

Ist Konsens bei der Leitlinienerstellung Nonsense?

Schwerpunkt

Hans-Konrad Selbmann, Institut für Medizinische Informationsverarbeitung, Universität Tübingen

Zusammenfassung

Kulturelle Hintergründe beeinflussen die Entwicklung von Leitlinien. In Deutschland gibt es die Klassifikation von S1 (niedrige Evidenz, keine strukturierte Konsensfindung) bis S3 (hohe Evidenz, Konsensfindung durch ein repräsentatives Gremium). Konsens ist notwendig bei der Entwicklung von Leitlinien, da jeder Schritt durch Präferenzen und Wertentscheidungen beeinflusst wird, von der Wahl der Output- und Outcome-Kriterien bis zur Rolle pathophysiologischer Modelle in der Bewertung neuer Therapien. Ein strukturierter Konsensfindungsprozess erhöht die Transparenz. Verschiedene Modelle für die Konsensfindung werden diskutiert.

Sachwörter: Konsensfindung – DELBI (Deutsches Leitlinien-Bewertungs-Instrument) – Evidenz-basierte Medizin – GRADE (Grading of Recommendations, Assessment, Development and Evaluation) – Leitlinien – HTA (Health Technology Assessment)
(Wie vom Gastherausgeber eingereicht)

Is reaching consensus on the development of guidelines nonsense?

Summary

Cultural background influences the development of guidelines. In Germany, classifications exist from S1 (low level of evidence, no structured process of consensus finding) up to S3 (high level of evidence, consensus finding by a representative body). Consensus is indispensable in the guideline development process because every step is influenced by preferences and values, from the choice of output and outcome criteria to the role of pathophysiological models in the assessment of new therapies. A structured process of consensus finding increases transparency. Different models of consensus finding are discussed.

Key words: consensus finding – DELBI (Deutsches Leitlinien-Bewertungs-Instrument) – evidence-based medicine – GRADE (Grading of Recommendations, Assessment, Development and Evaluation) – guidelines – HTA (Health Technology Assessment)
(As supplied by publisher)

Die etwas provokative Frage im Titel lässt sich ganz leicht beantworten, allerdings dauert die Begründung ein bisschen länger. Ich habe mir vorgenommen, Ihnen anhand einiger Punkte zu begründen, was ein Konsens eigentlich bewirken soll.

Konsens bezüglich der Aufgaben einer Leitlinie

Nach der Definition der Arbeitsgemeinschaft Wissenschaftlicher Medizinischer Fachgesellschaften (AWMF) sind Leit-

linien systematisch entwickelte Aussagen, die den gegenwärtigen Erkenntnisstand wiedergeben und dem behandelnden Arzt und seinem Patienten die Entscheidungsfindung für eine angemessene Behandlung in einer spezifischen Krankheitssituation erleichtern. Dieser Definition kann man entnehmen, wie in Abbildung 1 dargestellt, dass Leitlinien keine Richtlinien sein wollen, sondern Hilfen bei notwendigen Entscheidungen in bestimmten Beratungs- oder Behandlungssituationen sind, in der Arzt und Patient wegen eines individuellen Gesundheitsproblems

miteinander agieren müssen. Wenn in die Entscheidungshilfe das externe Wissen explizit integriert ist, nennt man das Ganze wohl „Evidence based Medicine“ im Sinne von David Sackett.

Beim externen Wissen spielen die Evidenzen, also die wissenschaftlichen Belege, eine große Rolle. Dazu gehören Belege sowohl aus der empirischen Forschung z. B. aus klinischen Studien als auch aus der theoretischen Forschung z. B. in Form von pathophysiologischen Modellen. Solche wissenschaftlichen Nachweise fasst man entweder in Leitlinien zusammen oder präsentiert sie in Evidenzreports oder, was sicher die Zukunft ist, in Wissensbanken.

Dabei wundert mich immer wieder, dass pathophysiologische Modellvorstellungen zurzeit so abgewertet werden (letzter Grad der Evidenzstärkeskala). Ein Arzneimittel zuzulassen, ohne pathophysiologische Vorstellungen über seine Wirkungen und Nebenwirkungen zu besitzen, ist heutzutage aber gar nicht mehr möglich. Die pathophysiologische Begründung für die Wirkung einer Therapie dürfen wir auch in Zeiten der Evidence basierten Medizin nicht aus den Augen lassen.

