all the results of clinical trials. How else can we know the truth about their products? *New Scientist* 2004, 6. März, S. 19. Darin zitiert: Moore T, *Deadly Medicine*. New York: Simon and Schuster, 1995.
4. Cowley AJ, Skene A, Stainer K, *et al.* The effect of lorcanide on arrhythmias and survival in patients with acute myocardial infarction: an example of publication bias. *International Journal of Cardiology* 1993; 40: 161–166.
5. Chalmers I. Evaluating the effects of care during pregnancy and childbirth. In: Chalmers I, Enkin M, Keirse MJNC, eds. *Effective care in pregnancy and childbirth*. Oxford: Oxford University Press, 1989: 3–38.
6. Ulfelder H. The stilbestrol disorders in historical perspective. *Cancer* 1980; 45: 3008–3011.
7. Office of Technology Assessment. *Identifying health technologies that work: searching for evidence*. Washington, DC: US Government Printing Office, 1994.
8. Hemminki E, McPherson K. Impact of postmenopausal hormone therapy on cardiovascular events and cancer: pooled data from clinical trials. *BMJ* 1997; 315: 149–153.
9. Anonymous. HRT: update on the risk of breast cancer and long-term safety. *Current Problems in Pharmacovigilance* 2003; 29: 1–3. Zitiert werden darin die Ergebnisse der randomisierten kontrollierten Studie der Women's Health Initiative (*JAMA* 2003; 289: 3243–3253) sowie der Million Women Study (*Lancet* 2003; 362: 419–427).
10. Roberts H. Hormone replacement therapy comes full circle. *BMJ* 2007; 335: 219–220.
11. Williams HC. Evening primrose oil for atopic dermatitis: time to say goodnight (editorial). *BMJ* 2003; 327: 1358–1359.
12. Hoare C, Li Wan Po A, Williams H. Systematic review of treatment for atopic eczema. *Health Technology Assessment* 2000; 4(37): 1–191.
13. Takwale A, Tan E, Agarwal S, *et al.* Efficacy and tolerability of borage oil in adults and children with atopic eczema: randomised, double blind, placebo controlled, parallel group trial. *BMJ* 2003; 327: 1385–1387.

Kapitel 3

1. Crile G. A plea against blind fear of cancer. *Life*, 31. Oktober 1955, S. 128–132.
2. Baum M, Houghton J. Contribution of randomised controlled trials to understanding and management of early breast cancer. *BMJ* 1999; 319: 568–571.
3. Veronesi U, Cascinelli N, Mariani L, *et al.* Twenty-year follow up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *New England Journal of Medicine* 2002; 347: 1227–1232.
4. Baum M. *Breast beating: a personal odyssey in the quest for an understanding of breast cancer, the meaning of life and other easy questions*. London: Anshan, 2010.
5. Japanese Breast Cancer Society. Results of questionnaires concerning breast cancer surgery in Japan 1980–2003. *Breast Cancer* 2005; 12(1).
6. Early Breast Cancer Trialists' Collaborative Group. Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer. An overview

- of 61 randomized trials among 28,896 women. *New England Journal of Medicine* 1988; 319: 1681–1692.
7. Website von Clinical Trial Service Unit: www.ctsu.ox.ac.uk.
 8. Website der Cochrane Collaboration: www.cochrane.org.
 9. Kolata G, Eichenwald K. Health business thrives on unproven treatment, leaving science behind. Sonderbericht in der *New York Times*, 2. Oktober 1999.
 10. Farquhar C, Marjoribanks J, Bassler R, *et al.* High dose chemotherapy and autologous bone marrow or stem cell transplantation versus conventional chemotherapy for women with early poor prognosis breast cancer. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Art. No.: CD003139.
 11. Farquhar C, Marjoribanks J, Bassler R, *et al.* High dose chemotherapy and autologous bone marrow or stem cell transplantation versus conventional chemotherapy for women with metastatic breast cancer. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Art. No.: CD003142.
 12. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, *et al.* Trastuzumab after adjuvant chemotherapy in HER-2-positive breast cancer. *New England Journal of Medicine* 2005; 353: 1659–1672.
 13. Romond EH, Perez EA, Bryant J, *et al.* Trastuzumab plus adjuvant chemotherapy for operable HER-2-positive breast cancer. *New England Journal of Medicine* 2005; 353: 1673–1684.
 14. Carlson GW, Woods WC. Management of axillary lymph node metastasis in breast cancer: making progress. *JAMA* 2011; 305: 606–607.

Kapitel 4

1. Raffle A, Gray M. *Screening: evidence and practice*. Oxford: Oxford University Press, überarb. Nachdruck, 2009.
2. Sense About Science. *Making sense of screening*. London: Sense About Science, 2009. Zugänglich unter: www.senseaboutscience.org.
3. Goodman MT, Gurney JG, Smith MA, *et al.* Sympathetic nervous system tumors. In: Ries LAG, Smith MA, Gurney JG, *et al.* (eds). *Cancer incidence and survival among children and adolescents: United States SEER Program 1975–1995*. National Cancer Institute, SEER Program. NIH Pub. No.994649. Bethesda, MD, 1999. SEER Pediatric Monograph. Zugänglich unter <http://seer.cancer.gov/publications/childhood>.
4. Mullassery D, Dominici C, Jesudason EC, *et al.* Neuroblastoma: contemporary management. *Archives of Disease in Childhood – Education and Practice* 2009; 94: 177–185.
5. Morris JK. Screening for neuroblastoma in children. *Journal of Medical Screening* 2002; 9: 56.
6. Raffle A, Gray M. *op. cit.*, S. 89–92.
7. Welch HG. *Should I be tested for cancer? Maybe not and here's why*. Berkeley and Los Angeles: University of California Press, 2004, S. 77.
8. Cosford PA, Leng GC, Thomas J. Screening for abdominal aortic aneurysm. *Cochrane Database of Systematic Reviews*, 2007, Issue 2, Art. No.: CD002945.

9. Welch HG. Screening mammography – a long run for a short slide? *New England Journal of Medicine* 2010; 363: 1276–1278.
10. Heath I. It is not wrong to say no. Why are women told only the benefits of breast screening and none of the possible harms? *BMJ* 2009; 338: 1534.
11. Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews* 2011, Issue 1. Art. No.: CD001877.
12. Kösters JP, Gøtzsche PC. Regular self-examination or clinical examination for early detection of breast cancer. *Cochrane Database of Systematic Reviews* 2003, Issue 2. Art. No.: CD003373. (keine Änderungen, Update, Issue 3, Juli 2008.)
13. McPherson K. Should we screen for breast cancer? *BMJ* 2010;340:c3106.
14. Cancer Research UK. Prostate cancer – UK incidence statistics. Aktualisiert am 23. Dezember 2010. <http://info.cancerresearchuk.org/cancerstats/types/prostate/incidence>.
15. Chapman S, Barratt A, Stockler M. *Let sleeping dogs lie? What men should know before getting tested for prostate cancer*. Sydney: Sydney University Press, 2010. pdf zugänglich unter: <http://ses.library.usyd.edu.au/bitstream/2123/6835/3/Let-sleeping-dogs-lie.pdf>.
16. Holmström B, Johansson M, Bergh A, et al. Prostate specific antigen for early detection of prostate cancer: longitudinal study. *BMJ* 2009; 339: b3537.
17. Djulbegovic M, Beyth RJ, Neuberger MM, et al. Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2010; 341: c4543.
18. Stark JR, Mucci L, Rothman KJ, et al. Prostate cancer screening: the controversy continues. *BMJ* 2009; 339: b3601.
19. National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine* 2011; 365: 395–409.
20. Moynihan R. Beware the fortune tellers peddling genetic tests. *BMJ* 2010; 341: c7233.
21. Thornton H. The screening debates: time for a broader approach? *European Journal of Cancer* 2003; 39: 1807–1809.
22. Adapted from Wilson JMG, Jungner G. *Principles and practice of screening for disease*. Public health paper no 34. Geneva: World Health Organization, 1968.
23. COMARE 12th Report: The impact of personally initiated X-ray computed tomography scanning for the health assessment of asymptomatic individuals. Pressemitteilung vom 19. Dezember 2007. www.comare.org.uk/12thReport-PressRelease.htm.
24. Department of Health. Better protection for patients having «MOT» scans: http://webarchive.nationalarchives.gov.uk/+www.dh.gov.uk/en/MediaCentre/Pressreleasesarchive/DH_115243.
25. Food and Drug Administration. Radiation-emitting products: Computed tomography and full-body CT scans – what you need to know: www.fda.gov/radiation-emittingProducts.

