

Prof. Dr. rer. oec. Axel C. Mühlbacher
ist Professor für Gesundheitsökonomie und Medizinmanagement an der Hochschule Neubrandenburg und ist u.a. Mitglied im wissenschaftlichen Beirat von „Monitor Versorgungsforschung“

„Neue Evidenz: Patientenpräferenzen statt Lotterie“

>> Die methodischen Grundlagen der Nutzenbewertung in Deutschland werden seit der Vorstellung des ersten Entwurfs des Methodenpapieres im Jahr 2009 durch das Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) fortgeschrieben. Führte bereits der Vorschlag zur Effizienzgrenze (EG) mit dem Maß des Gesamtnutzens zu kontroversen, wissenschaftlichen Diskussionen, dann werfen die Anpassungen an die Regelungen des AMNOG im Kontext der frühen Nutzenbewertung bzw. der Bestimmung des Ausmaßes des Zusatznutzens weitere methodische Fragen auf. Im Kern geht es darum, wie das latente Konstrukt des Patientennutzens operationalisiert werden kann. Die Operationalisierung erfolgt über die Abwägung und Aggregation von positiven und negativen Effekten einer Intervention (im Vergleich mit einer Alternative).

Evidenzbasierte Medizin als Grundlage

Heute – sechs Jahre nach dem ersten Entwurf – sind die Frage- und Problemstellungen des Messens von Zielkriterien und Nutzendimensionen im Zusammenhang mit der evidenzbasierten Medizin im Methodenpapier ausführlich und gut dargestellt. Die praktische Implementierung der evidenzbasierten Medizin, die höchsten Anforderungen an das Studiendesign und die tiefgründige Analyse der Ergebnissicherheit garantieren die bestmögliche Information für die Beurteilung der kausal begründeten Effekte einer Intervention im Vergleich mit dem Status Quo oder Placebo. Das Methodenpapier des IQWiG begleitet hier einen wesentlichen Fortschritt in der evidenzbasierten Medizin (EBM). Die Messung der klinischen und nicht-klinischen Effekte ist eine notwendige Bedingung für die Entscheidung über den Einsatz, die Erstattungsfähigkeit und die Festsetzung der Preise von Gesundheitstechnologien. Es ist eine notwendige, jedoch keine hinreichende Bedingung.

Eine rationale Entscheidung ist mit diesen Informationen nicht möglich. Über das gesamte Methodenpapier bleiben Fragen zur Bewertung der gemessenen klinischen oder nicht-klinischen Effekte offen. Informationen über Werturteile und Präferenzen der Betroffenen werden bei der Gewichtung der Zielkriterien nicht systematisch berücksichtigt. Es entsteht der Eindruck, dass der Nutzen einer Intervention bei der frühen Nutzenbewertung (FNB) anders bewertet werden soll, als bei der Kosten-Nutzen-Bewertung (KNB). Wenn ja, dann stellen sich folgende Fragen: Auf Basis welcher methodischen und theoretischen Grundlagen unterscheiden sich diese Bewertungsverfahren? Welche wissenschaftlichen Annahmen und Werturteile sind Basis der unterschiedlichen Bewertung? Unklar bleibt, mit welchen Annahmen und Werturteilen das Maß des Gesamtnutzens (MdG) bzw. das Ausmaß des Zusatznutzens (ZN) in einem eindimensionalen Nutzenwert quantifiziert werden soll. Da eine konsistente wissenschaftliche Nutzenbewertung für die FNB und die KNB des Institutes für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) eine wesentliche Voraussetzung für konsistente und valide Entscheidungen über die Erstattungsfähigkeit oder Preisfestsetzung sind. Es kann von einem erheblichen Handlungsbedarf ausgegangen werden, gelten doch die Ergebnisse der Nutzenbewertung als Grundlage der Verhandlungen zwischen Hersteller und GKV-Spitzenverband. Keiner der Beteiligten wünscht sich eine „Lotterie“, wonach die Preise der Gesundheitstechnologien nicht auf Basis der Nutzenbewertung abgeleitet werden können.