Der kulturelle Hintergrund, vor dem die medizinische Versorgung abläuft, ist in verschiedenen Ländern ganz unterschiedlich. Lynn Payer [1] hat das in ihrem berühmten Buch „Andere Länder, andere Leiden“ beschrieben:

– Die Deutschen, denen man ja so oft das Prädikat „romantisch veranlagt“ nachsagt, neigen einerseits zu autoritären und andererseits zu pessimistischen Einstellungen. Wir sind sehr leitlinien- und richtliniengläubig, ganz allgemein obrigkeitsgläubig und vielleicht auch wissenschaftsgläubig. Auch der Pessimismus, den wir an den Tag legen, schlägt sich in

der Versorgung nieder. Man versucht hier häufig, zunächst die Extreme auszuschließen, anstatt sich zu sagen, „Was häufig ist, das kommt auch häufig vor“, und sich zunächst auf das Häufigste zu konzentrieren.

- Die Franzosen dagegen neigen zum Cartesianismus, sie wollen alles logisch begründet haben. Wenn in der Versorgung etwas logisch begründbar ist, wird es auch gemacht.
- Bei den Engländern dagegen, denen man einen ausgeprägten Empirismus zuschreibt, muss alles erst einmal empirisch erprobt worden sein. Man wartet geduldig, bis auch wirklich alles und mehrfach geprüft ist, und solange bewahrt man eben Haltung. So gesehen ist Evidenzbasierte Medizin tatsächlich ein Import aus England.
- In den USA geht man sehr viel pragmatischer vor. Dort handelt man schnell entschlossen und ist stark technikorientiert. Für den Amerikaner reicht oft eine einzige Studie, um einen Standard zu setzen.

Es ist daher, glaube ich, immer wichtig, wenn man über Versorgung, Konsens und Leitlinien redet, sich auch über die Versorgungskulturen in den verschiedenen Ländern Gedanken zu machen. Die Ergebnisse einer exzellenten Studie, die in Japan gelaufen ist und die Evidenzklasse 1 besitzt, auf das deutsche Gesundheitswesen zu übertragen, ist ohne weiteres medizin-wissenschaftliches Nachdenken daher nicht erlaubt.

Konsens bezüglich des Entwicklungsverfahrens einer Leitlinie

Nach der konsentierten Festlegung der Funktionen einer Leitlinie besteht der nächste Schritt der Leitlinienentwicklung darin, sich auf ein Entwicklungsverfahren zu einigen. Dazu gehören z. B. die Wahl der Output-/Outcome-Kriterien, die Entscheidung bei Unsicherheiten in der Evidenzbasierung, die Übertragung der Evidenzen in den Alltag und der Umgang mit Evidenzlücken. In Deutschland – als einziges Land in Europa glaube ich – hat die AWMF die so genannten S-Klassen für Leitlinien eingeführt (Abb. 2). Die S1-Klasse bil-



Fig. 1.

	Typ	Charakteristika der Entwicklung	Wissens-Legitimation der Methode	Legitimation für Umsetzung
S1	Handlungsempfehlungen von Experten	Gremium selektiert Keine system. Ebasierung Keine strukt. Konsensf.	niedrig	gering
S2k	Konsensbasierte LL	Gremium repräsentativ Keine system. Ebasierung Strukt. Konsensfindung	niedrig	hoch
S2e	Evidenzbasierte LL	Gremium selektiert Systematische Ebasierung Keine strukt. Konsensf.	hoch	mittel
S3	Evidenz- und Konsensbasierte LL	Gremium repräsentativ Systematische Evid.bas. Strukt. Konsensfindung	hoch	hoch

Fig. 2.

den noch keine wirklichen Leitlinien, sondern Handlungsempfehlungen, die oft von erfahrenen Meinungsbildnern verfasst werden. Sie besitzen keine systematische Evidenzbasierung, und es gibt hier auch keine strukturierte Konsensfindung. In der S3-Klasse dagegen finden sich die evidenzbasierten und konsensbasierten Leitlinien, die zudem von einem repräsentativen Gremium entwickelt wurden. Je repräsentativer das Gremium ist, desto stärker ist auch die Bereitschaft zur späteren Umsetzung dieser Leitlinie.