Kapitel 5

1. Cabello JB, Burls A, Emparanza JI, *et al.* Oxygen therapy for acute myocardial infarction. *Cochrane Database of Systematic Reviews* 2010, Issue 6. Art No.: CD007160.
2. Glasziou P, Chalmers I, Rawlins M, *et al.* When are randomised trials unnecessary? Picking signal from noise. *BMJ* 2007; 334: 349–351.
3. Goh CL. Flashlamp-pumped pulsed dye laser (585nm) for the treatment of portwine stains: a study of treatment outcome in 94 Asian patients in Singapore. *Singapore Medical Journal* 2000; 41: 24–28.
4. Druker BJ, Talpaz M, Resta DJ, *et al.* Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *New England Journal of Medicine* 2001; 344: 1031–1037.
5. Goldman J. for the British Committee for Standards in Haematology. *Recommendations for the management of BCR-ABL-positive chronic myeloid leukaemia*. London: BSH, 2007.
6. Purohit N, Ray S, Wilson T, *et al.* The parent's kiss: an effective way to remove paediatric nasal foreign bodies. *Annals of the Royal College of Surgeons of England* 2008; 90: 420–422.
7. Sanghavi DM. How should we tell the stories of our medical miracles? *Lancet* 2010; 375: 2068–2069.
8. Léauté-Labrèze C, Dumas la Roque E, Hubische T, *et al.* Propranolol for severe hemangiomas of infancy. *New England Journal of Medicine* 2008; 358: 2649–2651.
9. Huikeshoven M, Koster PHL, de Borgie CAJM, *et al.* Re-darkening of port-wine stains 10 years after pulsed-dye-laser treatment. *New England Journal of Medicine* 2007; 356: 1235–1240.
10. Waner M. Recent developments in lasers and the treatment of birthmarks. *Archives of Disease in Childhood* 2003; 88: 372–374.
11. Anti-Thrombotic Trialists' (ATT) Collaboration. Clinical Trial Service Unit Website: www.ctsu.ox.ac.uk/projects/att.
12. Lin CWC, Moseley AM, Refshauge KM. Rehabilitation for ankle fractures in adults. *Cochrane Database of Systematic Reviews* 2008, Issue 3. Art. No.: CD005595.
13. Lindley RI. Persönliche Mitteilung, 2005.
14. Wardlaw JM, Murray V, Berge E, *et al.* Thrombolysis for acute ischaemic stroke. *Cochrane Database of Systematic Reviews* 2009, Issue 4. Art. No.: CD000213.
15. Schmidt B, Roberts RS, Davis P, *et al.* for the Caffeine for Apnea of Prematurity Trial Group. Long-term effects of caffeine therapy for apnea of prematurity. *New England Journal of Medicine* 2007; 357: 1893–1902.
16. Koffeinzitrat (Kommentar) in *Neonatal Formulary* 5. Zugänglich unter: www.blackwellpublishing.com/medicine/bmj/nnf5/pdfs/comment/caffeine1.pdf.
17. Kenyon S, Pike K, Jones DR, *et al.* Childhood outcomes after prescription of antibiotics to pregnant women with preterm rupture of the membranes: 7-year follow-up of the ORACLE I trial. *Lancet* 2008; 372: 1310–1318.
18. Kenyon S, Pike K, Jones DR, *et al.* Childhood outcomes after prescription of antibiotics to pregnant women with spontaneous preterm labour: 7-year follow-up of the ORACLE II trial. *Lancet* 2008; 372: 1319–1327.

19. Erythromycin (Kommentar) in: *Neonatal Formulary 5*. Zugänglich unter: www.blackwellpublishing.com/medicine/bmj/nnf5/pdfs/commentary/erythromycin.pdf.
20. Giuliano AE, Hunt KK, Ballman KV, *et al.* Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial. *JAMA* 2011; 305: 569–575.
21. General Medical Council. *Good Medical Practice*. London: GMC, 2006, S. 13.
22. Ashcroft R. Giving medicine a fair trial. *BMJ* 2000; 320: 1686.
23. Pritchard-Jones K, Dixon-Woods M, Naafs-Wilstra M, *et al.* Improving recruitment to clinical trials for cancer in childhood. *Lancet Oncology* 2008; 9: 392–399.
24. EQUATOR Network Resource Centre, Leitlinien für gute Forschungsberichte zu medizinischen Studien: www.equator-network.org.
25. Smithells RW. Iatrogenic hazards and their effects. *Postgraduate Medical Journal* 1975; 15: 39–52.

Kapitel 6

1. Hopkins WA. Patulin in the common cold. IV. Biological properties: extended trial in the common cold. *Lancet* 1943; ii: 631–635.
2. Sanders TAB, Woolfe R, Rantzen E. Controlled evaluation of slimming diets: use of television for recruitment. *Lancet* 1990; 336: 918–920.
3. Glasziou P, Chalmers I, Rawlins M, *et al.* When are randomised trials unnecessary? Picking signal from noise. *BMJ* 2007; 334: 349–351.
4. Pocock SJ. Randomised clinical trials. *BMJ* 1977; 1: 1661.
5. Balfour TG. Quoted in West C (1854). *Lectures on the Diseases of Infancy and Childhood*. London: Longman, Brown, Green and Longmans, S. 600.
6. King G, Gakidou E, Imai K, *et al.* Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme. *Lancet* 2009; 373: 1447–1454.
7. Peto J, Eden OB, Lilleyman J, *et al.* Improvement in treatments for children with acute lymphoblastic leukaemia: The Medical Research Council UKALL Trials, 1972–84. *Lancet* 1986; i: 408–411.
8. Noseworthy JH, Ebers GC, Vandervoort MK, *et al.* The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994; 44: 16–20.
9. Moseley JB, O'Malley K, Petersen NJ, *et al.* A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *New England Journal of Medicine* 2002; 347: 81–88.
10. Venning GR. Validity of anecdotal reports of suspected adverse drug reactions: the problem of false alarms. *BMJ* 1982; 284: 249–254.
11. McLernon DJ, Bond CM, Hannaford PC, *et al.* on behalf of the Yellow Card Collaborative. Adverse drug reaction reporting in the UK: a retrospective observational comparison of Yellow Card reports submitted by patients and health-care professionals. *Drug Safety* 2010; 33: 775–788.
12. Kocher T. Über Kropfexstirpation und ihre Folgen. *Archiv für Klinische Chirurgie* 1883; 29: 254–337.