Analogie

Für die Diskussion des Entwicklungspotenzials der Nutzenbewertung im Kontext der FNB und KNB im Kontext der Entscheidungen der Selbstverwaltung, soll hier die Analogie „von zwei Läufern in einem fairen Rennen“ des IQWiG-Methodenberichts aufgegriffen werden (IQWiG 2014, S. 7).

Messen der Leistung:

Anhand des Laufwettbewerbs kann eindrücklich die Problematik des Messens und der Ergebnissicherheit verdeutlicht werden. Es besteht die Möglichkeit, die Zeit oder aber auch die zurückgelegte Wegstrecke zu messen. Um einen fairen Wettkampf zu garantieren, müssen die Ausgangsvoraussetzungen für beide Läufer gleich sein. Laufen die Wettkämpfer gleichzeitig gegeneinander, braucht es noch nicht einmal eine Stoppuhr, um einen Sieger zu bestimmen (Ein Messinstrument wird erst benötigt, wenn der Wettkampf zeitversetzt stattfindet). Je näher die zwei Läufer beieinander liegen, desto genauer muss die Messung erfolgen. Je knapper das Ergebnis desto schwerwiegender ist die Ergebnisunsicherheit. Messverfahren und Ergebnisunsicherheit müssen daher transparent sein.

Entscheidung über die Platzierung:

Es können unterschiedliche Entscheidungsregeln zur Bestimmung der Platzierung, d.h. der Rangordnung, genutzt werden. Die Entscheidungsregeln haben auch Auswirkungen auf die Durchführung der Messung. Wird das Minimalprinzip für die Festlegung eines Siegers herangezogen, dann wird ein fixer Streckenabschnitt bestimmt, welchen die Läufer so schnell wie möglich zurücklegen müssen (um zu gewinnen, bedarf es der minimalen Zeit). Denkbar wäre aber auch das Maximalprinzip, wobei der Sieger des Wettkampfes die längste Strecke innerhalb eines bestimmten Zeitbudgets zurücklegen muss. Überträgt man dieses Beispiel auf die Bewertung von Gesundheitstechnologien, so stellt sich die Frage, ob die Entscheidungslogik von einem festen Budget ausgeht (Maximalprinzip) oder ob ein(e) zuvor bestimmte Leistung/Qualität/Nutzen mit dem minimalen Mitteleinsatz bewerkstelligt werden soll (Minimalprinzip). Unabhängig von der Entscheidungsregel, ist die Rangreihung – also die Feststellung des Siegers – in diesem Beispiel des Wettrennens sehr einfach, da es sich hier nur um eine Disziplin handelt. Dieses sehr einfache Entscheidungsproblem ist so in der Realität selten vorzufinden – schon gar nicht im Gesundheitswesen.

Komplexität des Entscheidungsproblems:

Das Messen der Zeit oder der Wegstrecke ist nur eine verkürzte Darstellung des Bewertungsproblems bei der Bewertung von Gesundheitstechnologien. Realistischer ist die Suche nach dem besten Athleten: hierzu bedarf es mehrerer Wettkämpfe, wie zum Beispiel bei einem Zehnkampf. Es zählen nicht nur Schnelligkeit, sondern auch andere Fähigkeiten wie Kraft, Ausdauer oder Koordination. Analog der klinischen

Realität müssen unterschiedliche Bewertungskriterien (Disziplinen) für die Auswahl des Siegers herangezogen werden, damit die Überlegenheit eines Athleten A gegenüber B bestimmt werden kann. Um einen fairen Wettkampf zu garantieren, müssen die Disziplinen (Zielkriterien) den Athleten vorab bekannt sein.

Vergleichbarkeit der Leistungen:

Man kann nicht davon ausgehen, dass einer der Wettkämpfer alle Disziplinen gewinnt. Dem Vergleich der Leistungen in allen zehn Disziplinen müssen Werturteile zugrunde gelegt werden. Aus diesem Grund ist die Basis der fairen Entscheidung, dass vor dem Vergleich die Annahmen über die Vergleichbarkeit bzw. die Methoden des Vergleichs der gemessenen Ergebnisse nachvollziehbar gemacht werden (vgl. das Punktesystem im Zehnkampf).