Warum ist in dem Prozess der evidenzbasierten Leitlinienentwicklung eine Konsensfindung nötig? Das liegt zunächst einmal daran, dass wir nicht für alles und jedes so gute wissenschaftliche Belege haben, wie wir sie gerne

hätten. Nach Schätzungen von Experten sind nur bis zu 20 Prozent der Tätigkeiten im ärztlichen Alltag mit hoher wissenschaftlicher Evidenz belegt. Selbst wenn es 40 Prozent wären, fehlten immer noch für 60 Prozent der Leistungen gute Evidenzen. Nach Neises und Windeler [2] differiert das Vorliegen guter Evidenzen zudem erheblich zwischen den medizinischen Disziplinen. Insgesamt kann man sagen, dass die guten Evidenzen in den operativen Fächern seltener sind als z. B. in der Allgemeinmedizin.

Im Entwicklungsverfahren von Leitlinien der AWMF gibt es drei Eckpunkte:

- Man braucht eine repräsentative Entwicklungsgruppe,
- man braucht eine systematische Evidenzbasierung und

– man braucht eine klinische Bewertung der Evidenzen, clinical judgement, wie die Engländer und Amerikaner sagen, an die sich eine strukturierte Konsensfindung anschließt.

„Andere Länder, andere Sitten“. Die verhältnismäßig ausgeprägte Trennung zwischen der systematischen Evidenzbasierung und der klinischen Bewertung der Evidenzen im AWMF-Verfahren ist ein wenig dadurch geprägt, dass in Deutschland die systematische Evidenzbasierung oft von Wissenschaftlern durchgeführt wird, deren Klinikerfahrung schon etwas länger zurückliegt, oder die gar keine Klinikerfahrung haben. Sie müssen die vorhandenen Studien auf ihre Aussagekraft hin bewerten, ohne den aktuellen Versorgungsalltag zu kennen.

In anderen Ländern sieht das anders aus. In Abbildung 3 ist der Leitlinienentwicklungsprozess dargestellt, wie ihn die Arbeitsgruppe „Grading of Recommendations, Assessment, Development and Evaluation“ (GRADE) sieht – von der Frage, wie soll die Leitlinie aussehen bis hin zu der Abgabe von Empfehlungsstärken. (Die Pluszeichen in der Abb. 3, die das Ausmaß von Werteabwägungen darstellen sollen, stammen von mir, deshalb steht dort auch ein em-ba = eminence based). Wie Sie sehen, sind bei jedem der Schritte ab Nummer 3 Konsensfragen zu stellen: Was sind die wichtigen Outcomes? Wie bewertet man die Qualität einer Evidenz? Wie wägt man Nutzen und Schaden ab? Das GRADE-Verfahren sieht im Gegensatz zum AWMF-Verfahren vor, dass das Abwägen von Wertevorstellungen innerhalb jedes einzelnen Schrittes erfolgt. Das ist dann möglich, wenn die „Evidenzbasierer“ zugleich über den klinischen Sachverstand für die klinische Bewertung verfügen. Dies ist ein personelles Problem, das, aus der Sicht der AWMF, in Zukunft auch in Deutschland gelöst werden muss.

Konsensbedarf an einzelnen Beispielen

Lassen Sie uns einige ausgewählte Aspekte bei der Leitlinienentwicklung betrachten, bei denen ein Konsens not-

wendig ist und beginnen wir mit der Wahl der Output- und Outcome-Kriterien.

Output- und Outcome-Kriterien

Woran messen wir den Nutzen einer Versorgung? Eine Theorie, die auf John Overtveit zurückgeht, geht von drei Interessentengruppen im Gesundheitswesen aus: den Patienten, den Leistungserbringern und den Kostenträgern bzw. Managern. Wie in Abbildung 4 dargestellt, hat jede dieser Agentengruppen unterschiedliche Vorstellungen davon, was ihr am wichtigsten ist. Der Kunde, unser Patient, will wissen, wie

es ihm nach der Behandlung geht. Der Leistungserbringer, der in der Regel keinen Werkvertrag, sondern einen Dienstleistungsauftrag hat, hält sich primär an den Behandlungsprozess, der dem gegenwärtigen Stand des Wissens entsprechen muss. Der Manager muss dagegen über die Ressourcen und über die Opportunitätskosten nachdenken. Die Zuordnung der drei Agentengruppen zu der Triade „Struktur, Prozess, Ergebnis“ von Avedis Donabedian ist unschwer zu erkennen. Aber selbst wenn man sich für eine dieser Qualitätsdimensionen entschieden hat, hört die Abwägung von Wertevorstellungen nicht auf. Dies lässt sich gut am Beispiel

