

Regina Kunz Khalid S. Khan Jos Kleijnen Gerd Antes

Systematische Übersichtsarbeiten und Meta-Analysen

Einführung in Instrumente
der evidenzbasierten Medizin für Ärzte,
klinische Forscher und Experten
im Gesundheitswesen



HUBER



Kunz / Khan / Kleijnen / Antes
Systematische Übersichten und Meta-Analysen

Verlag Hans Huber
Programmbereich Medizin

Bücher aus verwandten Sachgebieten

Greenhalgh

Einführung in die Evidence-based Medicine

Kritische Beurteilung klinischer Studien als Basis einer rationalen Medizin

2. Aufl. 2003. ISBN 978-3-456-83926-4

Bonita / Beaglehole / Kjellström

Einführung in die Epidemiologie

2. Aufl. 2008. ISBN 978-3-456-84535-7

Fletcher / Fletcher

Klinische Epidemiologie

Grundlagen und Anwendung

2. Aufl. 2007. ISBN 978-3-456-84374-2

Diepgen (Hrsg.)

Grundwissen Epidemiologie, medizinische Biometrie und medizinische Informatik

Querschnittsbereich 1

2008. ISBN 978-3-456-84180-9

Bachmann / Puhan / Steurer (Hrsg.)

Patientenorientierte Forschung

Einführung in die Planung und Durchführung einer Studie

2008. ISBN 978-3-456-84538-8

Kuhn / Wildner

Gesundheitsdaten verstehen

2006. ISBN 978-3-456-84355-1

Coggon

Statistik

Einführung für Gesundheitsberufe

2007. ISBN 978-3-456-84465-7

Hüsler / Zimmermann

Statistische Prinzipien für medizinische Projekte

4. Aufl. 2006. ISBN 978-3-456-84233-2

Hall (Hrsg.)

Publish or Perish

Wie man einen wissenschaftlichen Beitrag schreibt, ohne die Leser zu langweilen
oder die Daten zu verfälschen

1998. ISBN 978-3-456-82884-8

Reinhardt

Vom Wissen zum Buch

Fach- und Sachbücher schreiben

2008. ISBN 978-3-456-84521-0

Weitere Informationen über unsere Neuerscheinungen finden Sie im Internet unter
www.verlag-hanshuber.com.

Regina Kunz
Khalid S. Khan
Jos Kleijnen
Gerd Antes

Systematische Übersichtsarbeiten und Meta-Analysen

Einführung in Instrumente der evidenzbasierten
Medizin für Ärzte, klinische Forscher und Experten
im Gesundheitswesen

2., vollständig überarbeitete Auflage

Verlag Hans Huber

Korrespondenzanschrift:
Prof. Dr. Regina Kunz
Institut für Klinische Epidemiologie
Universitätsspital Basel
Hebelstrasse 10
CH-4031 Basel
Schweiz
rkunz@uhbs.ch

Lektorat: Dr. Klaus Reinhardt
Bearbeitung: Ulrike Boos
Herstellung: Anna Colosio
Umschlag: Atelier Mühlberg, Basel
Druckvorstufe: am-productions GmbH, Wiesloch
Druck und buchbinderische Verarbeitung: Hubert & Co, Göttingen
Printed in Germany

Bibliographische Information der Deutschen Bibliothek

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



Dieses Werk, einschließlich aller seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen sowie die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen oder Warenbezeichnungen in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen-Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürfen.

Anregungen und Zuschriften an:

Verlag Hans Huber
Lektorat Medizin
Länggass-Strasse 76
CH-3000 Bern 9
Tel: 0041 (0)31 300 4500
Fax: 0041 (0)31 300 4593
verlag@hanshuber.com
www.verlag-hanshuber.com

2. Auflage 2009

© 2009 by Verlag Hans Huber, Hogrefe AG, Bern

ISBN 978-3-456-84691-0

Inhaltsverzeichnis

1	Einleitung	1
1.1	<i>Kritische Bewertung systematischer Reviews</i>	3
1.2	<i>Erstellung eines systematischen Reviews</i>	5
1.3	<i>Aufbau des Buches</i>	6
2	Schritte eines systematischen Reviews	9
2.1	<i>Schritt 1: Reviewfragen formulieren</i>	10
2.1.1	Wie man Reviewfragen formuliert	10
2.1.2	Variation von Populationen, Interventionen und Endpunkten	11
2.1.3	Variationen im Studiendesign	16
2.1.4	Modifikation Ihrer Fragen während der Reviewerstellung	20
2.2	<i>Schritt 2: Relevante Literatur identifizieren</i>	22
2.2.1	Erstellen einer vorläufigen Literaturliste	22
2.2.2	Sammeln und Management von Literaturstellen	30
2.2.3	Auswahl der relevanten Studien	31
2.2.4	Publikationsbias und andere systematische Verzerrungen	35
2.3	<i>Schritt 3: Qualität der Literatur bewerten</i>	37
2.3.1	Entwicklung von Checklisten zur Qualitätsbewertung	37
2.3.2	Qualitätsbewertung in Reviews mit verschiedenen Studiendesigns	43
2.3.3	Wie zuverlässig ist die Qualitätscheckliste eines Reviews?	45
2.3.4	Wie man Qualitätsbewertungen in einem Review verwendet	46
2.4	<i>Schritt 4: Evidenz zusammenfassen</i>	50
2.4.1	Aufbau eines Evidenzberichts	50
2.4.2	Unterschiede in den Effekten zwischen den Studien untersuchen	54
2.4.3	Meta-Analyse (quantitative Synthese) der Therapieeffekte einzelner Studien	56
2.4.4	Klinische Heterogenität	59
2.4.5	Methodische Heterogenität	61
2.4.6	Durchführung einer Meta-Analyse bei ungeklärter Heterogenität	62
2.5	<i>Schritt 5: Ergebnisse interpretieren</i>	64
2.5.1	Aufspüren von Publikationsbias und ähnlichen Biasformen	65
2.5.2	Wie man die Aussagekraft der Reviewergebnisse bestimmt	66
2.5.3	Wie man die Reviewergebnisse klinisch umsetzt	70
2.5.4	Evidenzprofile und Praxisempfehlungen	73

3	Fallstudien	81
3.1	<i>Fallstudie 1: Identifizierung und Bewertung systematischer Reviews</i>	82
3.1.1	Szenario: Medikamentöse Behandlung einer beginnenden Schizophrenie	82
3.1.2	Auflösung des Szenarios	90
3.2	<i>Fallstudie 2: Review zu potenziellen Risiken einer Public-Health-Maßnahme</i>	91
3.2.1	Szenario: Risiken der Trinkwasserfluoridierung	91
3.2.2	Auflösung des Szenarios	99
3.3	<i>Fallstudie 3: Review zur Wirksamkeit einer Therapie</i>	100
3.3.1	Szenario: Antimikrobielle Therapie zur Behandlung chronischer Wunden	100
3.3.2	Auflösung des Szenarios	109
3.4	<i>Fallstudie 4: Reviews zur Testgenauigkeit</i>	110
3.4.1	Szenario: Ultraschall als bildgebender Test bei postmenopausalen Frauen mit Vaginalblutung	110
3.4.2	Auflösung des Szenarios	121
	Empfohlene Literatur	124
	Glossar	128
	Sachregister	142

Vorwort zur 2. Auflage

Evidenzbasierte Medizin und evidenzbasierte Gesundheitsversorgung brauchen eine Basis aus hochwertiger Evidenz. Aus guten Gründen haben systematische Übersichtsarbeiten und Meta-Analysen dafür einen festen Platz in der Versorgung von Patienten, bei der Erstellung von Leitlinien und für Entscheidungen im Gesundheitssystem gefunden. Hier tragen Wissenschaftler und Reviewer in systematischer und umfassender Weise das vorhandene Wissen zu klinisch wichtigen Fragestellungen zusammen. Sie liefern Ärzten, Patienten und Entscheidungsträgern fundierte Grundlagen für die Behandlung von kranken Menschen, für den (Stellen-)Wert eines diagnostischen Tests oder für die Prognose eines Krankheitsverlaufs. In Abhängigkeit von der Komplexität der Fragestellung und dem Umfang der verfügbaren Evidenz investiert ein Review-Team zwischen drei Monaten und drei Jahren, um die Literatur zu suchen, zu sichten, zu bewerten und zu analysieren. Wie bedeutsam diese Publikationen geworden sind und welchen Stellenwert ihnen als eine Grundlage ärztlichen Handelns beigemessen wird, zeigt allein die Tatsache, dass nahezu jede Ausgabe von hochrangigen internationalen medizinischen Fachzeitschriften wie dem JAMA (Journal of the American Medical Association), den Annals of Internal Medicine, dem Lancet oder dem British Medical Journal eine oder mehrere systematische Reviews und Meta-Analysen veröffentlicht. Seit kurzem wird die Cochrane Database of Systematic Reviews im ISI Web of Knowledge erstmals mit einem Impact Factor bewertet. Analysen aus den Vorjahren stuften die Zitierhäufigkeit und damit die Bedeutung der Reviews mit einem Impact Factor von 4.654 ein. Die Höhe des Wertes hat selbst überzeugte Befürworter von Cochrane Reviews positiv überrascht. Weitere wissenschaftliche Anerkennung erhalten systematische Reviews und Meta-Analysen durch das Bundesministerium für Bildung und Forschung in seinem Programm zur Förderung klinischer Studien. Hier werden systematische Übersichtsarbeiten und Meta-Analysen mit multizentrischen prospektiven, kontrollierten klinischen Studien gleichgestellt und Forschungsgelder nur beim Erfüllen hoher methodischer Auflagen vergeben. Mit dieser Gleichbehandlung steht das BMBF im internationalen Vergleich nicht alleine.