13. Silverman WA, Andersen DH, Blanc WA, *et al.* A difference in mortality rate and incidence of kernicterus among premature infants allotted to two prophylactic regimens. *Pediatrics* 1956; 18: 614–625.
14. Zhang J, Ding E, Song Y. Adverse effects of cyclooxygenase 2 inhibitors on renal and arrhythmia events: meta-analysis of randomized trials. *JAMA* 2006; 296: 1619–1621.
15. Vandenbroucke JP, Psaty BM. Benefits and risks of drug treatments: how to combine the best evidence on benefits with the best data about adverse effects. *JAMA* 2008; 300: 2417–2419.
16. Whittington CJ, Kendall T, Fonagy P, *et al.* Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004; 363: 1341–1345.

Kapitel 7

1. Antithrombotic Trialists' (ATT) Collaboration. Aspirin in the primary and secondary prevention of vascular disease: collaborative meta-analysis of individual participant data from randomised trials. *Lancet* 2009; 373: 1849–1860.
2. CRASH trial collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury – outcomes at 6 months. *Lancet* 2005; 365: 1957–1959.
3. CRASH-2 trial collaborators. Effects of tranexamic acid on death, vascular occlusive events, and blood transfusion in trauma patients with significant haemorrhage (CRASH-2): a randomised, placebo-controlled trial. *Lancet* 2010; 376: 23–32.
4. Askie LM, Brocklehurst P, Darlow BA, *et al* and the NeOProm Collaborative Group. NeOProm: Neonatal Oxygenation Prospective Meta-analysis Collaboration study protocol. *BMC Pediatrics* 2011; 11: 6.

Kapitel 8

1. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988; 332: 349–360.
2. Reynolds LA, Tansey EM, eds. *Prenatal corticosteroids for reducing morbidity and mortality after preterm birth*. London: Wellcome Trust Centre for the History of Medicine, 2005.
3. Dickersin K, Chalmers I. Recognising, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the World Health Organisation. James Lind Library 2010 (www.jameslindlibrary.org).
4. Cowley AJ, Skene A, Stainer K, *et al.* The effect of lorainide on arrhythmias and survival in patients with acute myocardial infarction: an example of publication bias. *International Journal of Cardiology* 1993; 40: 161–166.

5. Moore T. *Deadly Medicine*. New York: Simon and Schuster, 1995.
6. Stjernswärd J. Decreased survival related to irradiation postoperatively in early operable breast cancer. *Lancet* 1974; ii: 1285–1286.
7. Stjernswärd J. Meta-analysis as a manifestation of «bondföruft» («peasant sense»). JLL Bulletin: Commentaries on the history of treatment evaluation 2009 (www.jameslindlibrary.org).
8. Fugh-Berman AJ. The haunting of medical journals: how ghostwriting sold «HRT». *PLoS Medicine* 2010; 7(9): e1000335.
9. Whittington CJ, Kendall T, Fonagy P, *et al.* Selective serotonin-reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004; 363: 1341–1345.
10. Spielmans GI, Biehn TL, Sawrey DL. A case study of salami slicing: pooled analysis of duloxetine for depression. *Psychotherapy and Psychosomatics* 2010; 79: 97–106.
11. Antman EM, Lau J, Kupelnick B, *et al.* A comparison of results of metaanalysis of randomized control trials and recommendations of clinical experts. *JAMA* 1992; 268: 240–248.
12. Natanson C, Kern SJ, Lurie P, *et al.* Cell-free hemoglobin-based blood substitutes and risk of myocardial infarction and death: a meta-analysis. *JAMA* 2008; 299(19): 2304–2312.
13. Chalmers I. TGN1412 and *The Lancet's* solicitation of reports of phase 1 trials. *Lancet* 2006; 368:2206–7.
14. Jack A. Call to release human drug trial data. *Financial Times*, 8. August 2006.
15. Kenter MJH, Cohen AF. Establishing risk of human experimentation with drugs: lessons from TGN1412. *Lancet* 2006; 368: 1387–91.
16. McLellan F. 1966 and all that – when is a literature search done? *Lancet* 2001; 358: 646.
17. Horn J, Limburg M. Calcium antagonists for acute ischemic stroke. *Cochrane Database of Systematic Reviews* 2000, Issue 1. Art. No.: CD001928.
18. Horn J, de Haan RJ, Vermeulen M, *et al.* Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. *Stroke* 2001; 32: 2433–2438.
19. O'Collins VE, Macleod MR, Donnan GA, *et al.* 1,026 experimental treatments in acute stroke. *Annals of Neurology* 2006; 59: 467–477.
20. CRASH trial collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury – outcomes at 6 months. *Lancet* 2005; 365: 1957–1959.

Kapitel 9

1. Emanuel EJ, Menikoff J. Reforming the regulations governing research with human subjects. *New England Journal of Medicine* 2011; 10: 1056/NEJMsb1106942.NEJM.org.
2. Chalmers I, Lindley R. Double standards on informed consent to treatment. In: Doyal L, Tobias JS, eds. *Informed consent in medical research*. London: BMJ Books 2001, S. 266–275.

3. Fallowfield L, Jenkins V, Farewell V, *et al.* Efficacy of a Cancer Research UK communicating skills training model for oncologists: a randomised controlled trial. *Lancet* 2002; 359: 650–656.
4. Chalmers I. Regulation of therapeutic research is compromising the interests of patients. *International Journal of Pharmaceutical Medicine* 2007; 21: 395–404.
5. Roberts I, Prieto-Marino D, Shakur H, *et al.* Effect of consent rituals on mortality in emergency care research. *Lancet* 2011; 377: 1071–1072.

Kapitel 10

1. EQUATOR Network Resource Centre, Leitlinien für gute Forschungsberichte zu medizinischen Studien: www.equator-network.org.
2. European Carotid Surgery Trialists' Collaborative Group. Randomised trial of endarterectomy for recently symptomatic carotid stenosis: final results of the MRC European Carotid Surgery Trial (ECST). *Lancet* 1998; 351: 1379–1387.
3. Cina CS, Clase CM, Haynes RB. Carotid endarterectomy for symptomatic carotid stenosis. *The Cochrane Database of Systematic Reviews* 1999, Issue 3. Art. No.: CD001081.
4. The Magpie Trial Collaborative Group. Do women with pre-eclampsia, and their babies, benefit from magnesium sulphate? The Magpie Trial: a randomised, placebo-controlled trial. *Lancet* 2002; 359: 1877–1890.
5. Duley L, Gülmezoglu AM, Henderson-Smart DJ. Magnesium sulphate and other anticonvulsants for women with pre-eclampsia. *Cochrane Database of Systematic Reviews* 2003, Issue 2. Art. No.: CD000025.
6. Global Report. UNAIDS report on the Global AIDS epidemic 2010: www.unaids.org/globalreport/Global_report.htm.
7. Grimwade K, Swingle G. Cotrimoxazole prophylaxis for opportunistic infections in adults with HIV. *Cochrane Database of Systematic Reviews* 2003, Issue 3. Art. No.: CD003108.
8. Chintu C, Bhat GJ, Walker AS, *et al.* Co-trimoxazole as prophylaxis against opportunistic infections in HIV-infected Zambian children (CHAP): a double blind randomised placebo-controlled trial. *Lancet* 2004; 364: 1865–1871.
9. MRC News Release. Antibiotic drug almost halves AIDS-related death in children. London: MRC, 19. November 2004.
10. World Health Organization and UNICEF. *Co-trimoxazole prophylaxis for HIV-exposed and HIV-infected infants and children: practical approaches to implementation and scale up*. WHO and UNICEF, 2009.
11. Soares K, McGrath J, Adams C. Evidence and tardive dyskinesia. *Lancet* 1996; 347: 1696–1697.
12. Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ* 1998; 317: 1181–1184.
13. Howell CJ, Chalmers I. A review of prospectively controlled comparisons of epidural with non-epidural forms of pain relief during labour. *International Journal of Obstetric Anesthesia* 1992; 1: 93–110.
14. Horn J, Limburg M. Calcium antagonists for acute ischemic stroke. *Cochrane Database of Systematic Reviews* 2000, Issue 1. Art No.: CD001928.