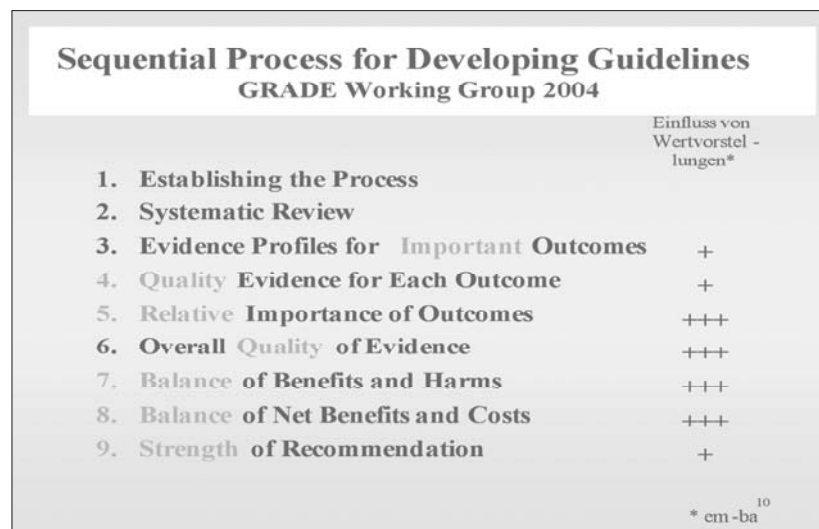


Fig. 3.

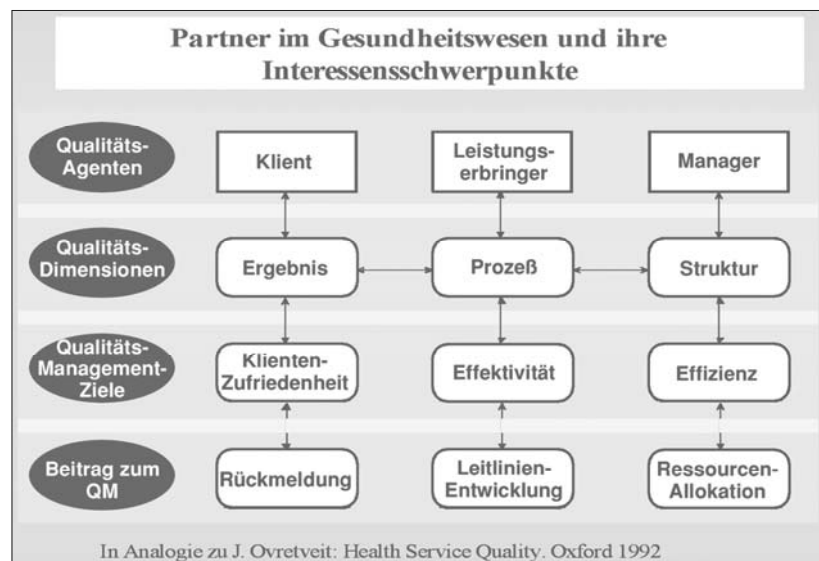


Fig. 4.



Wirksamkeit der hyperbaren Sauerstofftherapie bei Schädel-Hirn-Trauma-Patienten		
Nach einem Jahr	HbO-Gruppe	Kontrollen
Mortalität (0.04)	14 (17%)	26 (32%)
Schwere Morbidität	26 (31%)	12 (14%)
Gut erholt bzw. leicht behindert	44 (52%)	44 (54%)
100%	84	82

Rockswold et al.: Results of a prospective randomized trial for treatment of severely brain-injured patients with hyperbaric oxygen.

Fig. 5.

der Wirksamkeit der hyperbaren Sauerstofftherapie (HBO) nach Schädel-Hirn-Traumen beschreiben (Abb. 5).

Nach der besten verfügbaren Studie zu diesem Thema halbiert die HBO die Mortalität, lässt aber die leichte Morbidität bzw. die Wiederherstellung der Patienten unverändert. Wie soll hier zwischen Mortalität und Lebensqualität abgewogen werden, wenn es um die zukünftige Finanzierung der HBO bei Schädel-Hirn-Trauma-Patienten in der GKV geht und wer soll das tun? Eine solche Entscheidung kann man eigentlich nicht Methodikern oder Vertragspartnern überlassen; hier bedarf es eines Konsenses der Solidargemeinschaft. Natürlich müssen wir dabei die Wertevorstellungen der Patienten mit betrachten. Patienten interessieren sich nicht nur für die klassischen traditionellen Endpunkte wie Überlebensrate, rezidivfreie Zeit oder biochemische Variablen, sondern auch für Dinge wie Lebensqualität, Erfüllung ihrer Erwartungen oder soziales Stigma. Zwar fordern die Verfahrensordnungen des G-BA und des IQWiG, dass nicht ausschließlich die Mortalität, sondern auch die Lebensqualität bei der Bewertung berücksichtigt werden soll. Aber wenn man viele systematische Reviews und Meta-Analysen betrachtet, bauen die doch oft nur auf der Mortalität auf. Das liegt einfach daran, dass viele Studien nur die Mortalität als vergleichbare Endpunkte haben, die Lebensqualität dagegen mit vielerlei Maß – es gibt etwa

200 verschiedene Lebensqualitätsindizes – gemessen wird.