Was hat es nun mit systematischen Übersichtsarbeiten und Meta-Analysen für eine besondere Bewandnis, dass sie in der Krankenversorgung wie auch bei der öffentlichen Forschungsfinanzierung so große Aufmerksamkeit erhalten? Die Antwort auf diese Frage möchten wir Ihnen im vorliegenden Buch näher bringen. Wir beginnen mit der „Philosophie“ von systematischen Reviews und Meta-Analysen und führen Sie Schritt für Schritt durch jede Phase der Reviewerstellung, bis sich Ihnen die Evidenz in strukturierten Tabellen und grafischen Forestplots erschließt. Nach der Lektüre werden Sie in der Lage sein, veröffentlichte systematische Reviews kritisch zu bewerten und die daraus abgeleiteten Schlussfolgerungen und Empfehlungen zu beurteilen. Ein weiterer Schwerpunkt dieses Buches ist die Vermittlung von Grundwissen zur Planung und Erstellung solcher Übersichtsarbeiten und Meta-Analysen. Als Zielgruppe haben wir dabei Ärzte und Einsteiger in das Gebiet der Reviewerstellung im Auge, weniger die erfahrenen Epidemiologen und Statistiker. Mit diesem Buch sollten Sie sich in der Lage fühlen, einen einfachen Review zu beginnen. Das Buch wird Ihnen für jeden Schritt ein zuverlässiger Begleiter sein.

Ob als Nutzer von Reviews oder als Reviewer in den Startlöchern – wir wünschen Ihnen bei der Lektüre des Buches viel Vergnügen und bei der Umsetzung in ein eigenes Projekt viel Erfolg.

Die Autoren

Danksagungen

Ohne allseitige Unterstützung ist es kaum möglich, ein solches Buch fertigzustellen. Unser besonderer Dank gilt Susan Hahné, Anjum Doshani, Peter J. Thompson und Jack Cohen für die kritische Durchsicht einer früheren Version dieses Buches; Mary Publicover für ihre kritische Stellungnahme zu Schritt 2; Christine Anne Clark und Anne Marie Bagnall für die Durchsicht von Fallstudie 1 und Sue O'Meara für die Durchsicht von Fallstudie 3. Edith Motschall hat uns mit vielen hilfreichen Hinweisen zur Literatursuche in Deutschland unterstützt. Robert Wolff und Katja Suter danken wir für ihre Hilfe bei der redaktionellen Überarbeitung der 2. Ausgabe, vor allem der Aktualisierung der Internet-Seiten und der grafischen Ausgestaltung.

Khalid S. Khan widmet dieses Buch Heinke Kunst, Zarie und Jani.

Anmerkungen

Medizin als Wissenschaft ist einem sehr schnellebigen Wandel unterworfen. In dem Maße, wie neue Forschung und klinische Erfahrung unser Wissen verbreitert, verändern sich Behandlungsstandards und Medikamente. Wir Autoren haben lange diskutiert, ob wir die Beispiele der ersten Auflage aktualisieren bzw. durch neue Beispiele ersetzen sollen. Da jedoch die Illustration methodischer Zusammenhänge und Vorgehensweisen (und weniger der spezifische medizinische Inhalt) im Vordergrund steht und Leser in ihrem Feedback gerade die Beispiele als besonders anschaulich hervorgehoben haben, beschlossen wir, hier auf Veränderungen zu verzichten. Leser werden daher gebeten, die medizinischen Informationen des Buches mit anderen Quellen abzugleichen.

Internetseiten unterliegen einem beständigen Wechsel. Die Internetadressen in diesem Buch wurden zuletzt im Juli 2008 überprüft und aktualisiert.

Die Autoren haben in den Fallstudien und Beispielen des vorliegenden Buches die Evidenz im Kontext spezifischer Szenarien interpretiert. Bei der Interpretation der systematischen Reviews in anderen Zusammenhängen sind natürlich auch andere Auflösungen dieser Szenarien denkbar.

Über die Autoren

Mit zusammen mehr als 150 systematischen Reviews zählen wir schon zu den „alten Hasen“. Im Laufe der Jahre haben wir gemeinsam mit Ärzten, klinischen Forschern und Experten Reviews erstellt, deren Ergebnisse in den klinischen Alltag und Entscheidungen im Gesundheitswesen eingeflossen sind. Mit anderen Epidemiologen und Statistikern arbeiten wir daran, die Methoden zur Erstellung systematischer Reviews zu verbessern. Wir arbeiten im akademischen Bereich und beschäftigen uns mit der Erstellung und Förderung systematischer Reviews. Zwei von uns arbeiten zusätzlich im klinischen Bereich und wenden die wissenschaftlichen Erkenntnisse aus Reviews und Meta-Analysen in der praktischen Arbeit mit Patienten an.

Regina Kunz ist klinische Epidemiologin am Basel Institut für Klinische Epidemiologie und Nephrologin am Universitätsspital Basel. Sie hat umfangreiche Erfahrungen in der Erstellung von systematischen Übersichtsarbeiten, Meta-Analysen und HTA-Berichten und weiss aus ihrer klinischen Tätigkeit um den Wert hochwertiger Evidenzsynthesen in der Patientenversorgung. Aufgrund ihrer langjährigen Zugehörigkeit zur GRADE Working Gruppe und ihrer Tätigkeit im Vorstand von G-I-N, dem Guidelines International Network, kennt sie den Bedarf nach Reviews und Meta-Analysen in allen Bereichen des Gesundheitswesens, um Ärzten, Patienten und Entscheidungsträger fundierte Empfehlungen zur Verfügung zu stellen, und die Schwierigkeiten, in Abwesenheit guter Reviews überzeugende Empfehlungen zu formulieren. Sie ist Gründungsmitglied des Deutschen Netzwerkes EbM, Mitherausgeberin des Werkes *Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis* (Deutscher Ärzteverlag 2007) und verfügt über umfangreiche Erfahrung in der Leitung von Kursen zu systematischen Reviews und EbM.

Khalid Khan ist Professor für Gynäkologie/Geburtshilfe und Klinische Epidemiologie an der University of Birmingham und Direktor des WHO Collaborating Centre for Research Synthesis in Reproductive Health in Birmingham. Er hat umfassende Erfahrung in der Erstellung systematischer Reviews zu therapeutischen Interventionen und diagnostischer Testverfahren und der zugrundeliegenden Methodik. Mit seiner Qualifikation in Medical Education ist er stellvertretender Direktor des Education Resource Centre am Birmingham Women's Hospital, wo er mit Aktivitäten wie Journal Clubs für eine evidenzbasierte Patientenversorgung, EbM-Visiten oder Kursen in kritischer Studienbewertung einen Schwerpunkt in Evidenzbasierter Medizin aufgebaut hat. Er leitet die Abteilung Research and Development der Klinik, lehrt in der Aus- und Weiterbildung, und ist für mehrere klinische Zeitschriften als Reviewer im Peer-Review-Verfahren tätig.

Jos Kleijnen ist Direktor der Firma Kleijnen Systematic Reviews Ltd in York (Großbritannien) und ehemaliger Direktor des Centre for Reviews and Dissemination in York. Nach Abschluss des Medizinstudiums hat er die Laufbahn eines klinischen Epidemiologen eingeschlagen. Er verfügt über vielfältige Erfahrungen in der Erstellung und Verbreitung systematischer Reviews sowie anderer Forschungsarbeiten. Er gehört verschiedenen Steering Groups und wissenschaftlichen Beiräten von Auftraggebern für Reviewerstellung an. Weiterhin war er Gründungsdirektor des niederländischen Cochrane Zentrums und ist Mitglied verschiedener Methods Working Groups der Cochrane Collaboration. Daneben ist er Editor der Cochrane Peripheral Vascular Diseases Review Group, leitet in mehreren Ländern Kurse zu systematischen Reviews und EbM und pflegt eine enge Zusammenarbeit mit dem Horten-Zentrum in Zürich sowie der Vienna School of Clinical Research in Wien.

Gerd Antes ist Direktor des Deutschen Cochrane Zentrums in Freiburg. Als Biometriker verfügt er über eingehende Kenntnisse der mathematischen Grundlagen von Meta-Analysen, Heterogenität, Funnel-Plots u.ä. Abgesehen von seinem Interesse an methodischer Forschung, statistischen Auswertungen und medizinischer Informatik hat er die deutsche Cochrane-Initiative ins Leben gerufen und widmet einen Großteil seiner Zeit der Förderung von EbM und systematischen Reviews in Deutschland. Er ist Gründungsmitglied und war Sprecher des Deutschen EbM-Netzwerkes. Er gehörte mehrere Jahre der *Steering Group* der Cochrane Collaboration an.

Aus unserer Überzeugung heraus, dass Ärzte und andere Gesundheitsberufe viel aus Reviews und Leitlinien lernen können, haben wir dieses Buch geschrieben. Umgekehrt benötigen Reviews und Leitlinien den Input von Ärzten und anderen Gesundheitsberufen. Wir hoffen, dass Ärzte, klinische Forscher und Experten in der Gesundheitsversorgung sich durch dieses Buch ermutigt fühlen, systematische Reviews kritisch zu lesen, effektiv anzuwenden und sich auch daran wagen, den ersten eigenen Review zu erstellen.

Häufig verwendete Abkürzungen

CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Trials
CER	Control Event Rate, Ereignisrate in der Kontrollgruppe
CI	Konfidenzintervall
DARE	Database of Abstracts of Reviews of Effects
EbM	Evidenzbasierte Medizin
EER	Experimental Event Rate, Ereignisrate in der Experimentalgruppe
ES	Effect Size, Effektgröße für kontinuierliche Daten
HTA	Health Technology Assessment
ITT	Intention-to-Treat-Analyse
LR	Likelihood Ratio, Wahrscheinlichkeitsverhältnis (LR+, LR bei positivem Testergebnis; LR-, LR bei negativem Testergebnis)
MeSH	Medical Subject Heading
NNT	Number Needed to Treat
OR	Odds Ratio [nicht zu verwechseln mit dem Boole'schen OR- (≙ ODER-)Operator in elektronischen Literaturrecherchen]
RCT	Randomisierte, kontrollierte Studie
RD	Risikodifferenz (oder ARR, Absolute Risikoreduktion)
RR	Relatives Risiko
SD	Standard Deviation, Standardabweichung
SE	Standard Error, Standardfehler

1 Einleitung

- 1.1 Kritische Bewertung systematischer Reviews – 3
- 1.2 Erstellung eines systematischen Reviews – 5
- 1.3 Aufbau des Buches – 6

Ein **systematischer Review** ist ein wissenschaftlicher Artikel, in dem relevante Studien identifiziert, ihre Qualität bewertet und ihre Ergebnisse nach wissenschaftlichen Methoden zusammengefasst werden.