Gewichtung der Leistungen:

In einem weiteren Schritt ist zu klären, welcher Wert den einzelnen Disziplinen beigemessen wird. Im Falle des Zehnkampfes gehen wir davon aus, dass jede Disziplin mit dem gleichen Anteil in die Bewertung der Gesamtleistung eingeht (Was auch eine bewusste Entscheidung sein kann, dann aber einer entsprechenden Rationale bedarf). Denkbar wäre es aber auch, dass der Sprint höher gewichtet werden soll als die Langstrecke, die Langstrecke höher als das Kugelstoßen und das Kugelstoßen höher als der Speerwurf. Demnach müssen die gemessenen Leistungen entsprechend ihrer Relevanz bei der Ermittlung des Gesamtsiegers gewichtet werden. Diese Gewichtung könnte abhängig sein von den Wertmaßstäben der Athleten oder von Experten. Die Gewichtung und die Annahmen über die Vergleichbarkeit haben erheblichen Einfluss auf die Gesamtwertung, deshalb sollten die zugrundeliegenden Werturteile transparent gemacht werden. Vorab sollten ausreichend Informationen vorliegen, um über die Gewichtung der Disziplinen zu entscheiden.

Verzerrungen der Wettkampfergebnisse:

Werden die einzelnen Wettkampfdisziplinen festgelegt, nachdem der Sportler bereits trainiert hat, kann dies zu Verzerrungen und strategischen Entscheidungen führen. Sind weder dem Publikum noch dem Sportler die Annahmen über die Vergleichbarkeit oder die Gewichtung der einzelnen Disziplinen bekannt, werden die Rangplätze intransparent und ohne rationale Begründung vergeben. Vor diesem Hintergrund ist ein Training für den Sportler nicht sinnvoll, da die Rangordnung einer gewissen Beliebigkeit entspricht und durch den Athleten nicht in ausreichendem Maße beeinflusst werden kann. Weiterhin kann man davon ausgehen, dass die Beliebigkeit des Gesamtergebnisses nicht vom Publikum akzeptiert wird.

Bewertung im Gesundheitswesen:

Bei der Bewertung von Gesundheitstechnologien werden häufig mehrere Zielkriterien formuliert (z. B. Mortalität, Morbidität, Lebensqualität), durch Nutzendimensionen abgebildet (z.B. Verlängerung der Lebensdauer, Verbesserung des Gesundheitszustandes, Verkürzung der Krankheitsdauer, Verringerung der Nebenwirkung, Verbesserung der Lebensqualität) und durch patientenrelevante Endpunkte gemessen. Weder bei der Identifikation der Zielkriterien, noch bei der Gewichtung der gemessenen Zielerreichung patientenrelevanter Endpunkte berücksichtigt das IQWiG Evidenz. Damit verfügen die „Schiedsrichter" weder über wissenschaftliche Studien hinsichtlich der Prioritäten der Patientenpopulation(en), noch machen sie ihre eigene Einschätzung der Relevanz der Bewertungskriterien transparent. Fraglich ist, wie ein „fairer Wettkampf" in diesem Kontext zustande kommen kann.

Skalenniveau:

Das Ergebnis der Bewertung kann in unterschiedlichen Skalen präsentiert werden. Die Anforderung an das Skalenniveau ist abhängig von der Art der Entscheidung. Bei einem sportlichen Wettkampf muss eine Aussage getroffen werden, ob ein Unterschied zwischen Athlet A und Athlet B vorliegt (nicht in welcher Höhe). Das heißt: Eine Verortung des Bewertungsergebnisses auf einer Ordinalskala ist ausreichend. (Dem Wettkämpfer wird auf einer Ordinalskala ein Rangplatz zugewiesen: Wir kennen den Gewinner, wissen aber nicht, um wie viel besser der Gewinner ist). Soll aber ein Preisgeld entsprechend der sportlichen Leistung verliehen werden, d.h. entsprechend der Höhe des Erfolges von A über B, dann muss das gemessene und bewertete Ergebnis auf einer Kardinalskala (Intervallskala) abgetragen werden können (d.h. der Bewertungsprozess muss im Ergebnis den Abstand zwischen A und B abbilden können). Der unterschiedliche Informationsbedarf bei der Entscheidung stellt unterschiedliche Anforderungen an die Methoden und Instrumente der (Nutzen-)Bewertung, ändert aber nichts an der Forderung einer sachgerechten Auswahl der Disziplinen, der Vergleichbarkeit der Disziplinen und der Gewichtung der Wettkämpfe hinsichtlich ihrer Bedeutung.

Fazit und zukünftige Potenziale

Die Bewertung der Zielkriterien und der Vergleich der gemessenen Effekte ist ein wesentlicher Bestandteil der Nutzenbewertung, egal ob bei der Operationalisierung der gemessenen Effekte, in der frühen Nutzenbewertung oder der Kosten-Nutzen-Bewertung. Eine faire Entscheidung bei multiplen Zielkriterien ohne eine explizite Gewichtung ist unmöglich. Die Aggregation eines Maßes für den Gesamtnutzen oder die Bestimmung des Ausmaßes des Zusatznutzens beim Vergleich von Therapie A gegen B ist ohne ein Werturteil im Hinblick auf die Vergleichbarkeit und Gewichtung der Zielkriterien ebenso nicht möglich. Dieses Werturteil kann subjektiv oder objektiv getroffen werden. Werturteile können transparent oder intransparent im Entscheidungsprozess verwendet werden. Jeder Sportler trainiert im Vertrauen auf einen fairen Wettkampf. Ohne eine transparente und objektive Transformation der gemessenen Ergebnisse in eine Entscheidung über Erfolg und Misserfolg sind die Entbehrungen im Training irrational. Es könnte argumentiert werden, dass die Beurteilung der Messung Bestandteil der Aufgaben des IQWiG ist, wobei das Treffen von Wertentscheidungen in den Aufgabenbereich der Selbstverwaltung fällt bzw. beim G-BA liegt. Konkret: Das IQWiG bewertet die Messung, also die Begründung kausaler Effekte, und der G-BA trifft auf Basis des Berichtes bei der Feststellung des Ausmaßes des Zusatznutzens die notwendigen Wertentscheidungen. Diese Argumentation greift zu kurz, da eine rationale Entscheidung durch den G-BA die Informationen über die Wertmaßstäbe der Bevölkerung und Präferenzen der Patienten benötigt. Diese In-

Literatur:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Hrsg.) (2014): Allgemeine Methoden; Entwurf für Version 4.2 vom 18.06.2014; Köln.

Quelle:

Dieser Kommentar basiert in wesentlichen Teilen auf einer Stellungnahme des Autors im Stellungnahmeverfahren des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) zum Methodenpapier 4.2.

formationen sollten demnach Gegenstand des Dossiers und Gegenstand des Berichtes des IQWiG an den G-BA sein.

Die Berücksichtigung der Wertvorstellungen der Versicherten und die Präferenzen der Patienten ist eine notwendige Bedingung für eine systematische und transparente Entscheidungsfindung. Es ist nicht ausreichend geklärt, auf welcher Basis bzw. welcher Evidenz die Abwägung des Nutzens und Schadens erfolgen soll. Dies ist sowohl aus der Perspektive der evidenzbasierten Medizin als auch aus Sicht der Gesundheitsökonomie eine wesentliche Lücke des Status Quo. Gegenwärtig werden diese Wertentscheidungen nicht transparent gemacht. Wertentscheidungen können explizit auf Basis subjektiver oder objektiver Daten getroffen werden. Zunehmend wird die Forderung laut, implizites Wissen der Entscheidungsgremien durch explizites Wissen über diese Wertentscheidungen zu ersetzen. Der Anspruch an die Wissenschaftlichkeit des Bewertungsverfahrens macht eine eindeutige und klare Kommunikation der Gewichtung der Zielkriterien notwendig. Explizites Wissen kann systematisch dokumentiert und über eine wissenschaftliche Veröffentlichung der Bürger-, Versicherten oder Patientenpräferenzen kommuniziert und korrigiert werden.