Entscheidung bei Unsicherheit in der Evidenzbasierung

Die Formulierung „Entscheidung bei Unsicherheit in der Evidenzbasierung“ wird Sie vielleicht verwundern. Die Unsicherheit entsteht dadurch, dass es beispielsweise Studien in ganz unterschiedlicher methodischer Qualität gibt. Als Produzent von Evidenzen – als an klinischen Studien beteiligter medizinischer Statistiker – weiß ich, dass es keine perfekte Studie gibt. Alle haben ihre Probleme. Mal sind die vergleichenden Behandlungen nicht adäquat, wenn zum Beispiel falsche Dosierungen eingesetzt werden, mal haben Rando-

misierung oder Verblindung nicht geklappt, mal sind Dropout-Raten zu groß oder unbalanciert, mal ist die Auswertung verbesserungsbedürftig. Wie geht man zum Beispiel bei der Auswertung mit fehlenden Werten um?

Je zahlreicher die Abweichungen von der Idealstudie sind, desto stärker könnten die Verzerrungen der Studienaussagen sein, müssen aber nicht. Es stellt sich also die Frage, wie mit den Ergebnissen von Studien mit Verzerrungen umzugehen ist, wenn nur diese scheinbar verzerrten Studien existieren. Oft ist es gerade so, dass dort, wo die beste verfügbare Evidenz nur mäßig ist, Ärzte von Leitlinien eine Hilfestellung bei ihren Entscheidungen erwarten.

Extrapolation von Studienergebnissen in den Alltag

Die Übertragung von Studienergebnissen in den Versorgungsalltag ist ein weiterer kritischer Punkt. Stellen Sie sich eine typische Health Technology Assessment (HTA)-Frage vor: Ist das Verfahren A dem Standardverfahren B vorzuziehen und unter welchen Umständen? In Abbildung 6 stellt der große Kreis den beanspruchten Gültigkeitsbereich des HTA-Berichts dar.

Normalerweise gibt es viele verschiedene Studien (kleine Kreise) dazu, deren Gültigkeitsbereiche zusammen keineswegs den angestrebten Gültigkeitsbereich des HTA-Berichts überdecken. Wenn es zu Überschneidungen der Gültigkeitsbereiche der kleinen Studien kommt, könnte man Meta-Analysen

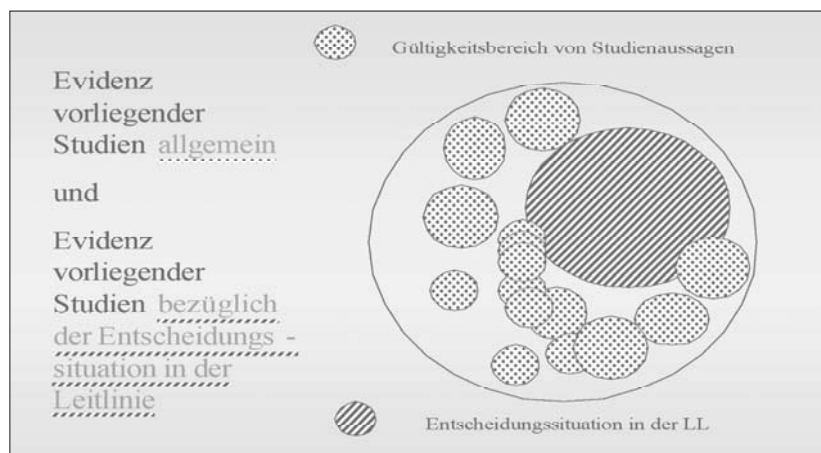


Fig. 6.

anfertigen. Aber es bleiben auch viele Lücken bestehen, also Entscheidungssituationen, die durch die Gültigkeitsbereiche der verschiedenen kleinen Studien und Meta-Analysen nicht abgedeckt sind. Der Kreis mittlerer Größe in Abbildung 6 stellt den Gültigkeitsbereich der Leitlinie dar. Er fasst die Entscheidungssituationen des Alltags zusammen. Aber auch hier zeigt es sich oft, dass sich die existierenden Studien nicht ohne Probleme auf die jeweilige Entscheidungssituation übertragen lassen. Kann man die Ergebnisse der Studien auf den Gültigkeitsbereich der Leitlinie extrapolieren und wenn ja, wie? Mit Sicherheit wird es nicht ohne pathophysiologische oder andere Modellvorstellungen gehen. Hier öffnet sich ein weites Forschungsfeld für die Medizin.