Eine **Meta-Analyse** ist nicht mit einem **systematischen Review** identisch, sondern vielmehr ein Teil des **Reviews**. Sie ist ein statistisches Verfahren, bei dem die Ergebnisse mehrerer Einzelstudien zu einem Gesamtergebnis zusammengefasst werden. *Nicht alle Publikationen, die unter der Bezeichnung Meta-Analyse herausgegeben werden, sind systematische Reviews!*

Im weiteren Verlauf beziehen wir uns mit dem Begriff **Reviews** stets auf **systematische Reviews**, die ausschließlich nach systematischen Methoden verfasst sein sollten. In diesem Buch verwenden wir also beide Begriffe synonym.

Heutzutage gibt es kaum noch medizinische Fachzeitschriften, in denen man keine Übersichtsartikel (Reviews) findet. Woher kommt das? Übersichtsarbeiten fassen wissenschaftliche Erkenntnisse (die sog. Evidenz) zu einem bestimmten Thema aus einer Vielzahl von Einzelstudien zusammen. Häufig sind Forschungsergebnisse, die für unsere ärztliche Tätigkeit von Interesse sind, aber über die gesamte medizinische Literatur verstreut und meist auch in uns fremden Sprachen publiziert. Ein Übersichtsartikel hingegen verschafft uns problemlos einen schnellen Überblick über die gesamte Bandbreite der Evidenz zu dem uns interessierenden Thema. Deshalb sind Übersichtsarbeiten auch so beliebt. Wir können uns damit auf dem Laufenden halten, ohne sämtliche, für unsere Praxis relevanten Einzelstudien durcharbeiten zu müssen. Wem kämen solche Übersichtsarbeiten angesichts zunehmender Belastungen im Berufsleben und dem damit verbundenen Mangel an Zeit nicht gelegen? Seien wir doch ehrlich: Selbst wenn wir Zeit und Mittel hätten, relevante Studien zu finden und zu bewerten, würden viele von uns trotz allem lieber zu einem Review greifen.

Aber hier ist ein Wort der Warnung angebracht: Die Art und Weise, wie in konventionell erstellten Reviews Studien recherchiert, wissenschaftliche Belege gesammelt und Schlussfolgerungen abgeleitet werden, ist oft genug fragwürdig. Schlimmstenfalls sind der gesamte Reviewprozess und seine Schlussfolgerungen von der persönlichen Sichtweise des Autors beeinflusst. Denn bei vielen Reviews, die wir lesen, handelt es sich um Beiträge, die auf Einladung verfasst wurden, und nicht um genuin wissenschaftliche Arbeiten. Können wir ausschließen, dass uns Übersichtsarbeiten dieser Art nicht in die Irre leiten? Wohl kaum. Und dies ist auch der Grund, weshalb konventionelle Übersichtsarbeiten heutzutage mehr und mehr durch systematische Reviews verdrängt werden.

Sorgfältig erstellte systematische Reviews der medizinischen Literatur sind eigenständige Forschungsarbeiten. Sie identifizieren relevante Studien, bewerten ihre Qualität und fassen deren Ergebnisse anhand wissenschaftlicher Methoden zusammen. Dadurch unterscheiden sie sich von konventionellen Übersichtsarbeiten und aus dem Stehgreif erstellten „Experten“-Beiträgen. Wichtiger noch, die Empfehlungen systematischer Reviews beruhen auf ausgewogenen Schlussfolgerungen, die aus der gesammelten Evidenz abgeleitet sind und geben nicht die persönliche Meinung von „Experten“ wieder.

Das vorliegende Buch beschreibt die grundlegenden Prinzipien der systematischen Auswertung von Primärstudien, die sich mit den Wirkungen medizinischer Interventionen befassen. Mit Hilfe dieses Buches soll es Ihnen möglich werden, einen Review selbstbewusst auf seine Qualität und Aussage hin kritisch zu bewerten und auch mit der Erstellung einer eigenen Übersichtsarbeit zu beginnen.

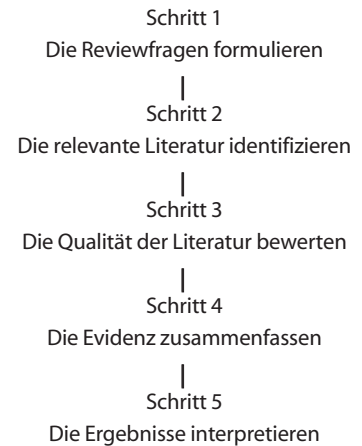
1.1 Kritische Bewertung systematischer Reviews

Immer häufiger stützen sich gesundheitspolitische Entscheidungen auf klar strukturierte, umfassende Zusammenfassungen von Informationen, die durch systematische Reviews der relevanten Literatur gewonnen wurden. Damit erfordert das Praktizieren einer evidenzbasierten Versorgung in der heutigen Zeit mehr als nur die kritische Bewertung einzelner Studien. Medizinische Leitlinien, deren Empfehlungen häufig auf systematischen Übersichten beruhen, liefern ein ausgezeichnetes Beispiel dafür, dass systematische Reviews aus unserem Berufsleben nicht mehr wegzudenken sind.

Vermutlich kommen systematische Reviews einem Quantensprung in der Erstellung von Übersichten gleich. Allerdings sollte man ihnen nicht unbedingt blind vertrauen, denn ebenso wie einzelne Studien können auch Reviews und Leitlinien von unterschiedlichster Qualität sein. Es gibt zahlreiche Beispiele dafür, dass selbst Reviews aus hervorragenden medizinischen Fachzeitschriften oder Leitlinien von wissenschaftlichen Fachgesellschaften gleichwohl von schlechter Qualität sein können. Deswegen müssen wir auch in scheinbar glaubwürdigen Reviews und Leitlinien mit irreführenden Schlussfolgerungen rechnen. Für jeden praktizierenden Arzt sollte dies Grund genug sein, sich eingehender mit den Prinzipien systematischer Reviews zu befassen. Was aber, wenn Sie im Bereich Gesundheitsforschung lediglich über methodische Grundkenntnisse verfügen und Ihnen die Bewertung von Reviews als sehr mühsam erscheint? Dieses Buch wird Ihnen helfen, sich den Prozess der systematischen Auswertung medizinischer Literatur mit all seinen Tücken zu erschließen, und Sie werden lernen, besser zwischen hochwertigen und weniger hochwertigen Reviews und Leitlinien zu unterscheiden.

Verschiedene Quellen mit bereits existierenden Reviews, an denen wir unsere Entscheidungen in der Patientenversorgung orientieren können, sind in **■** Tabelle 1.1 aufgelistet. Bei der Suche nach geeigneten Reviews wird uns gelegentlich aber schmerzlich bewusst, dass zu vielen Fragen aus der täglichen Praxis entweder gar keine relevanten Reviews existieren oder vorhandene Reviews nicht adäquat sind. Warum also nicht selbst einen Review erstellen, wenn sich keiner finden lässt, der den Anforderungen gerecht wird?

Welche Schritte muss man beim Auffinden der Bewertung und der Anwendung von Reviews berücksichtigen?



Evidenzbasierte Medizin (EbM) ist der gewissenhafte, ausdrückliche und abwägende Gebrauch der gegenwärtig besten wissenschaftlichen Evidenz für Entscheidungen in der medizinischen Versorgung individueller Patienten. Systematische Reviews liefern aussagekräftige Evidenz zur Unterstützung der EbM.

Leitlinien sind systematisch entwickelte Entscheidungshilfen für Ärzte und Patienten über die angemessene Vorgehensweise bei speziellen Gesundheitsproblemen. Häufig, aber *nicht* immer, stützen sie sich auf Evidenz aus systematischen Reviews.

■ Tabelle 1.1. Ausgewählte Quellen für systematische Reviews und Leitlinien

Cochrane Library* (www.thecochranelibrary.com)

Die Cochrane Library umfasst drei Datenbanken veröffentlichter Reviews und Reviews in Vorbereitung:

— Cochrane Database of Systematic Reviews (CDSR; Cochrane Reviews)

Abstracts und Volltexte von systematischen Reviews zu medizinischen Interventionen, die nach der Methodik der Cochrane Collaboration erstellt wurden und regelmäßig aktualisiert werden, sowie Protokolle zu laufenden Reviews. Die Produkte einzelner Cochrane Review Groups (CRGs) finden sich unter „About Cochrane“ und umfassen eine Liste mit den Publikationen aller 51 CRGs.



■ **Tabelle 1.1.** Fortsetzung

■ **Database of Abstracts of Reviews of Effects (Dare; other Reviews)**

Enthält kurze Inhaltsangaben mit kritischen Bewertungen von systematischen Reviews, die nicht als Cochrane Reviews erstellt wurden. Die Publikationen werden durch regelmäßige Recherchen bibliographischer Datenbanken, Handsuchen in wichtigen medizinischen Fachzeitschriften sowie durch Sichtung der grauen Literatur identifiziert (kostenloser Zugriff unter www.york.ac.uk/inst/crd/crddatabases.htm).

■ **Health Technology Assessments Database (HTA; Technology Assessment)**

Abstracts abgeschlossener Technologiebewertungen und laufender Projekte, die von den Mitgliedern des International Network of Agencies for Health Technology Assessment (INAHTA) sowie anderer HTA-Einrichtungen durchgeführt werden. Die meisten umfassen systematische Reviews (kostenloser Zugriff unter www.york.ac.uk/inst/crd/crddatabases.htm)

Es existieren wesentlich mehr systematische Reviews, als man gemeinhin denkt. Allein in der 3. Ausgabe der Cochrane Library 2008 befanden sich 3364 abgeschlossene Reviews und 2052 Protokolle über laufende Cochrane Reviews, 8349 Abstracts von Reviews mit strukturierter methodischer Bewertung unter „Other Reviews“ und 7486 Abstracts von abgeschlossenen Technologie-Bewertungen („technology assessments“).

Allgemeine elektronische Datenbanken (siehe dazu auch Tabelle 2.6)

- MEDLINE – PubMed Clinical Queries bei Verwendung der Suchoption „Systematic Reviews“; zugänglich unter www.ncbi.nlm.nih.gov/entrez/query/static/clinical.html
- CINAHL, EMBASE, PsycLIT etc. können nach Reviews durchsucht werden, indem die Suchfilter des Centre for Reviews and Dissemination (Kombination aus Textwörtern, Deskriptoren und Schlagwörtern zur Erfassung relevanter Artikel) entsprechend angepasst werden. Die Suchstrategien des Centre for Reviews and Dissemination sind zugänglich unter www.york.ac.uk/inst/crd/search.htm.