15. Horn J, de Haan RJ, Vermeulen RD, Luiten PGM, *et al.* Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. *Stroke* 2001; 32: 2433–2438.
16. Fergusson D, Glass KC, Hutton B, *et al.* Randomized controlled trials of aprotinin in cardiac surgery: using clinical equipoise to stop the bleeding. *Clinical Trials* 2005; 2: 218–232.
17. Tallon D, Chard J, Dieppe P. Relation between agendas of the research community and the research consumer. *Lancet* 2000; 355: 2037–2040.
18. Cream J, Cayton H. New drugs for Alzheimer’s disease – a consumer perspective. *CPD Bulletin Old Age Psychiatry* 2001; 2: 80–82.
19. Cohen CL, D’Onofrio A, Larkin L, *et al.* A comparison of consumer and provider preferences for research on homeless veterans. *Community Mental Health Journal* 1999; 35: 273–279.
20. Griffiths KM, Jorm AF, Christensen H, *et al.* Research priorities in mental health, Part 2: an evaluation of the current research effort against stakeholders’ priorities. *Australian and New Zealand Journal of Psychiatry* 2002; 36: 327–339.
21. Oliver S, Gray J. *A bibliography of research reports about patients’, clinicians’ and researchers’ priorities for new research.* London: James Lind Alliance, Dezember 2006.
22. Chalmers I. Current controlled trials: an opportunity to help improve the quality of clinical research. *Current Controlled Trials in Cardiovascular Medicine* 2000; 1: 3–8. Zugänglich unter: <http://cvm.controlled-trials.com/content/1/1/3>.
23. Editorial. Safeguarding participants in controlled trials. *Lancet* 2000; 355: 1455–1463.
24. Fugh-Berman A. The haunting of medical journals: how ghostwriting sold «HRT». *PLoS Medicine* 2010; 7(9): e10000335.
25. Heimans L, van Hylckama V, Dekker FW. Are claims of advertisements in medical journals supported by RCTs? *Netherlands Journal of Medicine* 2010; 68: 46–49.
26. Lexchin J, Bero LA, Djulbeovic B, *et al.* Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003; 326: 1167–1176.
27. Weatherall D. Academia and industry: increasingly uneasy bedfellows. *Lancet* 2000; 355: 1574.
28. Angell M. Is academic medicine for sale? *New England Journal of Medicine* 2000; 342: 1516–1518.
29. Grant J, Green L, Mason B. From bench to bedside: Comroe and Dripps revisited. HERG Research Report No. 30. Uxbridge, Middlesex: Brunel University Health Economics Research Group, 2003.
30. Pound P, Ebrahim S, Sandercock P, *et al.* Reviewing Animal Trials Systematically (RATS) Group. Where is the evidence that animal research benefits humans? *BMJ* 2004; 328: 514–517.
31. Weatherall D. The quiet art revisited. *Lancet* 2011; 377: 1912–1913.
32. Pirmohamed M. Cited in Mayor S. Fitting the drug to the patient. *BMJ* 2007; 334: 452–453.
33. Editorial. The human genome at ten. *Nature* 2010; 464: 649–650.

34. Mackillop WJ, Palmer MJ, O'Sullivan B, *et al.* Clinical trials in cancer: the role of surrogate patients in defining what constitutes an ethically acceptable clinical experiment. *British Journal of Cancer* 1989; 59: 388–395.
35. The Psoriasis Association: www.psoriasis-association.org.uk.
36. National Psoriasis Association. Statistische Angaben zu Psoriasis: www.psoriasis.org/netcommunity/learn_statistics.
37. Jobling R. Therapeutic research into psoriasis: patients' perspectives, priorities and interests. In: Rawlins M, Littlejohns P, ed. *Delivering quality in the NHS 2005*. Abingdon: Radcliffe Publishing Ltd, S. 53–56.

Kapitel 11

1. Oliver S, Clarke-Jones L, Rees R, *et al.* Involving consumers in research and development agenda setting for the NHS: developing an evidence-based approach. *Health Technology Assessment Report* 2004; 8(15).
2. NIHR Guy's and St. Thomas' and King's College London's Biomedical Research Centre. *Involving users in the research process: a «how to» guide for researchers*. Version 1, April 2010. Zugänglich unter: www.biomedicalresearchcentre.org.
3. Cartwright J, Crowe S. *Patient and public involvement toolkit*. London: Wiley-Blackwell and BMJI Books, 2011.
4. European Science Foundation/EMRC. *Implementation of medical research in clinical practice – a growing challenge*. Strasbourg: ESF, 2011.
5. Hanley B, Truesdale A, King A, *et al.* Involving consumers in designing, conducting, and interpreting randomised controlled trials: questionnaire survey. *BMJ* 2001; 322: 519–523.
6. Koops L, Lindley RI. Thrombolysis for acute ischaemic stroke: consumer involvement in design of new randomised controlled trial. *BMJ* 2002; 325: 415–417.
7. Staley K. *Exploring impact: public involvement in NHS, public health and social care research*. Eastleigh: INVOLVE, 2009. Zugänglich unter: www.invo.org.uk.
8. Petit-Zeman S, Firkins L, Scadding JW. The James Lind Alliance: tackling research mismatches. *Lancet* 2010; 376: 667–669.
9. Patient Partner Project. An EU programme «Identifying the needs for patients partnering in clinical research»: www.patientpartner-europe.eu.
10. Thornton H, Edwards A, Elwyn G. Evolving the multiple roles of «patients» in health-care research: reflections after involvement in a trial of shared decision-making. *Health Expectations* 2003; 6: 189–197.
11. Dixon-Woods M, Agarwal S, Young B *et al.* Integrative approaches to qualitative and quantitative evidence. NHS Health Development Agency, 2004.
12. Kushner R. *Breast cancer: a personal history and an investigative report*. New York: Harcourt Brace Jovanovitch, 1975.
13. Lerner BH. *The breast cancer wars: hope, fear, and the pursuit of a cure in twentieth-century America*. New York: Oxford University Press, 2003.
14. Institute of Medical Ethics Working Party on the ethical implications of AIDS: AIDS, ethics, and clinical trials. *BMJ* 1992; 305: 699–701.