In diesem Zusammenhang möchte ich noch ein weiteres, bekanntes Phänomen ansprechen: die Relevanz des zu erwartenden Outcomes unter Alltagsbedingungen. Studienergebnisse sind – eben wegen der Studienbedingungen – häufig sehr viel besser als die Ergebnisse, die später im klinischen Alltag erzielt werden. Um die Alltagsrelevanz beurteilen zu können, liegen häufig jedoch keine Studien mit hoher Evidenzstärke vor. Auch hier ist ein Konsens ebenso gefragt wie für die Übertragbarkeit der Ergebnisse auf bestimmte Patienten und Anwenderzielgruppen.

Abwägen zwischen Outcomes mit unterschiedlicher Evidenzstärke

Probleme gibt es auch bei der Abwägung zwischen verschiedenen Outcomes, für die Studien unterschiedlicher Evidenzstärken existieren. Nehmen wir die Situation, dass die Wirksamkeit einer Therapie mit randomisierten klinischen Studien belegt ist, aber es zur Frage der unerwünschten Nebenwirkungen nur Beobachtungsstudien gibt. Welche Gesamtevidenzstärke lässt sich damit für den Einsatz der Therapie festlegen? All dies sind Punkte, über die ein Konsens hergestellt werden muss.

Strukturierte Konsensfindung

Es gibt einen Health Technology Assessment Bericht des NHS in Großbritannien aus dem Jahre 1998 über den Einsatz von Konsensfindungstechniken bei der Leitlinienentwicklung [5].

Die Autoren weisen darin auf fünf wesentliche Voraussetzungen für eine Konsensfindung in Gruppen hin. Zum einen ist die Zahl der Beteiligten wichtig – wenn eine größere Gruppe an der Konsensfindung mitwirkt, ist die Wahrscheinlichkeit für die Richtigkeit der Entscheidung sicher größer, als wenn Einzelpersonen die Entscheidung treffen. Zum zweiten spielt die Gruppenautorität eine Rolle. Wichtig sind drittens die Ermöglichung einer rationalen Argumentation und Diskussion während der Konsensfindung sowie viertens ein nach formalen Regeln gesteuerter Konsensfindungsprozess. Fünftens bleibt schließlich noch die Feststellung der Autoren, dass formale Konsensfindungsmethoden den Anforderungen genügen, die man an wissenschaftliche Methoden stellt. Eine Aktualisierung dieses HTA-Berichts unter gleichzeitiger Berücksichtigung der neueren AWMF-Erfahrungen wäre sicher sinnvoll.

Das Hauptziel der strukturierten Konsensfindung der AWMF ist das Erlangen angemessener, manipulationsfreier und reproduzierbarer Empfehlungen für anstehende Entscheidungen. Dazu gehört, dass man von vornherein die

richtigen Leute an der Konsensfindung beteiligt: Experten, Anwender und Patienten. Diese Gruppe muss auf Seiten der Experten und Anwender interdisziplinär und multiprofessionell zusammengesetzt sein, und jedes einzelne Mitglied muss eine möglichst große Unabhängigkeit besitzen. Wir wissen aber, dass es weltweit keine unabhängigen Experten gibt – entweder ist jemand unabhängig oder er ist Experte. Daher wird versucht, durch die Pluralität der Abhängigkeiten und einen strukturierten Dialog den Konsens ausgewogen und stabil zu machen. Natürlich muss dazu jedes potenzielle Mitglied der Konsensgruppe seine Abhängigkeiten der Führungsgruppe der Leitlinie offenbaren.