Ausgewählte Internetseiten (in alphabetischer Reihenfolge)

- CMA Infobase – www.mdm.ca/cpgsnew/cpgs/index.asp
- Guidelines and Guidelines in Practice – www.eguidelines.co.uk
- GIN Guidelines International Network – www.g-i-n.org
- HTA-Programm des National Institute for Health Research (NIHR) – www.nchta.org/project/htapubs.asp
- National Electronic Library for Health – www.nelh.nhs.uk
- National Institute for Health and Clinical Excellence (NICE) – www.nice.org.uk
- National Guidelines Clearinghouse – www.guideline.gov
- Medizin-Suchmaschine von Intute – omni.ac.uk (verwendet eine fortgeschrittene Suchstrategie und spezifiziert Leitlinien nach Quellentyp)
- SchARR-Lock's Guide to the evidence – www.shef.ac.uk/scharr/ir/scebm.html
- SIGN-Leitlinien – www.sign.ac.uk
- Turning Research Into Practice (TRIP) – www.tripdatabase.com

Ausgewählte Printpublikationen

- Clinical Evidence – www.clinicalevidence.org
- Effective Health Care Bulletin – www.york.ac.uk/inst/crd/ehcb.htm

* Für ein Recherchebeispiel der Cochrane Library siehe Fallstudie 1
+ zugänglich unter www.york.ac.uk/inst/crd

Webseiten sind beständigen Änderungen unterworfen. Die Internet-Adressen in diesem Buch stammen von Juli 2008.

1.2 Erstellung eines systematischen Reviews

Internetrecherchen, die elektronische Beschaffung oder Fernleihe von Artikeln, benutzerfreundliche Software für Meta-Analysen usw. schaffen die Voraussetzungen für eine neue Form der Literaturlauswertung. Da die entsprechenden Ressourcen in zunehmendem Maße auch im klinischen Bereich verfügbar sind, ist die Erstellung systematischer Reviews für praktizierende Ärzte heutzutage eine durchaus realistische Option. Weshalb aber sollten Ärzte eigene Reviews erstellen?

Gründe dafür gibt es genug. Manch einer will vielleicht eine Übersichtsarbeit erstellen, um damit eine evidenzbasierte Versorgung zu fördern oder um klinische Empfehlungen oder Strategien zu untermauern. Andere sehen darin einen Beitrag zur eigenen beruflichen Weiterentwicklung oder eine Gelegenheit für eine Publikation in einem Journal mit Peer-Review. Wiederum andere möchten das Einleitungskapitel zu einer Dissertation oder einem Forschungsprojekt schreiben, eine Präsentation für eine wissenschaftliche Tagung vorbereiten oder einen Gastbeitrag verfassen. Dazu müssen wir das Rad natürlich nicht jedesmal neu erfinden, sondern sollten uns dabei weitestmöglich auf schon vorhandene Reviews und Leitlinien stützen. Denn vielleicht gibt es bereits aktuelle hochwertige Reviews mit sämtlichen Informationen, die wir brauchen. Existieren aber zu einem Thema noch keine Reviews oder sind die vorhandenen veraltet bzw. von nicht ausreichender Qualität, dann können wir entweder „Experten“ um Rat fragen, vorhandene Primärstudien auswerten oder einen systematischen Review erstellen.

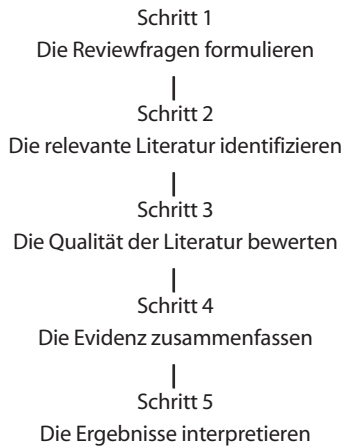
Wir sollten uns aber darüber klar sein, dass eine „Expertenmeinung“ unter Umständen nicht evidenzbasiert ist bzw. von anderen nicht unbedingt akzeptiert wird, denn zu jeder „Expertenmeinung“ gibt es bekanntlich auch immer eine Gegenmeinung. Die alleinige Bewertung von einzelnen Studien kann selbstverständlich kein vollständiges Bild ergeben. Und eben dies ist der Ansatzpunkt für einen neuen Review. Viele Cochrane-Reviews entstehen auf genau diese Weise, und sind sie erst einmal publiziert, nutzen sie Ärzten, Entscheidungsträgern, Patienten. Die Erstellung eines neuen systematischen Reviews ist zwar sehr mühsam, doch die Mühe lohnt sich.

Was ist unser nächster Schritt? Es gibt nur einen Weg. Als Akademiker im Gesundheitswesen, ohne tiefere Kenntnisse in Epidemiologie und Statistik, mögen wir uns zwar mit der Veröffentlichung von Leitartikeln, Stellungnahmen und Kommentaren auskennen, doch steigt der Druck seitens der Herausgeber von Zeitschriften, dass wir in unseren Arbeiten systematischer vorgehen. Warum also keinen systematischen Review als nächsten Beitrag einreichen? Möglicherweise haben wir Bedenken, weil wir nicht über die nötigen Kenntnisse oder Fertigkeiten zur Abfassung von Reviews verfügen. Doch dem kann abgeholfen werden: Das vorliegende Buch enthält alle wichtigen Informationen, die man zur Planung und Erstellung von Reviews der medizinischen Literatur braucht.

In erster Linie wendet sich dieses Buch an Ärzte und andere Gesundheitsberufe sowie noch unerfahrene Reviewer, weniger an Epidemiologen und Statistiker. Es versetzt den Leser nicht nur in die Lage, ohne die Unterstützung professioneller Reviewer eigene Reviews zu verfassen, sondern enthält außerdem Hinweise auf weiterführende Literatur und Ratschläge, wie man sich bei schwierigen Fragestellungen erfolgreich um professionelle Unterstützung bemüht. Angesichts des Arbeitsaufwandes, der in den verschiedenen Schritten der Reviewerstellung steckt,

Die Cochrane Collaboration ist ein internationales Netzwerk von Ärzten und Wissenschaftlern. Ihr Ziel ist es, systematische Reviews zu medizinischen Interventionen zu erstellen, zu aktualisieren und zugänglich zu machen, um auf diese Weise einen Beitrag zur informierten Entscheidungsfindung im Gesundheitswesen zu leisten (www.cochrane.org).

Bei einem **Cochrane-Review** handelt es sich um einen systematischen Review, der nach der Methodik der **Cochrane Collaboration** erstellt und in der **Cochrane Library** erfasst wurde.



Das vorliegende Buch befasst sich vorwiegend mit wissenschaftlichen Reviews, in denen die Auswirkungen von *Interventionen* oder *Expositionen* auf die Gesundheit untersucht werden.

ist man gut beraten, einen oder mehrere Reviewer zu finden, die Lust haben, sich an dieser Arbeit zu beteiligen. Zudem ist es für Anfänger ratsam, an lokalen Workshops oder Kursen zu systematischen Reviews teilzunehmen. Viele davon werden von der Cochrane Collaboration veranstaltet. Wenden Sie sich doch einfach an das Deutsche Cochrane Zentrum (www.cochrane.de) und erkundigen sich dort nach dem nächsten Workshop-Termin.

1.3 Aufbau des Buches

Die Anleitungen in diesem Buch sollen Ihnen helfen, sich die Prinzipien systematischer Reviews zu erschließen. Schritt für Schritt wird der Reviewprozess in den folgenden Kapiteln erklärt. Es sind nur fünf Schritte. Jeder Schritt enthält präzise Anleitungen. Diese wiederum werden anhand von Beispielen erläutert, die wir veröffentlichten Reviews entnommen haben. Viele dieser Beispiele werden Sie durch die verschiedenen Reviewschritte begleiten, so dass die Verbindung zwischen den einzelnen Schritten deutlich erkennbar wird. Darüber hinaus wird die Anwendung der Theorie anhand von Fallstudien veranschaulicht. Jeder Fall besteht aus einem Szenario, für das wir Evidenz aus Reviews benötigen, aus einer Demonstration verschiedener Reviewmethoden sowie einem Vorschlag zur Auflösung des Szenarios. Das Erarbeiten der verschiedenen Schritte, Beispiele und Fallstudien gewährt dabei tiefere Einblicke in die kritische Bewertung und Durchführung eines systematischen Reviews.

Wenn Sie sich dazu entschlossen haben, mit einem Review zu beginnen, raten wir Ihnen dringend, vorher eine Projektskizze oder ein Protokoll mit entsprechenden Hintergrundinformationen zu erstellen, in dem Sie das zu behandelnde Problem sowie die für Ihren Review vorgesehene Methodik darlegen. Während des gesamten Reviewprozesses wird dieses Protokoll Sie stets an Ausgangspunkt und Ziel Ihrer Arbeit erinnern. Außerdem kann man das Protokoll einem Peer-Review-Verfahren unterziehen, bevor Sie mit der eigentlichen Arbeit beginnen. Gelegentlich wird vorgeschlagen, ein solches Protokoll ins Internet einzustellen, um eine breitere Rückmeldung zu erhalten. Sie sollten sich aber bewusst sein, dass die Kommentare der Besucher Ihrer Homepage von unterschiedlichster Qualität sein können. Realistisch gesehen steigen Ihre Chancen auf einen fachlich fundierten Kommentar deutlich, wenn Sie das Protokoll gezielt einem Kollegen mit Reviewerfahrung zu lesen geben oder wenn Sie Ihr Protokoll bei der entsprechenden Reviewgruppe der Cochrane Collaboration registrieren lassen.

Bei der Erstellung eines solchen Protokolls wie auch im weiteren Verlauf der Reviewarbeit wird Ihnen das vorliegende Buch ein nützlicher Begleiter sein. Werfen wir noch einen Blick auf den Reviewprozess. Als erstes müssen Sie die Probleme, die Sie angehen wollen, als gut strukturierte Fragen formulieren (Schritt 1). Dies ist ein Schlüsselschritt: Alle weiteren Aspekte des Reviews leiten sich unmittelbar aus diesen Fragen ab. Zweitens müssen in einer gründlichen Literaturrecherche alle potenziell relevanten Studien identifiziert werden (Schritt 2). Damit wird aus einer Übersichtsarbeit ein systematischer Review. Drittens ist die Qualität der ausgewählten Studien zu bewerten (Schritt 3), und viertens werden die Studiencharakteristika und die Studienergebnisse zusammengefasst (Schritt 4). Falls durchführbar und sinnvoll, kann bei der Zusammenfassung der Ergebnisse auch eine statistische

Meta-Analyse hilfreich sein. Ausgehend von der Interpretation und Prüfung der klinischen Relevanz der Reviewergebnisse lassen sich schließlich entsprechende Schlussfolgerungen und Empfehlungen ableiten (Schritt 5).