15. Thornton H. The patient's role in research. [Vortrag, gehalten anlässlich der *The Lancet* «Challenge of Breast Cancer» Conference, Brügge, April 1994.] In: Health Committee Third Report. *Breast cancer services. Volume II. Minutes of evidence and appendices*. London: HMSO, July 1995, S. 112–114.
16. Concorde Coordinating Committee. Concorde: MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. *Lancet* 1994; 343: 871–881.
17. Perehudoff K, Alves TL. *Patient and consumer organisations at the European Medicines Agency: financial disclosure and transparency*. Amsterdam: Health Action International, 2010. Zugänglich unter: www.haieurope.org.
18. Herxheimer A. Relationships between the pharmaceutical industry and patients' organisations. *BMJ* 2003; 326: 1208–1210.
19. Consumers' Association. Who's injecting the cash? *Which?* 2003, April, S. 24–25.
20. Koops L, Lindley RI. Thrombolysis for acute ischaemic stroke: consumer involvement in design of new randomised controlled trial. *BMJ* 2002; 325: 415–417.
21. Donovan J, Mills N, Smith M, *et al* for the ProtecT Study Group. Quality improvement report: improving design and conduct of randomised trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study. *BMJ* 2002; 325: 766–770.

Kapitel 12

1. Edwards A, Elwyn G, Atwell C, *et al*. Shared decision making and risk communication in general practice – a study incorporating systematic literature reviews, psychometric evaluation of outcome measures, and quantitative, qualitative and health economic analyses of a cluster randomised trial of professional skill development. Report to Health in Partnership programme, UK Department of Health. Cardiff: Department of General Practice. University of Wales College of Medicine, 2002.
2. Farrell C, ed. *Patient and public involvement in health: The evidence for policy implementation. A summary of the results of the Health in Partnership research programme*. London: Department of Health Publications, April 2004. Zugänglich unter: www.dh.gov.uk/en/Publicationsandstatistics/Publications/Publications-PolicyAndGuidance/DH_4082332.
3. In Anlehnung an: Marshall T. Prevention of cardiovascular disease. Risk and benefit calculator. Zugänglich unter: www.haps.bham.ac.uk/publichealth/cardiovascular/index.shtml.
4. Evans I, Thornton H. Transparency in numbers: the dangers of statistical illiteracy. *Journal of the Royal Society of Medicine* 2009; 102: 354–356.
5. Irwig L, Irwig J, Trevena L, *et al*. *Smart health choices: making sense of health advice*. London: Hammersmith Press, 2008.
6. Salzburg Global Seminar Website: www.salzburgglobal.org.
7. Salzburger Erklärung zur partizipativen Entscheidungsfindung: Salzburg Global Seminar. *BMJ* 2011; 342: d1745. Zugänglich unter: www.bmj.com/content/342/bmj.d1745.full.

Kapitel 13

1. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; 374: 86–89.
2. Roberts I, Yates D, Sandercock P, *et al*; CRASH trial collaborators. Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial. *Lancet* 2004; 364: 1321–1328.
3. Cochrane Injuries Group Albumin Reviewers. Human albumin administration in critically ill patients: systematic review of randomised controlled trials. *BMJ* 1998; 317: 235–240.
4. Finfer S, Bellomo R, Bryce N, *et al* (SAFE Study Investigators). A comparison of albumin and saline for fluid resuscitation in the intensive care unit. *New England Journal of Medicine* 2004; 350: 2247–2256.
5. Edwards A, Elwyn G, Atwell C, *et al*. Shared decision making and risk communication in general practice – a study incorporating systematic literature reviews, psychometric evaluation of outcome measures, and quantitative, qualitative and health economic analyses of a cluster randomised trial of professional skill development. Report to Health in Partnership programme, UK Department of Health. Cardiff: Department of General Practice. University of Wales College of Medicine, 2002.
6. Farrell C, ed. *Patient and public involvement in health: The evidence for policy implementation. A summary of the results of the Health in Partnership research programme*. London: Department of Health Publications, April 2004. Zugänglich unter: www.dh.gov.uk/en/Publicationsandstatistics/Publications/Publications-PolicyAndGuidance/DH_4082332.
7. Donovan J, Mills N, Smith M, *et al* for the ProtecT Study Group. Quality improvement report: improving design and conduct of randomised trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study. *BMJ* 2002; 325: 766–770.
8. Dickersin K, Chalmers I. Recognising, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the World Health Organization. James Lind Library, 2010 (www.jameslindlibrary.org).
9. Chan A-W, Hróbjartsson A, Haahr MT, *et al*. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004; 291: 2457–2465.
10. Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005; 365: 1159–1162.
11. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008; 336: 1472–1474.
12. Clarke M, Hopewell S, Chalmers I. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. *Lancet* 2010; 376: 20–21.

Zusätzliche Quellen

Wenn Sie an weiteren allgemeinen Informationen über *Testing Treatments* («Wo ist der Beweis?») interessiert sind

Websites

Testing Treatments Interactive

www.testingtreatments.org

Auf dieser Seite finden Sie eine kostenlos zugängliche elektronische Version der 2. Auflage von *Testing Treatments*, die im Laufe der kommenden Jahre um Übersetzungen und weitere Materialien ergänzt werden soll. Zudem sind auf dieser Seite bereits Übersetzungen der 1. Auflage von *Testing Treatments* in arabischer, chinesischer, deutscher, italienischer, polnischer und spanischer Sprache zugänglich.

James Lind Library

www.jameslindlibrary.org

Cochrane Collaboration

www.cochrane.org

NHS Choices

www.nhs.uk (im Suchfenster *research* eingeben)

UK Clinical Research Collaboration

www.ukcrc.org

Healthtalkonline

www.healthtalkonline.org

US National Cancer Institute

Schulungsmaterial zu klinischen Studien <http://cancertrials.nci.nih.gov/clinicaltrials/learning>

Bücher

Ben Goldacre. *Bad science*. London: Harper Perennial, 2009.

Bengt D Furberg, Curt D Furberg. *Evaluating clinical research: all that glitters is not gold*. 2nd edition. New York: Springer, 2007.

Steven Woloshin, Lisa Schwartz, Gilbert Welch. *Know your chances: understanding health statistics*. Berkeley: University of California Press, 2008. Kostenlos zugänglich unter: www.jameslindlibrary.org.

Les Irwig, Judy Irwig, Lyndal Trevena, Melissa Sweet. *Smart health choices: making sense of health advice*. London: Hammersmith Press, 2008. Kostenlos zugänglich unter: www.jameslindlibrary.org.

Trisha Greenhalgh. *How to read a paper: the basics of evidence-based medicine*. 4th edition. Oxford and London: Wiley-Blackwell and BMJI Books, 2010. Dt.: *Einführung in die Evidence-based Medicine*. Bern: Hans Huber, 2. Aufl. 2003.

H Gilbert Welch, Lisa M. Schwartz, Steven Woloshin. *Overdiagnosed: making people sick in the pursuit of health*. Boston: Beacon Press, 2011.