Bei der Methode der Konsensfindung sind drei verschiedene Techniken international akzeptiert: die nominale Gruppentechnik, die Delphi-Methode und die strukturierte Konsensuskonferenz. An der nominalen Gruppentechnik sind etwa zwölf bis 15 Personen beteiligt, während es bei der postalisch ablaufenden Delphi-Methode auch einmal 200 Mitwirkende sein können. Die strukturierte Konsensuskonferenz, bei der sich die Beteiligten strukturiert mittels TED-Abstimmung zu bestimmten Entscheidungsthemen äußern, funktioniert mit 50 bis 60 Teilnehmern am besten. Alle drei Methoden zur strukturierten Konsensfindung haben sich bewährt. Allerdings soll nicht verschwiegen werden, dass diesbezüglich

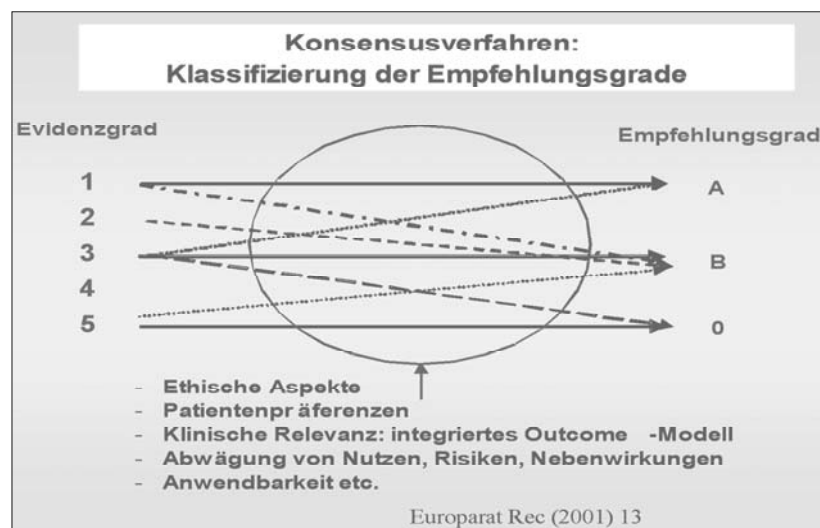


Fig. 7.

in Deutschland noch ein großer Forschungs- und ein Überzeugungsbedarf existieren.

Was ist die Aufgabe dieser das Entwicklungsverfahren abschließenden Konsensfindung? Es geht darum, die ungelösten Bewertungsprobleme aus dem Bereich der systematischen Evidenzbewertung und der Übertragung der Evidenzen auf die angezielten Leistungserbringer und Patienten aufzuarbeiten, die Konsensstärke zu messen und schließlich Empfehlungsgrade – z.B. starke Empfehlung (A), Empfehlung (B) und Option (O) – zur Kennzeichnung der Aussagen in der Leitlinie zu vergeben. Dabei kann z.B. passieren, wie es die Abbildung 7 aus dem Report des Europarates (6) zeigt, dass randomisierte Studien (Evidenzstärke 1) beim Empfehlungsgrad B (Empfehlung) landen, weil z.B. die Studienergebnisse nicht direkt auf das deutsche Gesundheitssystem übertragbar sind oder es sich um indirekte Vergleiche handelt. Umgekehrt kann eine Kohortenstudie (Evidenzstärke 3) zum Empfehlungsgrad A aufgewertet werden, weil die Konsens Teilnehmer mit einem sehr starken Konsens der Studienaussage eine besonders hohe Wirksamkeit im ärztlichen Alltag attestieren. Letztendlich werden die Empfehlungsgrade, ähnlich wie bei Koalitionsverhandlungen, im Konsens verabschiedet, in dem man das Risiko, einer Verzerrung aufzusitzen, gegen den potenziellen Nutzen für die aktuellen Patienten abwägt.

Transparenz der Leitlinienentwicklung

Noch ein Wort zur Transparenz der Leitlinienentwicklung, die gefordert wird, um die Qualität der Leitlinienentwicklung überprüfen zu können. Für diese Überprüfung gibt es seit Juni 2005 ein Werkzeug, das früher die deutsche Übersetzung des internationalen Instruments AGREE (Appraisal of Guidelines for Research and Evaluation in Europe) war und heute DELBI (DEutsches Leitlinien-Bewertungs-Instrument www.delbi.de) heißt. Das von der AWMF und dem ÄZQ entwickelte DELBI umfasst die früheren AGREE-Kriterien – also den Geltungsbereich der Leitlinie, die Betei-