Am Ende jedes einzelnen Schrittes werden die wichtigsten Aspekte der Bewertung und Durchführung von Reviews zusammengefasst. Fallstudien illustrieren die Anwendung der Reviewmethodik, die in den fünf Schritten entwickelt wird. Möglicherweise wird sich der eine oder andere Leser zunächst mit der Reviewmethodik befassen wollen, bevor er sich den Fallstudien zuwendet. Wieder andere lesen die Fallstudien lieber parallel zu den theoretischen Grundlagen in den einzelnen Schritten. Die Literaturempfehlungen enthalten Hinweise auf weiterführende Texte zu theoretischen und methodischen Fragestellungen, die über den hier abgedeckten Kernbereich hinausgehen.

Mit diesem Buch bieten wir Ihnen Grundlagen, die sich sowohl für den Anwender systematischer Reviews als auch den unerfahrenen Reviewer eignen. Das Buch ist nicht als fertiges Rezept zur Bewertung bzw. Erstellung systematischer Reviews gedacht; vielmehr präsentiert es eine Reihe von Modulen, die je nach Fragestellung und Kontext variiert werden können.

Übersicht

Schlüsselaspekte dieses Buches

- Dieses Buch versetzt den Leser in die Lage, die Qualität veröffentlichter Reviews zuverlässig zu bewerten und eigene Reviews zu verfassen.
- Es beschreibt die wichtigsten Prinzipien, wie zu spezifischen Fragen der Gesundheitsversorgung systematische Übersichtsarbeiten erstellt werden und wendet sich in erster Linie an Ärzte und andere Gesundheitsberufe.
- Schritt für Schritt wird anhand von anschaulichen Beispielen und Fallstudien erklärt, wie man Übersichtsarbeiten bewertet und durchführt.
- Am Ende jedes einzelnen Reviewschrittes werden die Schlüsselaspekte der kritischen Bewertung und Durchführung eines Reviews zusammengefasst.

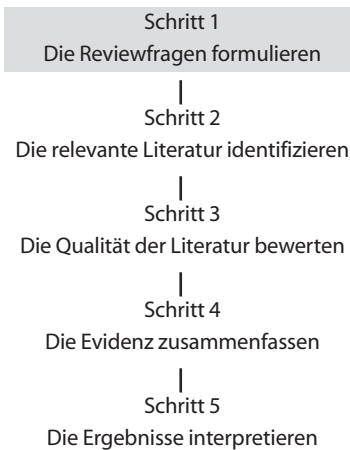
2 Schritte eines systematischen Reviews

- 2.1 Schritt 1: Reviewfragen formulieren – 10**
 - 2.1.1 Wie man Reviewfragen formuliert – 10
 - 2.1.2 Variation von Populationen, Interventionen und Endpunkten – 11
 - 2.1.3 Variationen im Studiendesign – 16
 - 2.1.4 Modifikation Ihrer Fragen während der Reviewerstellung – 20
- 2.2 Schritt 2: Relevante Literatur identifizieren – 22**
 - 2.2.1 Erstellen einer vorläufigen Literaturliste – 22
 - 2.2.2 Sammeln und Management von Literaturstellen – 30
 - 2.2.3 Auswahl der relevanten Studien – 31
 - 2.2.4 Publikationsbias und andere systematische Verzerrungen – 35
- 2.3 Schritt 3: Qualität der Literatur bewerten – 37**
 - 2.3.1 Entwicklung von Checklisten zur Qualitätsbewertung – 37
 - 2.3.2 Qualitätsbewertung in Reviews
 - mit verschiedenen Studiendesigns – 43
 - 2.3.3 Wie zuverlässig ist die Qualitätscheckliste eines Reviews? – 45
 - 2.3.4 Wie man Qualitätsbewertungen in einem Review verwendet – 46
- 2.4 Schritt 4: Evidenz zusammenfassen – 50**
 - 2.4.1 Aufbau eines Evidenzberichts – 50
 - 2.4.2 Unterschiede in den Effekten zwischen den Studien untersuchen – 54
 - 2.4.3 Meta-Analyse (quantitative Synthese) der Therapieeffekte einzelner Studien – 56
 - 2.4.4 Klinische Heterogenität – 59
 - 2.4.5 Methodische Heterogenität – 61
 - 2.4.6 Durchführung einer Meta-Analyse bei ungeklärter Heterogenität – 62
- 2.5 Schritt 5: Ergebnisse interpretieren – 64**
 - 2.5.1 Aufspüren von Publikationsbias und ähnlichen Biasformen – 65
 - 2.5.2 Wie man die Aussagekraft der Reviewergebnisse bestimmt – 66
 - 2.5.3 Wie man die Reviewergebnisse klinisch umsetzt – 70
 - 2.5.4 Evidenzprofile und Praxisempfehlungen – 73

Abschnitt A begleitet Sie Schritt für Schritt durch die einzelnen Phasen der Reviewerstellung. Es sind insgesamt nur fünf Schritte. In jedem Schritt werden die grundlegenden Prinzipien eines Reviews anhand von Beispielen aus veröffentlichten Reviews erläutert. Viele Beispiele werden Ihnen in den einzelnen Schritten immer wieder begegnen und Ihnen auf diese Weise die Verknüpfung der verschiedenen Reviewphasen vor Augen führen.

- Schritt 1: Die Reviewfragen formulieren
- Schritt 2: Die relevante Literatur identifizieren
- Schritt 3: Die Qualität der Literatur bewerten
- Schritt 4: Die Evidenz zusammenfassen
- Schritt 5: Die Ergebnisse interpretieren

2.1 Schritt 1: Reviewfragen formulieren



Fragekomponenten

- Patienten
- Interventionen
- Endpunkte
- Studiendesigns

Von systematischen Reviews erwarten wir Antworten auf fokussierte Fragen für eine optimale Versorgung unserer Patienten und der Bevölkerung. Der Schlüssel zum Erfolg eines Reviewprojektes liegt in der Fähigkeit des Reviewers, die Fragen und Probleme, die der Review beantworten soll, genau zu spezifizieren. Dies ist ein ganz entscheidender Schritt, denn wie Sie im Verlauf noch sehen werden, ergeben sich alle weiteren Schritte des Reviews aus eben diesen Fragen. Wir werden daher in diesem ersten Schritt nicht nur detailliert beschreiben, wie eine Reviewfrage formuliert wird, sondern auch kurz erläutern, wie man die Auswirkungen untersucht, die Änderungen in der Definition der einzelnen Bestandteile einer Reviewfrage möglicherweise mit sich bringen.

2.1.1 Wie man Reviewfragen formuliert

Die Formulierung einer Frage ist längst nicht so einfach, wie es auf den ersten Blick aussieht. Ein strukturiertes Vorgehen erweist sich hierbei als besonders nützlich. Das Gerüst einer Reviewfrage besteht aus vier Komponenten: den *Populationen*, *Interventionen* (bzw. *Expositionen*) und *Endpunkten*, die sich auf das gestellte Problem beziehen, sowie den *Studiendesigns*, die zur Beantwortung der Reviewfrage geeignet erscheinen. ■ Tabelle 2.1 zeigt für eine Vergleichsstudie, in welcher Beziehung die verschiedenen Bestandteile einer Reviewfrage zueinander stehen.

Die Formulierung unserer Reviewfragen in ■ Tabelle 2.1 mag für einen unerfahrenen Reviewer abschreckend wirken. Aber Sie sollten sich nicht gleich geschlagen geben. Wir werden Sie in den folgenden Abschnitten Schritt für Schritt durch die einzelnen Phasen der Fragenformulierung begleiten, damit Sie für Ihren Review gleich von Anfang an einen guten Start hinlegen. Zugegeben, es fällt auch erfahrenen Ärzten nicht immer leicht, Fragen für eine evidenzbasierte Praxis richtig zu formulieren. Daher kann es niemanden ernsthaft verwundern, wenn gerade Anfänger zu Beginn ihres Reviews mit Schwierigkeiten zu kämpfen haben. Dieser erste Schritt ist zwar mühsam, doch werden Sie bald merken, dass die Mühe sich auszahlt, da sich die weiteren Reviewsschritte wie von selbst direkt aus diesen Fragen ergeben.

Noch bevor die eigentliche Arbeit am Review beginnt, investieren die meisten ernsthaften Reviewer beträchtliche Zeit und Mühe, das Problem in die richtigen Fragen umzuwandeln und vermeiden somit, dass sie die Fragen zu einem späteren Zeitpunkt noch einmal ändern müssen. Sie sollten da keine Ausnahme machen. Fällt es Ihnen schwer, die einzelnen Komponenten einer Frage auszutüfteln, können Sie Ihr Problem zunächst als freie Frage niederschreiben. Wie ■ Tabelle 2.2 beispielhaft zeigt, lässt sich die freie Frage anschließend in eine strukturierte Frage umwandeln. Die *Population* beschreibt dabei die Teilnehmer- oder Patientengruppe, für die Ihr Review die Evidenz liefern soll. Die *Interventionen* definieren, vereinfacht dargestellt, Maßnahmen und ihre jeweiligen Alternativen, die für Ihre *Population* in Betracht kommen. *Endpunkte* sind Kenngrößen dafür, was sich die *Population* von den *Interventionen* erhofft, z. B. die Vermeidung von Krankheit oder Tod. Und schließlich müssen Sie sich überlegen, welches *Design* die Studie braucht, um die Auswirkung Ihrer *Interventionen* fehlerfrei untersuchen zu können. So lässt sich z. B. der Effekt einer *Intervention* als „Krankheit, die durch die Intervention vermieden wurde“ messen, indem man in einer *Population* die *Endpunkte* der Gruppen einer *Population* mit und ohne *Intervention* vergleicht.