<http://www.inkanet.de/db/special/trias/index.html>

Wenn Sie Informationen darüber suchen, was man über die Wirkungen von Therapien weiß

Cochrane Library

www.thecochranelibrary.com

NHS Evidence

www.evidence.nhs.uk

Informed Health Online

www.informedhealthonline.org

PubMed Health

www.pubmed.gov/health

Patienteninformationen des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

www.gesundheitsinformation.de

Der Krebsinformationsdienst

www.krebsinformationsdienst.de

Patientenportal von Kassenärztlicher Vereinigung und Bundesärztekammer

www.patienten-information.de

Das Deutsche Cochrane Zentrum

www.cochrane.de

Wenn Sie Informationen darüber suchen, was man nicht über die Wirkungen von Therapien weiß

UK Database of Uncertainties about the Effects of Treatments (UK DUETs)

www.evidence.nhs.uk

Wenn Sie an Informationen über aktuelle Forschungsprojekte interessiert sind, die Unsicherheiten in Bezug auf Therapieeffekte zum Thema haben

WHO International Clinical Trials Registry Platform

www.who.int/trialsearch

US National Institutes of Health Clinical Trials Registry

www.clinicaltrials.gov

EU Clinical Trials Register

<https://www.clinicaltrialsregister.eu>

Australian Cancer Trials

www.australiancancertrials.gov.au

Das Deutsche Register Klinischer Studien

www.drks.de

Wenn auch Sie dazu beitragen möchten, die Relevanz und Qualität der Forschung über Therapieeffekte zu verbessern

James Lind Alliance

www.lindalliance.org

Fördert die Kooperationspartnerschaft zwischen Patienten und Ärzten zum Zwecke der Identifizierung und Priorisierung relevanter Unsicherheiten bezüglich der Wirkungen von Therapien

National Institute for Health Research

NIHR Health Technology Assessment

www.nchta.org

Beteiligt die Nutzer aktiv an allen Phasen ihrer Projekte

NIHR Clinical Research Network Coordinating Centre

www.crncc.nihr.ac.uk/ppi

Setzen sich für die Mitwirkung von Patienten, Betreuungspersonen und Öffentlichkeit an klinischen Studien und ihre aktive Beteiligung als «Forscher» ein

Cochrane Consumer Network

www.consumers.cochrane.org

Unterstützt die Mitwirkung von Patienten an den von der *Cochrane Collaboration* erstellten systematischen Übersichtsarbeiten von Therapiestudien

UK Clinical Research Network

www.ukcrn.org.uk

Deutsches Netzwerk Evidenzbasierte Medizin (Fachbereich Patienteninformation)

www.ebm-netzwerk.de

Wenn Sie an einer Schulung zur Bewertung von Forschungsergebnissen teilnehmen möchten

Critical Appraisal Skills Programme

www.casp-uk.net

Organisiert Workshops und andere Maßnahmen, mit deren Hilfe interessierte Teilnehmer die Fertigkeiten erwerben können, die sie zum Auffinden und Verstehen wissenschaftlicher Evidenz befähigen

US Cochrane Center

Understanding Evidence-based Healthcare: A Foundation for Action

us.cochrane.org/understanding-evidence-based-healthcarefoundation-action

Ein Internetkurs, der Interessierten helfen soll, die grundlegenden Konzepte und Fertigkeiten der evidenzbasierten Gesundheitsversorgung zu verstehen

Kompetenztraining für Patientinnen und Patienten

www.fgz.co.at/Wissen-macht-stark-und-gesund.335.0.html

Hintergrundmaterial der Universität Hamburg, Fachwissenschaft Gesundheit

www.patienteninformation.de

Wenn Sie mehr über partizipative Entscheidungsfindung erfahren möchten

The Foundation for Informed Medical Decision Making

www.informedmedicaldecisions.org

Dartmouth-Hitchcock Medical Center: Center for Shared Decision Making

patients.dartmouth-hitchcock.org/shared_decision_making.html

Salzburger Erklärung

www.bmj.com/content/342/bmj.d1745.fullwww.salzburgglobal.org

Patient als Partner

www.patient-als-partner.de

**Wenn Sie sich für systematische Übersichtsarbeiten
von Tierstudien interessieren**

www.sabre.org.uk

www.camarades.info

Wenn Sie Fachbegriffe nachschlagen wollen

Glossar des Deutschen Netzwerks Evidenzbasierte Medizin e.V.

www.ebm-netzwerk.de/was-ist-ebm/grundbegriffe/glossar

Über die Autoren

Imogen Evans praktizierte und lehrte Medizin in Kanada und Großbritannien, bevor sie sich in der britischen Fachzeitschrift *The Lancet* dem Medizinjournalismus zuwandte. Von 1996 bis 2005 arbeitete sie für das britische *Medical Research Council*, zuletzt auf dem Gebiet der Forschungsethik, und war als Vertreterin der britischen Regierung im *Committee on Bioethics* des Europarates tätig.

Hazel Thornton wurde nach einer Routine-Mammographie zur Teilnahme an einer klinischen Studie eingeladen, an der sie wegen der unzureichenden Patienteninformationen jedoch nicht teilnahm. Diese Erfahrung veranlasste sie aber, sich öffentlich für die Erforschung patientenrelevanter Endpunkte einzusetzen. Zu diesem Thema hat sie zahlreiche Beiträge verfasst und Vorträge gehalten.

Iain Chalmers arbeitete als Arzt in Großbritannien und Palästina, bevor er sich der *Health Services*-Forschung zuwandte und die nationale Einrichtung für perinatale Epidemiologie und später das britische Cochrane Centre leitete. Seit 2003 koordiniert er die *James Lind Initiative*, die vor allem durch eine stärkere Einbindung der Öffentlichkeit mit besser kontrollierten Studien eine bessere Gesundheitsversorgung zu erreichen sucht.

Paul Glasziou ist nicht nur in der medizinischen Forschung tätig, sondern arbeitet in Teilzeit auch als Hausarzt. Wegen der zwischen beiden Bereichen bestehenden Kluft hat er seinen Arbeitsschwerpunkt auf die Identifizierung und Beseitigung von Barrieren gelegt, die der Anwendung hochwertiger Forschungserkenntnisse im Praxisalltag entgegenstehen. Beim *British Medical Journal* (BMJ) war er

Herausgeber des *Journal of Evidence-Based Medicine* und von 2003 bis 2010 Direktor des *Centre for Evidence-Based Medicine* in Oxford. Er hat mehrere Bücher zum Thema evidenzbasierte Praxis verfasst. Derzeit hält er eine Forschungsprofessur des australischen *National Health and Medical Research Council*, die er im Juli 2010 an der *Bond University* angetreten hat.