ligung von Interessengruppen, die Evidenzbasierung und Konsensfindung, die klare Gestaltung und Anwendbarkeit sowie die redaktionelle Unabhängigkeit der Teilnehmer – und ist um eine neue Domäne ergänzt. Diese enthält Empfehlungen für die sektorenübergreifende Versorgung, etwa Aussagen zu den Nahtstellen zur ambulanten Versorgung, wenn es um Krankenhausleitlinien geht. Weiterhin wird erfragt, ob in den Leitlinien auch negative Empfehlungen (Unterlassungsempfehlungen) enthalten sind. Zu den neuen DELBI-Kriterien gehört ferner die Orientierung am Behandlungsablauf, für den Algorithmen formuliert sein können, die Strategie für die Verbreitung und Hilfen für die Implementierung der Leitlinie. Das letzte und wichtigste neue DELBI-Kriterium betrifft die Existenz eines Berichts über die Entstehung der Leitlinie. Dieser sogenannte Leitlinienreport ist Grundvoraussetzung für die Überprüfung der methodischen Qualität der Leitlinienentwicklung und macht zugleich die Entstehung der Leitlinie transparent. Insbesondere der Prozess der Konsensfindung wird so nachvollziehbarer: Wer war dabei, wie ist das gelaufen, wie wurde abgestimmt, und dergleichen mehr.

Der Leitlinienreport dient also als Grundlage für die Messung der internen Validität der Leitlinie. Ob jedoch das, was am grünen Tisch entwickelt worden ist, sich in der Praxis auch bewährt, ist mit DELBI nur zu vermuten. Die Effektivität und Effizienz einer Leitlinie kann erst nach ihrer Implementierung, entweder in einer vergleichenden Studie oder in einer Kohortenstudie im Alltag, überprüft werden.

Schlussbemerkungen

Ich habe versucht, Ihnen aufzuzeigen, wo überall Konsens bei der Leitlinienentwicklung notwendig ist. Wir müssen uns darüber im Klaren sein, dass eine Leitlinie keine Richtlinie sein will. Wir müssen einen Konsens herstellen, wie wir bei der Leitlinienentwicklung vorgehen wollen. Wir müssen uns verdeutlichen, welche patientenorientierten Ergebnisse uns wichtig sind und dabei auch die Lebensqualität ernst nehmen,

obwohl sie schlecht messbar und in systematischen Reviews nicht gut erfasst ist. Es gibt noch viele weitere Punkte, die zeigen, dass das Verfahren der Evidenzbasierung selbst keiner festgeschriebenen Richtlinie folgt, sondern dass stets unerschwellige Konsensfindungstechniken mitwirken. Bei der Übertragung der Evidenzen in den Alltag und beim Umgang mit Evidenzlücken ist das, glaube ich, noch am besten zu erkennen.

Die Frage, die mir gestellt wurde, hieß:

Ist Konsens bei der Leitlinienerstellung Nonsense?

Sie lässt sich mit einem einfachen Satz beantworten:

Ohne Konsens ist bei der Leitlinienerstellung alles Nonsense.

Literatur

- [1] Payer L. Andere Länder, andere Leiden. Campus Fachbuch 1993.
- [2] Neises G, Windeler J. Wie viel ist „evidenzbasiert“? Eine Übersicht zum aktuellen Forschungsstand; Z Äztl Fortbild Qualitätssich 2001;95: 95–104.
- [3] Geraedts M, Selbmann HK. Wer sollte die Qualität der Gesundheitsversorgung definieren – Patienten, Ärzte, Krankenkassen oder Gesundheitspolitiker? In Helmich P (Hrsg.) Primärärztliche Patientenbetreuung – Lehre – Forschung – Praxis. Stuttgart, Schattauer-Verlag 1997:246–57.
- [4] Rockswold GL, Ford SE, Anderson DC et al. Results of a Prospective Randomized Trial for Treatment of Severely Brain-Injured Patients with Hyperbaric Oxygen. J Neurosurgery 1992;76:929–34.
- [5] Murphy MK, Black NA, Lamping DL et al. Consensus Development Methods, and their Use in Clinical Guideline Development. Health Technology Assessment 1998;2(3):83.
- [6] Europarat Rec (2001). 13. Entwicklung einer Methodik für die Ausarbeitung von Leitlinien für die optimale medizinische Praxis. Z Äztl Fortbild Qualitätssich 2002;96 Suppl 3:1–60.

Korrespondenzadresse:

Prof. Dr. Hans-Konrad Selbmann
Institut für Medizinische Informationsverarbeitung
der Universität Tübingen
Westbahnhofstr. 55, D-72070 Tübingen
Tel.: (+49) 07071 - 2985218
Fax: (+49) 07071 - 49527
E-Mail: Hans-Konrad.Selbmann@med.uni-tuebingen.de