Der springende Punkt bei der Formulierung von Reviewfragen ist das strukturierte Vorgehen. Die in ■ Tabelle 2.1 dargestellte Struktur sollte aber nicht als starres Korsett gesehen werden, vielmehr lässt sie sich je nach Situation an die Erfordernisse der freien Frage anpassen. Bei einer epidemiologischen Frage zur Ätiologie lässt sich z. B. die Komponente *Intervention* problemlos gegen *Exposition* austauschen. Die Frage lautet dann, inwieweit sich die *Endpunkte* in den jeweiligen *Populationen* unterscheiden, die bestimmten Substanzen oder Risikofaktoren ausgesetzt waren oder nicht (■ Tabelle 2.2). Bei Fragen zur Genauigkeit von Screeninguntersuchungen oder diagnostischen Tests lässt sich die Komponente *Intervention* durch *Test* ersetzen und *Endpunkt* durch den *Referenzstandard*, an dem die Genauigkeit des untersuchten Tests gemessen wird (■ Tabelle 2.2). In diesem Sinne ist die vorgeschlagene Struktur variabel und auf ein breites Spektrum von Fragen anwendbar.

2.1.2 Variation von Populationen, Interventionen und Endpunkten

Haben Sie erst einmal einen Blick dafür entwickelt, wie die Fragen strukturiert sind (■ Tabelle 2.1), werden Sie erkennen, dass es sich bei einem systematischen Review um die Analyse vorhandener Studien mit den darin definierten *Populationen*, *Interventionen* und *Endpunkten* handelt. Anfangs waren Sie möglicherweise skeptisch, eine Reviewfrage auf diese Weise zu formulieren. Wenn Ihnen aber klar wird, dass sich hinter der freien Frage unterschiedliche *Populationen*, *Interventionen* und *Endpunkte* verbergen, ergeben sich am Ende wahrscheinlich doch deutlich mehr Fragen als nur diese eine. Wenn nicht, sollten Sie genau prüfen, inwieweit jede einzelne Fragekomponente Variationen zulässt. Dies ist ein wichtiger Aspekt. Selbst bei so unkomplizierten Fragen wie der Behandlung chronischer Wunden mit antimikrobiellen Präparaten (■ Tabelle 2.2) sollte klar sein, dass es zahlreiche Arten von chronischen Wunden (*Populationen*), antimikrobiellen Präparaten (*Interventionen*) und Möglichkeiten gibt, die Wundheilung zu messen (*Endpunkte*) (■ Tabelle 2.3).

Freie Frage: Sie beschreibt in sprachlich einfacher Form (wie vage auch immer) das Anliegen, für das Sie in Ihrem Review eine Lösung suchen.

Strukturierte Frage: Der Reviewer überführt eine freie Frage nach einem strukturierten Verfahren in eine klare, explizite Form

In diesem Buch befassen wir uns im Kontext **vergleichender Studiendesigns** hauptsächlich mit Fragen zu den **quantitativen** Effekten von *Interventionen* (Therapie, Prävention, Sozialversorgung etc.) bzw. *Expositionen* (Umweltfaktoren, Risikofaktoren etc.).

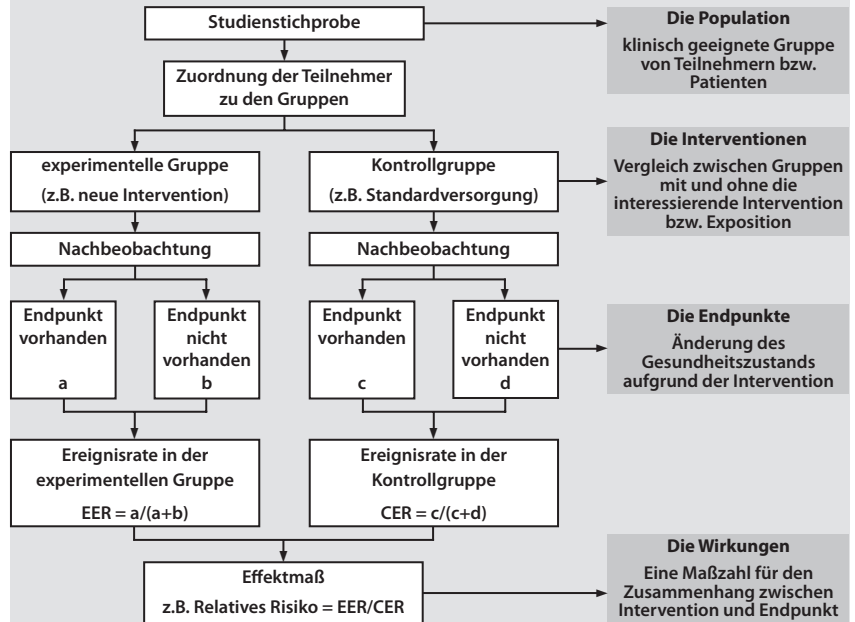
■ **Tabelle 2.1.** Formulierung von strukturierten Fragen für systematische Reviews

Fragekomponenten

- Die Populationen Die präzise Beschreibung einer Gruppe von Teilnehmern oder Patienten, des klinischen Problems sowie der Versorgungsebenen (z.B. primär / sekundär / tertiär, ambulant / stationär)
- Die Interventionen (bzw. Expositionen) Die wichtigste(n) in Betracht gezogene(n) Maßnahme(n), z.B. Therapien, Versorgungsabläufe, soziale / pädagogische Interventionen, Risikofaktoren, diagnostische Tests etc.
- Die Endpunkte Die klinischen Änderungen des Gesundheitszustands (Morbidität, Mortalität) sowie andere damit verbundene Änderungen, z.B. Nutzung der Ressourcen im Gesundheitswesen
- Das Studiendesign Das angemessene Vorgehen bei der Rekrutierung von Teilnehmern bzw. Patienten für eine wissenschaftliche Studie, bei der Verabreichung von Interventionen sowie der Auswertung ihrer Ergebnisse.

Beziehung zwischen den Fragekomponenten in einer Vergleichsstudie

Eine Vergleichsstudie bewertet die Wirksamkeit einer *Intervention* (oder *Exposition*) anhand von Vergleichsgruppen. Beispielsweise können Studienteilnehmer oder Patienten aus einer relevanten *Population* (mit oder ohne Randomisierung) alternativen *Interventions-* (oder *Expositions-*) Gruppen zugeteilt und nachbeobachtet werden, um die Auswirkung der *Interventionen* (bzw. *Expositionen*) auf die *Endpunkte* zu bestimmen.



Siehe dazu das entsprechende Studiendesign in Tabelle 2.4.

■ **Tabelle 2.2.** Einige Beispielfragen

Eine Frage zur klinischen Effektivität

Freie Frage: Welches der zahlreichen antimikrobiellen Präparate fördert bei Patienten mit chronischen Wunden die Heilung?

Strukturierte Frage

- Die Population Fördern bei Erwachsenen mit verschiedenen Formen chronischer Wunden im ambulanten Bereich
- Die Interventionen systemische oder topische antimikrobielle Präparate
- Die Endpunkte die Wundheilung?
- Das Studiendesign eine Vergleichsstudie, in der Patienten mit chronischen Wunden den zu untersuchenden Behandlungsalternativen zugeteilt werden und die Wirkung dieser Interventionen auf die Wundheilung bestimmt wird (z.B. randomisierte, kontrollierte Studie).

Zu einem entsprechenden Review siehe Fallstudie 3.

Ein Beispiel zur Ätiologie

Freie Frage: Gibt es eine Beziehung zwischen der Exposition mit Benzodiazepinen während der Schwangerschaft und Missbildungen bei Neugeborenen?

Strukturierte Frage

- Die Population Verursacht bei Schwangeren
- Die Exposition eine Einnahme von Benzodiazepinen in der Frühschwangerschaft...
- Die Endpunkte Missbildungen beim Neugeborenen?
- Die Studiendesigns
 - eine Studie, die Frauen in einem frühen Schwangerschaftsstadium rekrutiert, ihre Einnahme von Benzodiazepinen (*Exposition*) erfasst, diese Frauen nachbeobachtet und ihre Neugeborenen untersucht, um die Missbildungsrate der Kinder von Frauen mit – im Vergleich zu Frauen ohne - Exposition zu bestimmen (Kohortenstudie)
 - eine Studie, die die Einnahme von Benzodiazepinen in der Frühschwangerschaft retrospektiv bei Frauen mit einem missgebildeten Kind – im Vergleich zu Frauen mit einem gesund geborenen Kind – untersucht (Fall-Kontroll-Studie).

Zur entsprechenden Meta-Analyse siehe Tabelle 2.20.

Eine Frage zur Testgenauigkeit

Freie Frage: Lässt sich bei postmenopausalen Frauen mit abnormer Vaginalblutung durch Ultraschall des kleinen Beckens ein Uteruskarzinom exakt ausschließen?

Strukturierte Frage

- Die Population Kann bei Frauen mit postmenopausaler Vaginalblutung in einem ambulanten Setting (Primärversorgung)
- Der Test durch Ultraschall des kleinen Beckens
- Der Referenzstandard die histologisch gesicherte Diagnose „Uteruskarzinom“ exakt vorhergesagt werden ?
- Das Studiendesign eine Studie, die Frauen aus einer relevanten Population rekrutiert, Test (Ultraschall) plus Referenzstandard anwendet, um das Vorliegen eines Karzinoms (Histologie) zu bestätigen oder auszuschließen, und die Genauigkeit bestimmt, mit der der Test das Karzinom nachweisen kann (siehe Tabelle 3.13).

Zu einem entsprechenden Review siehe Fallstudie 4.

■ **Tabelle 2.3.** Formulierung von Reviewfragen: Variationen in den *Populationen, Interventionen, Endpunkten und Studiendesigns*

Fragebeispiel zur klinischen Effektivität

Freie Frage: Welches der zahlreichen antimikrobiellen Präparate fördert bei Patienten mit chronischen Wunden die Heilung?