Sachregister

- Ablin, Richard 78
- Abwarten, beobachtendes 196
- Abweichung von der
 - Behandlungszuteilung 123
- active monitoring 196
- AIDS 170, 189
- , Medikament 190
- Aktionsbündnis Arzt–Patient 231
- Aktionsplan 231
- Albumin 224
- Allgemeinwissen 227
- Allokation 118
- Al-Razi 113
- Analogpräparat 178
- Anämie 90
- Anekdote 33
- Aneurysma 73
- Antibiotikum, bei vorzeitigen
 - Wehen 98
- Antipsychotikum 171
- Aortenaneurysma 85, 205
- Aprotinin 175
- Arrhythmie 46
- Arzneimittelwirkung,
 - unerwünschte 128 ff.
- Arzt, Umgang mit Unsicherheiten 102
- Arzt-Wissenschaftler 188
- Aspirin 55, 137, 211
 - und Herzinfarkt 94
- Atemnotsyndrom bei
 - Frühgeborenen 173
- Atkins, Hedley 61
- Auseinandersetzen mit
 - Unsicherheiten 97
- Avandia 38
- Balfour, Thomas Graham 117
- Bauchaortenaneurysma-Screening 73
- Behandlung, unnötige 216
- Behandlungseffekt, moderater 93
- Behandlungsergebnis, faire
 - Erfassung 126
- Behandlungswirkung, unerwartete 129
- Behandlungszuteilung,
 - Abweichung 123
- Belladonna 118
- Berichterstattung, selektive 225
- Berichtstandard 224
- Bestätigungsbias 109
- Bewusstsein der Öffentlichkeit 227
- Bias 61, 122, 139, 232
 - in systematischen Reviews 145
- Bilirubin 130
- Björk-Shiley-Herzklappe 40
- Blindheit bei Säuglingen 35
- Blutarmut 90
- Bluthochdruck 206
- Blutschwamm 91
- Blutung, intra- und postoperativ 175
- Brustkrebs 41, 192
 - , Chirurgie 56 ff.
 - , Screening 73
 - , Therapie 56 ff., 100
- Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) 128
- Bürgerlobby im
 - Gesundheitswesen 194
- CancerBACKUP 189
- Checkliste 102
- Clement-Jones, Vicky 189
- Clinical Trials Directive 157

- Cochrane Collaboration 186
 Cochrane Database of Systematic
 Reviews 216
 compassionate release 190
 confidence interval (CI) 135
 confirmation bias 109
 Contergan 34
 Cotrimoxazol 170
 Crile, George 60
 Crossover-Studie 120
- Datenmanipulation bei systematischen
 Reviews 148
 Deklaration von Helsinki 157
 Diabetes 90
 Diabetiker, Definition 217
 Diagnostik, übermäßige 232
 Diethylstilbestrol (DES) 47 ff., 111, 130
 Direct-to-Consumer-(DTC-)Gentest 81
 Diuretika 55
 Doppelblindstudie 126
 Doppelstandard 229
 Dosis, richtige 55
 Duchenne-Muskeldystrophie 81
- Ehrlichkeit 228
 Einverständnis zur
 Studienteilnahme 160
 Einwilligungserklärung 161, 196
 – der Teilnehmer 160
 Einzelstudie 141
 Eklampsie 169
 Ekzem 51
 Entfernen von Fremdkörpern aus der
 Nase 91
 Entscheidungsfindung,
 partizipative 200, 217
 Entwurf für eine bessere
 Zukunft 221 ff.
 Epidemie von Blindheit 35
 Epiduralanalgesie bei
 Wehenschmerzen 173
 Erdbeer-Hämangiom 91
 Ethik, Voreingenommenheit 160
 Ethik-Kommission 156
 European Medicines Agency
 (EMA) 39, 193
- Evidenz 96, 133, 199
 –, relevante verlässliche 141 ff.
 –, systematische Prüfung 149
 –, unverzerrte 64
 –, wissenschaftliche 209
 –, zuverlässig? 214
- Fehlbildung bei Säuglingen 34
 Fehler, systematischer 61, 122, 139,
 232
 Fehlgeburt 47
 Fersen-Pricktest 72
 Feuerarm, Laserbehandlung 90
 Food and Drug Administration
 (FDA) 36
 Forschung
 –, bessere 222
 –, Einbindung von Patienten und
 Öffentlichkeit 186
 –, gute 167 ff.
 –, Informationen zusammenfassen 143
 –, klinische 167 ff.
 –, medizinische 155 ff.
 –, Mitwirkung von Patienten 188 ff.
 –, richtige 185 ff.
 –, schlechte 171 ff.
 –, überflüssige 173 ff.
 –, Verbesserung 185 ff.
 –, Verschwendung von
 Ressourcen 152
 Forschungsbericht
 –, Schlüsselfragen 141
 –, unbrauchbarer 222
 –, unverzerrter und brauchbarer 226
 Forschungsergebnis 143
 –, Einordnung 154
 –, selektive Berichterstattung 225
 Forschungsförderung 229
 Forschungsfragen, richtige 222
 Forschungsgemeinschaft,
 Ineffizienz 229
 Forschungspriorität, verkehrte 176 ff.
 Forschungsregulierung 164
 Forschungsvorhaben, richtig planen
 und durchführen 224
 Fragen
 –, für Patienten relevante 176

- zur Umsetzung von wissenschaftlicher Evidenz in die Praxis 204
- Franklins Gesetz 25
- früher = besser? 67 ff.
- Frühgeborenen-Apnoe 98
- Frühgeborenen-Retinopathie 35
- Frühgeborenes
 - , Atemnotsyndrom 173
 - , Koffein-Studie 97, 161
- gamma-linolenic acid (GLA) 51
- Ganzkörper-CT 84 ff.
- Genetik 181
- Gentest,
 - Direct-to-Consumer-(DTC-) 81
- Gesetz der großen Zahlen 133 ff.
- Gespräch Arzt–Patient 202
- Gesundheitsforschung 167 ff.
- Gesundheitsversorgung, bessere 199 ff.
- Gesundheitswesen, Verschwendung von Ressourcen 152
- Ghostwriting wissenschaftlicher Veröffentlichungen 179
- Gleiches mit Gleichem vergleichen 113 ff.
- Good Medical Practice 101
- Heilkraft, natürliche 107
- Herceptin 41 ff., 65, 192
- Herzinfarkt und Aspirin 94
- Herzklappe, mechanische 40
- Herzrhythmusstörung 46
- Hexamethonium 153
- Hill, Austin Bradford 141
- HIV 190
 - , Infektion bei Kindern 170
- Hochschulmedizin – käuflich? 181
- hormone replacement therapy (HRT) 49 ff.
- Hormonersatztherapie (HRT) 49 ff., 116
- Humanalbumin-Lösung 224
- Huntington-Krankheit 81
- Imatinib 91
- Industrie, pharmazeutische 177 ff.
- Ineffizienz in der
 - Forschungsgemeinschaft 229
- Information
 - und Einverständnis 160
 - aus der Forschung zusammenfassen 143
- Informationsquelle, zuverlässige 215
- Informationstransparenz 230
- informed consent 160, 196
- Institutional Review Board (IRB) 156
- Intensivtherapie, Brustkrebs 56 ff.
- Intention-to-Treat-Analyse 122
- Interferon 191
- James Lind Alliance 186
- Kampf um unverzerrte Evidenz 64
- Kniegelenkarthrose 201
- Knochenmarktransplantation 62
- Kocher, Theodor 130
- Koffein-Studie an Frühgeborenen 97, 161
- Konfidenzintervall 135
- Krankheit
 - , erbliche 212
 - , lebensbedrohliche 204
 - , unbehandelte 108
- Kushner, Rose 62, 189
- Laserbehandlung von Feuermalen 90
- Leadtime-Bias 70
- Length-Time-Bias 69
- Leukämie
 - , lymphatische 124
 - , myeloische 91
- Lind, James 32, 113
- Lorcaïnid 146
- Lumpektomie 59
- Lungenkrebs-Screening 78
- Macmillan Cancer Support 189
- Magnesiumsulfat 169
- Magnetismus, animalischer 126
- Magpie-Studie 169
- Mammakarzinom 100
- Mammographie-Screening 73
- Maskierung 125