Strukturierte Frage (in Erweiterung von Tabelle 2.2)

— Die Population	Erwachsene mit verschiedenen Formen chronischer Wunden:	<ul style="list-style-type: none"> – diabetische Ulzera – venöse Ulzera – Druckulzera
— Die Interventionen	antimikrobielle Präparate: <i>versus</i> Vergleichspräparat:	<ul style="list-style-type: none"> – systemische Präparate – topische Präparate – <i>versus</i> – andere Präparate
— Die Endpunkte	klinisch (verschiedene Möglichkeiten zur Messung der Wundheilung)	<ul style="list-style-type: none"> – komplette Heilung – verbleibende Wundfläche – Heilungsscores
— Die Studiendesigns	experimentelle und Beobachtungsstudien: (siehe Tabelle 2.4)	<ul style="list-style-type: none"> – randomisierte, kontrollierte Studien – experimentelle Studien ohne Randomisierung – Kohortenstudien mit zeitgleichen Kontrollen

Zu einem entsprechenden Review siehe Fallstudie 3

Ein weiteres Beispiel zur klinischen Effektivität

Freie Frage: Führen strukturierte Hausbesuchsprogramme bei älteren Menschen zu einer Besserung des Gesundheitszustands?

Strukturierte Frage:

— Die Population	ältere Menschen verschiedener Altersgruppen	<ul style="list-style-type: none"> – 65 bis 74 Jahre – 75 bis 85 Jahre – über 85 Jahre
— Die Interventionen	Hausbesuchsprogramm mit strukturierter Einschätzung von Basisfunktionen (Assessment): <i>versus</i> Vergleichsbehandlung:	<ul style="list-style-type: none"> – Intensität des Assessments – Häufigkeit des Assessments – <i>versus</i> – Standardversorgung
— Die Endpunkte	klinisch (verschiedene Möglichkeiten zur Messung von Gesundheit und der Nutzung von Ressourcen im Gesundheitssystem):	<ul style="list-style-type: none"> – Mortalität – Funktionszustand – Einweisung in ein Pflegeheim
— Das Studiendesign	experimentelle Studien: (siehe Tabelle 2.4)	<ul style="list-style-type: none"> – randomisierte, kontrollierte Studien – experimentelle Studien ohne Randomisierung

Zu einer entsprechenden Meta-Analyse siehe Tabelle 2.18.



■ **Tabelle 2.3.** Fortsetzung

Fragebeispiel zur klinischen Wirksamkeit und Kosteneffektivität

Freie Frage: In welchem Umfang lässt sich bei Patienten mit bevorstehender Hüftoperation das postoperative Infektionsrisiko durch antimikrobielle Prophylaxe verringern, und sind die Kosten dafür gerechtfertigt?

Strukturierte Frage:

■ Die Population	Patienten, bei denen ein Hüftgelenkersatz ansteht:	– verschiedene Verfahren
■ Die Interventionen	antimikrobielle Prophylaxe: <i>versus</i> Vergleichsbehandlung:	– verschiedene Arten von Antibiotika – <i>versus</i> – Placebo – keine Antibiotika
■ Die Endpunkte	klinisch: ökonomisch:	– postoperative Infektion – Kosten pro vermiedener Infektion
■ Das Studiendesign	klinisch: ökonomisch:	– experimentelle Studien (siehe Tabelle 2.4) – Kosteneffektivitätsanalysen

Eine entsprechende Bewertung der Studienqualität siehe Tabelle 2.12.

Deshalb müssen Sie sorgfältig überlegen, in welchen Aspekten sich *Populationen*, *Interventionen* und *Endpunkte* von Studie zu Studie unterscheiden könnten. Solcherlei Unterschiede sind sowohl für die Definition der Studienauswahlkriterien (Schritt 2) als auch für die Planung der Tabellarisierung der Studienergebnisse (Schritt 4) von großer Bedeutung. Anhand der Unterschiede fällt es Ihnen auch leichter zu verstehen, warum die Effekte der *Interventionen* von Studie zu Studie unterschiedlich ausfallen können (Schritt 4), und zu überprüfen, inwieweit die Studienergebnisse anwendbar sind (Schritt 5). Das bedeutet auch, dass die Schlussfolgerungen der einzelnen Studien und Reviews voneinander abweichen können, je nach dem, wie sehr sich die Merkmale ihrer *Populationen*, die Art oder Verabreichung ihrer *Interventionen* sowie die gewählten *Endpunkte* voneinander unterscheiden. Wir werden diese Aspekte an anderer Stelle noch im Detail erörtern. Hier wollen wir lediglich kurz auf die Konsequenzen für die Formulierung der Reviewfragen eingehen.

Studienpopulationen können sich bezüglich Alter und Geschlecht der Patienten, Schweregrad der Erkrankung, Vorliegen von Komorbidität etc. voneinander unterscheiden. Bei einer Untersuchung zur Auswirkung von Hausbesuchen bei älteren Menschen (■ Tabelle 2.3) beispielsweise war die *Intervention* bei den 65- bis 74-jährigen wirksamer als bei über 85-jährigen Studienteilnehmern. Ebenso können auch die Charakteristika der *Interventionen* (z. B. Versorgungsumgebung, Compliance oder Intensität, sonstige Routineversorgung etc.) mit jeweils unterschiedlichen Effekten verbunden sein. So sind etwa Hausbesuchsprogramme bei älteren Menschen wirksamer, wenn die formale Einschätzung von Fertigkeiten des täglichen Lebens („Assessment“) in mehreren Dimensionen und/oder mit häufigen Nachbeobachtungen durchgeführt wird.

Für die Bewertung der Wirksamkeit muss man alle klinisch relevanten und wichtigen *Endpunkte* identifizieren, mit denen sich Erfolg oder Misserfolg einer *Intervention* messen lassen. Bei der Erstellung Ihres Reviews könnte sich herausstellen, dass in den vorhandenen Studien nicht die *Endpunkte* verwendet wurden, die Ihnen selbst als relevant erscheinen. An sich ist bereits die Aufdeckung solcher Unzulänglichkeiten in einer Studie von Bedeutung. Manchmal ist es jedoch schwierig, die klinisch wichtigen *Endpunkte* zu erheben. In diesen Situationen könnte man mit sog. intermediären, Surrogat- oder Ersatz*endpunkten* arbeiten. Will man beispielsweise die Wirkung der Fluoridtherapie zur Prävention von Frakturen aufdecken, könnte man auf die Idee kommen, die Knochendichte als einen Surrogat*endpunkt* zu untersuchen, da die entsprechenden Daten relativ leicht zugänglich sind. Wie irreführend ein solches Vorgehen aber sein kann, lässt sich an einer im *N Engl J Med* 1990; 322: 802–9 veröffentlichten randomisierten, kontrollierten Studie nachweisen: Bei den mit Fluorid behandelten Studienteilnehmern nahm zwar die Knochendichte signifikant zu (im Vergleich zu Placebo und je nach Messort zwischen 10 % und 35 %), die Häufigkeit nicht-vertebraler Frakturen stieg jedoch unerwartet um das nahezu Dreifache an (Kontrollgruppe 24 vs. Fluoridgruppe 72, $p = 0,01$). Dieses Beispiel macht deutlich, dass Schlussfolgerungen aus wissenschaftlichen Untersuchungen, die sich auf Surrogat*endpunkte* stützen, in die Irre führen können und damit als Grundlage für Entscheidungen in der Patientenversorgung weniger zuverlässig sind.

Bei der Festlegung der *Endpunkte* für eine Reviewfrage muss man überlegen, was man unter Gesundheit verstehen möchte. Bedeutet Gesundheit lediglich die Abwesenheit von Krankheit oder Leiden? In diesem Buch wollen wir uns hauptsächlich auf quantitative *Endpunkte* wie Morbidität oder Mortalität konzentrieren. Außerdem beobachtet man immer öfter, dass sich Reviews mit der Frage befassen, wie sich optimale klinische *Ergebnisse* mit möglichst geringem Einsatz von Ressourcen erzielen lassen. Mit diesem Ansatz lässt sich erfassen, inwieweit sich die Investition in die jeweiligen *Interventionen* lohnt. Die *Endpunkte* müssen sich neben den klinischen *Ergebnissen* also auch mit den Kosten für die Gesundheitsversorgung befassen (■ Tabelle 2.3). In unserem Buch wollen wir dieser Problematik jedoch nur insoweit nachgehen, als sie die Formulierung der Reviewfragen berührt.

2.1.3 Variationen im Studiendesign

Glaubwürdige (valide) Ergebnisse sollen unverzerrt, d.h. frei von **Bias** sein. **Bias** kann zur Über- oder Unterschätzung des „wahren“ Effektes einer *Intervention* oder *Exposition* führen.

Die **Qualität** einer Studie ist davon abhängig, wie sehr *Design*, Durchführung und Analyse die Anfälligkeit einer Studie für **Bias** minimieren.

Wenden wir uns nun dem vierten Bestandteil einer Reviewfrage zu, dem *Studiendesign* (■ Tabelle 2.1). Reviews sind eine Zusammenfassung der Ergebnisse unterschiedlich *angelegter* Studien für definierte *Populationen*, *Interventionen* und *Endpunkte* (■ Tabelle 2.2). Weshalb spielt das *Design* dabei eine so wichtige Rolle? Nun, das *Studiendesign* determiniert die Validität der beobachteten Effekte. Anders ausgedrückt: Das Vertrauen, dass die Ergebnisse einer Studie den „wahren Effekt“ bei den untersuchten Patienten oder Studienteilnehmern auch möglichst genau wiedergeben, hängt wesentlich davon ab, wie solide das *Design* dieser Studie ist. So betrachtet lässt sich das *Design* als Qualitätsindikator verstehen. Man kann die Bedeutung dieser Komponente gar nicht stark genug betonen. Denn wie zuverlässig die Schlussfolgerungen eines Reviews sind, hängt entscheidend von der Qualität des *Studiendesigns* der eingeschlossenen Studien ab.

Manche Reviewer halten bestimmte *Studiendesigns* für besser, weil sie glauben, dass das *Design* einen Wert an sich hat. Bei der Erstellung eines Reviews berücksichtigen sie daher ausschließlich randomisierte Studien. Ein solches Vorgehen lässt jedoch die Tatsache außer Acht, dass unterschiedliche Fragestellungen auch unterschiedliche *Studiendesigns* erforderlich machen können. So setzt beispielsweise eine Frage zur Testgenauigkeit ein *Studiendesign* voraus, bei dem a.) sämtliche geeignete Patienten prospektiv (ohne Randomisierung) rekrutiert werden, b.) der *Test* und ein entsprechender *Referenzstandard* angewendet werden, um das Vorliegen einer Erkrankung zu bestätigen bzw. auszuschließen, und bei dem c.) die Genauigkeit bestimmt wird, mit der die Erkrankung durch diesen Test korrekt identifiziert wird (siehe Fallstudie 4). Zur Bewertung später auftretender oder seltener *Endpunkte*, insbesondere bei Untersuchungen zu den Risiken von *Interventionen* (wie in Fallstudie 2) bietet sich ein beobachtendes *Studiendesign* eher an als ein experimentelles. So lässt sich etwa die Auswirkung einer Einnahme von Benzodiazepinen in der Schwangerschaft (*Exposition*) auf selten vorkommende Missbildungen des Neugeborenen besser anhand von Kohorten- oder Fall-Kontroll-Studien erfassen als durch randomisierte Studien.