- Mastectomy Association 189
 Mastektomie, radikale 58, 188
 Medicines and Healthcare Products
 Regulatory Agency (MHRA) 129
 Medizin
 –, marketingbasierte 146
 –, personalisierte 211
 Medizinprodukt, klinische
 Prüfung 157
 mehr = besser? 55 ff.
 Mehrfachpublizieren 146
 Mehr-ist-besser-Ansatz 55 ff.
 Menopause 48, 116
 Mesmer, Anton 126
 Mesmerismus 126
 Meta-Analyse 140, 144
 –, kumulative 150
 Me-too-Präparat 178
 Mother's Kiss 91
 multiple Sklerose (MS) 127, 191
- Nachtkerzenöl 51 ff.
 National Cancer Institute 189
 National Health Service (NHS) 41, 111
 National Institute for Health und
 Clinical Excellence (NICE) 42
 Nebenwirkung 129
 Negativbeweis 26
 neu = besser? 31 ff.
 Neuroblastom-Screening 68
 N-gleich-1-Studie 210
 NHS Choices 216
 Nichtpublizieren 145
 Nichtsteroidale Antiphlogistika
 (NSAID) 35
 Nimodipin 174
 n-of-1 trial 210
- Operation, verstümmelnde 57
 Optimismus und Wunschdenken 109
 Outcome 126, 138
 Overreporting 146, 149
- Paroxetin 148
 Patient
 –, relevante Fragen 176
 –, Studienteilnahme 213
 –, vermeidbare Nachteile 149
 – und Wissenschaftler,
 Zusammenarbeit 195 ff.
 Patienten-Einwilligung 162
 Patientengruppe, Vergleich 116
 Patienteninformation 205
 Patientenorganisation 187, 193
 Patientenvertretung 193
 Patient-Wissenschaftler 189
 personalisierte Medizin 211
 Pester-Power 193
 Peto, Richard 62
 Pharmaceutical Management Agency
 (PHARMAC) 42
 Pharmaindustrie 221
 – und Ärzte 179
 –, Einfluss der 177 ff.
 Phenylketonurie (PKU) 72
 Phenylketonurie-Screening 72
 Placebo 110, 125
 plötzlicher Kindstod 45, 131
 Präeklampsie in der
 Schwangerschaft 168
 Propranolol 92
 Prostatakarzinom 196, 208, 225
 –, Screening 74
 –, Überdiagnostizierung 77
 prostataspezifisches Antigen (PSA) 75
 Pseudokarzinom 100
 Psoriasis 183
 Publikationsbias 146
 PubMed Health 216
- Randomisierung 61
 Regulierung von Therapietests 155 ff.
 Regulierungssystem für
 Therapiestudien 158 ff.
 Research Ethics Committee (REC) 156
 Review, systematischer 100, 137,
 143 ff.
 –, Bias 145
 –, Datenmanipulation 148
 –, Einfluss des Zufalls 147
 Revolution, genetische 181
 Rhabdomyolyse 128
 Richtlinien besserer Forschung 222
 Risiko 26

- Risikoreduktion, absolute und relative 208
- Risikoverhältnis 133
- risk ratio 133
- Rofecoxib 36, 131
- Röntgen-Thorax-Screening 79
- Rosuvastatin 128
- Royal College of Physicians 33
- Salamitaktik 149
- Salzburger Erklärung zur partizipativen Entscheidungsfindung 217
- Säugling
- , Blindheit 35
 - , Fehlbildung 34
 - , Schlafposition 45
- Scharlach 118
- Scheinbehandlung 110, 125
- Schizophrenie 172
- Schlafposition beim Säugling 45
- Schlaganfall 122, 152, 167, 174, 195, 206
- , Prävention 85
- Schlüsselfragen beim Forschungsbericht 141
- Screening 67 ff.
- , Sinn und Zweck 82
- Screeningprogramme – sinnvoll? 87
- Screeninguntersuchung, Nutzen und Schaden 72 ff.
- Screening-Zirkus 85
- Selbsthilfegruppe 188
- Seroxat 148
- Shared Decision Making 200, 217
- Signifikanz, statistische 136
- Sildenafil (Viagra) 111
- Skorbut 32
- Spiral-CT 79
- Spock, Benjamin 45, 131
- Spontanremission 108
- Stammzellrettung 62
- Statistik 206
- statistisch signifikant 138
- Störung, psychiatrische 171
- Studie, «negative» 225
- Studie
- , cluster-randomisierte 121
 - , randomisierte 119
 - , randomisierte mit Crossover-Design 120
 - , wissenschaftliche 101
- Studienprotokoll 123, 232
- Studienteilnehmer 124
- , vermeidbare Nachteile 151
- Teilnahme an einer klinischen Studie 232
- Teilnehmerzahl 138
- Test
- , fairer 31 ff., 101, 107 ff., 141, 209
 - , genetischer 81
- Thalidomid 34
- Therapie
- , mit dramatischen Effekten 114
 - , drastische 57
 - , systemische 59
 - , Überprüfung 231
 - , Überprüfung und Nützlichkeit 44
 - , mit wichtigen, aber mäßigen Effekten 115
- Therapieeffekt 231
- , dramatischer 89
 - , Unsicherheiten 89 ff.
- Therapiestudie, Regulierungssysteme 158 ff.
- Therapietest, Regulierung 155 ff.
- Therapievergleiche 113 ff.
- Therapiewirkung, unerwünschte 128 ff.
- Therapiezureitung, unverzerrte prospektive 117 ff.
- Totgeburt 47
- Tranexamsäure 138
- Transparenz der Informationen 230
- Trastuzumab 41, 192
- Typ-2-Diabetes 38
- Überdiagnose-Bias 69
- Überdiagnostizierung 71
- des Prostatakarzinoms 77
- Überprüfung von Therapien 231
- Übersichtsarbeit, systematische 228
- Überwachung, aktive 196
- Underreporting 145

- Unsicherheit 26, 228, 231
 - , Auseinandersetzen mit 97
 - beim Therapieeffekt 89 ff.
 - , Umgang mit 102
- Unterschied, signifikanter 136
- Untersuchung
 - , genetische 211
 - zu unerwünschten
 - Therapiewirkungen 128 ff.
- US Veterans Health
 - Administration 111
- Verbesserung der Forschung 185 ff.
- Verblindung 125
- Versorgung, medizinische, im Rahmen
 - fairer Tests 101
- Vertrauensintervall 135
- Verzerrung 230
- Viagra 111
- Vioxx 35, 131
- Vitamin B₁₂ 90
- Voreingenommenheit der Ethik 160
- Vorlaufzeit-Bias 70
- Wahrscheinlichkeit 26
 - watchful waiting 196
 - Wehen, vorzeitige 98
- Weltärztebund 157
- Wirkung
 - , erhoffte und nicht eingetretene 45 ff.
 - , unerwartete negative 34
- World Medical Association 157
- Yellow-Card-System 129
- Zidovudin 190
- Zufall, Einfluss auf systematische
 - Reviews 147
- Zufallsfaktor 133 ff., 232
- Zufallszuteilung zu
 - Therapieversuchen 120 ff.
- Zulassung, mitfühlende 190
- Zusammenarbeit
 - Patient–Wissenschaftler 195 ff.
- Zuteilung
 - , unverzerrte prospektive 117 ff.
 - , zufällige 120 ff.