Und selbst bei Fragen zur Wirksamkeit von *Interventionen*, bei denen man im Allgemeinen randomisierte Studien bevorzugt, kann es in einem Review schwierig werden, die Eingrenzung auf diesen Studientyp zu rechtfertigen. In erster Linie trifft dies auf ethisch nicht vertretbare Studien zu. Manchmal besteht aber auch schlicht ein Mangel an randomisierten Studien. So wurden beispielsweise für den Review über antimikrobielle Präparate zur Behandlung chronischer Wunden (Fallstudie 3) trotz umfassender Recherchen lediglich vier eindeutig randomisierte Studien aufgefunden. Daher war man gezwungen, auch andere *Designs* zu berücksichtigen. Und in Fallstudie 2 über die Risikobewertung der Trinkwasserfluoridierung waren überhaupt keine randomisierten Studien veröffentlicht worden, so dass auch hier verschiedene andere *Designs* in Betracht gezogen werden mussten. Gelegentlich werden in einem Review eine Reihe unterschiedlicher, aber verwandter Fragestellungen untersucht. Soll z. B. ein Review neben der Wirksamkeit auch die Kosteneffektivität einer *Intervention* bewerten, dann benötigt man dazu *Studiendesigns* zur ökonomischen Evaluation (■ Tabelle 2.3). Bei manchen Reviewfragen muss man also unter Umständen mit mehreren *Studiendesigns* gleichzeitig arbeiten. Diese *Designvielfalt* wirkt sich nicht nur auf die Bewertung der Studienqualität (Schritt 3), sondern auch auf die Datensynthese aus (Schritt 4).

Insistiert ein Reviewer darauf, in Reviews nur randomisierte Studien ein- und damit gleichzeitig alle anderen Arten von Evidenz auszuschließen, kann die Erstellung des Reviews daran praktisch scheitern, wenn sich nämlich für einen Review keinerlei randomisierte Studien finden lassen. Vor die Entscheidung gestellt ist es in der Praxis allemal besser, die beste verfügbare Evidenz zu verwenden als gar keine. Sie müssen sich daher über die Art Ihrer Fragestellungen klar werden: Geht es um Wirksamkeit, Ätiologie, Kosten-Nutzen oder Testgenauigkeit? und über die Optionen, diese zu beantworten, d. h. die in Frage kommenden *Populationen*, *Interventionen* und *Endpunkte*. Danach sollten Sie die *Studiendesigns* auswählen, die sich am ehesten dazu eignen, zuverlässige Antworten zu geben, und für diese eine Hierarchie festzulegen, die zu Ihrer Reviewfrage passt. Mit diesem Vorgehen lassen sich die Ein- und Ausschlusskriterien zur Auswahl von Studien und ihre qualitativen Mindestanforderungen festlegen (Schritt 2).

Wirksamkeit (effectiveness)

bezeichnet das Ausmaß, in dem eine *Intervention* (Therapie, Prävention, Diagnostik, Screening, Schulungen, Sozialversorgung etc.) unter gewöhnlichen Alltagsbedingungen zu positiven *Ergebnissen* (*Outcomes*) führt.

Effizienz (cost-effectiveness)

bezeichnet das Maß, in dem zwischen dem Input (Kosten) und dem Output (*Ergebnisse*) von *Interventionen* ein ausgewogenes Preis-Leistungs-Verhältnis besteht.

Jede Fragestellung hat ihre eigene *Design*hierarchie. In diesem Buch beschäftigen wir uns vorwiegend mit Fragen zu den gesundheitlichen Auswirkungen von *Interventionen* und *Expositionen*. Solche Fragen konzentrieren sich in der Regel auf einen Vergleich zwischen verschiedenen *Interventionen* bzw. *Expositionen*. Eine *Design*hierarchie für Studien, in denen diese Art Fragen angesprochen werden, ist in ■ Tabelle 2.4 aufgestellt. Beim zuverlässigsten *Studiendesign* (auch als Evidenz der Stufe I bezeichnet), werden die Teilnehmer aus einer relevanten *Population* (unter Geheimhaltung der Gruppenzuordnung) den alternativen *Interventionen* nach dem Zufallsprinzip zugeteilt. Ein solches *Design* dient der Vermeidung von Selektionsbias. Sind solche Studien zudem gut durchgeführt, rangieren sie, was die Evidenz zur Wirksamkeit von *Interventionen* betrifft, an der Spitze der *Studiendesign*hierarchie. Die nächste Stufe nehmen Studien der Evidenzstärke II ein. Experimentelle Studien, bei denen jedoch der Untersucher die Zuordnung von Patienten oder Teilnehmern kontrolliert und keine echte Randomisierung mit verdeckter Behandlungszuteilung erfolgt, haben im Allgemeinen einen höheren Stellenwert als Kohortenstudien, die ihrerseits einen höheren Rang als Fall-Kontroll-Studien einnehmen. Wie oben bereits angedeutet, sind bei vielen Reviews hochwertige experimentelle Studien schlicht nicht (Fallstudie 2) oder nur recht selten (Fallstudie 3) vorhanden. Folglich müssen Reviews mitunter auch auf der Grundlage von Studien mit weniger zuverlässigem *Design* oder durch Integration verschiedener *Studiendesigns* erstellt werden. Umfasst Ihr Review mehrere *Studiendesigns*, sind Sie gut beraten, die Qualitätsbewertung sorgfältig zu planen (Schritt 3), die Datensynthese nach *Design* und Qualität zu stratifizieren (Schritt 4), die Ergebnisse zurückhaltend zu interpretieren und sich dabei besonders auf die methodisch fundierten Studien zu stützen (Schritt 5).

■ **Tabelle 2.4.** Hierarchie der *Studiendesigns* für Fragen zur *Wirksamkeit* medizinischer Interventionen

Beschreibung des <i>Designs</i>	Evidenzstufen in Abhängigkeit von der <i>Design</i> qualität
<p>Experimentelle Studie Eine vergleichende Studie*, in der der Untersucher die Teilnehmer unterschiedlichen <i>Interventionen</i> zuordnet.</p> <ul style="list-style-type: none"> — randomisierte, kontrollierte Studie (mit verdeckter Zuordnung) Randomisierte Zuordnung der Teilnehmer zu einer <i>Interventions-</i> und einer Kontrollgruppe (z.B. Placebo oder Standardversorgung) mit Nachbeobachtung, um die <i>Ergebnis</i>unterschiede zwischen beiden Gruppen zu untersuchen. Randomisierung (mit Verdeckung der Zuordnungsfolge gegenüber dem einschließenden Studienarzt) vermeidet systematische Verzerrungen (Bias), da abgesehen von der Intervention bekannte wie auch unbekannte Ergebnisdeterminanten im Durchschnitt gleichmäßig auf beide Gruppen verteilt sind. 	<p>1</p>

■ **Tabelle 2.4.** Fortsetzung

Beschreibung des <i>Designs</i>	Evidenzstufen in Abhängigkeit von der <i>Designqualität</i> ⁺
<ul style="list-style-type: none"> — experimentelle Studie ohne Randomisierung (zuweilen fälschlich als quasi- oder pseudo-randomisierte bzw. quasi-experimentelle Studie bezeichnet) Eine Studie, in der die Zuordnung der Teilnehmer zu den verschiedenen Behandlungsalternativen vom Untersucher gesteuert wird, aber die Methode der Zuordnung nicht einer echten Randomisierung entspricht (z.B. alternierende Zuordnung oder Zuordnung nach geradem/ungeradem Datum). Mit solchen Verfahren kann die Zuordnungsfolge vor den einschließenden Studienärzten nicht geheimgehalten werden. 	} II
<p>Beobachtungsstudie mit Kontrollgruppe Eine vergleichende Studie*, in der die Zuordnung der Teilnehmer zu den verschiedenen <i>Interventionen</i> nicht durch den Untersucher erfolgt, sondern dieser lediglich beobachtet.</p>	
<ul style="list-style-type: none"> — Kohortenstudie Nachbeobachtung von Teilnehmern, die eine <i>Intervention</i> erhalten (die nicht vom Untersucher zugeteilt wurde) im Vergleich zu einer Kontrollgruppe (z.B. Teilnehmer, die keine Behandlung erhalten), um die Unterschiede in den Behandlungsergebnissen zu erfassen. 	
<ul style="list-style-type: none"> — Fall-Kontroll-Studien Vergleich der <i>Interventionsraten</i> zwischen Personen mit einem bestimmten <i>Endpunkt</i> (Fälle) und solchen ohne den <i>Endpunkt</i> (Kontrollen). 	
<p>Beobachtungsstudie ohne Kontrollgruppen</p>	} III
<ul style="list-style-type: none"> — Querschnittstudie Untersuchung der Beziehung zwischen <i>Behandlungsergebnissen</i> und anderen interessierenden Parametern (einschl. <i>Interventionen</i>), die zu einem gegebenen Zeitpunkt in einer relevanten <i>Population</i> erfasst werden. 	
<ul style="list-style-type: none"> — Vorher-Nachher-Studie Vergleich der Endpunkte bei Studienteilnehmern vor und nach einer <i>Intervention</i>. — Fallserien Beschreibung einer Anzahl von Fällen, die eine <i>Intervention</i> erhalten haben, und ihre <i>Ergebnisse</i>. 	
<p>Fallberichte Pathophysiologische Studien oder Laborforschung Expertenmeinung oder Expertenkonsens</p>	} IV
<p>* Eine Vergleichsstudie bewertet die Wirkung einer Intervention anhand von Vergleichsgruppen. (Im Flussdiagramm in Tabelle 2.1 finden Sie beispielhaft den Aufbau einer solchen Studie.)</p>	