Oftmals wird in diesem Kontext darauf hingewiesen, dass das IQWiG mit der Feststellung zum Ausmaß klinischer Effektmasse betraut ist. Bei der Beurteilung der Evidenz sollten möglichst wenige Wertentscheidungen (etwa zur unterschiedlichen „Wertigkeit" von Indikationen) getroffen werden, da diese Aufgabe dem G-BA aufgrund seiner Legitimation vorbehalten sein sollte. Die Trennung zwischen der Informationsaufbereitung und der tatsächlichen Entscheidung, also der Trennung der Aufgaben des IQWiG und des G-BA macht Sinn. Jedoch muss der G-BA in seiner Funktion auch in der Lage sein, auf den durch das IQWiG berichteten Daten zu entscheiden. Keinen Sinn macht die Tatsache, dass der G-BA durch den Bericht des IQWiG, keinerlei belastbare Informationen über die Werturteile der Versicherten bzw. Präferenzen der Patienten bekommt. In diesem Kontext müsste klar herausgestellt werden, inwiefern diese Informationen über die Präferenzen und Werturteile durch den Antragsteller im Dossier bereitgestellt werden bzw. wie diese Informationen für die Entscheidung des G-BA zusammenfassend beurteilt werden. <<

„Das reale Labor"

>> Im Rahmen eines Symposiums haben das DFG-geförderte Netzwerk „Klinische Studien in der Allgemeinmedizin" und die Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM) mit zahlreichen Experten aus Gesundheitspolitik und Gesundheitswesen über die Notwendigkeit einer strukturellen Förderung allgemeinmedizinischer Forschung diskutiert. Prof. Ferdinand M. Gerlach, Präsident der DEGAM, betonte in seinem Grußwort, dass hochspezialisierte Forschung an Universitätskliniken allein nicht ausreiche: „Auch in Deutschland ist das reale Labor die Praxis." Die Signale des Symposiums waren durchweg positiv: So soll zum Beispiel im Innovationsfonds der Bundesregierung ein Schwerpunkt auf die Primärversorgung gelegt werden.

Ein zentrales Ziel liegt nach Aussagen auf dem Symposium darin, forschungsfördernde Institutionen von der Notwendigkeit Klinischer Forschung und Versorgungsforschung zu überzeugen: „Die Förderung der Forschung in der Allgemeinmedizin muss vor allem nachhaltig sein", richtete Dr. Ildikó Gágyor (Univ. Göttingen) eine Botschaft aus den Vorträgen an die Entscheidungsträger.

Im Rahmen einer Podiumsdiskussion, moderiert von DEGAM-Vizepräsident Prof. Martin Scherer, herrschte bezüglich der Notwendigkeit eines Strukturwandels bei Klinischen Studien in der Allgemeinmedizin weitgehend Konsens. Auch Mitglieder des G-BA stimmten darin überein, dass Versorgungsprobleme in der ambulanten Versorgung stärker in den Fokus der Förderprogramme und des Innovationsfonds rücken sollten. <<

Gesundheit und Einwanderung

>> Wenn heute rund 20 Prozent der Bevölkerung eine familiäre Einwanderungsgeschichte haben, verändert diese Tatsache gesellschaftliche Systeme und Rahmenbedingungen. Dies gilt auch für das Gesundheits- und Pflegesystem, denn der Anteil der Seniorinnen und Senioren mit Migrationsgeschichte an der Bevölkerung steigt. Erstens werden sie voraussichtlich mehr medizinische und pflegerische Leistungen in Anspruch nehmen. Zweitens stellt sich die Aufgabe der interkulturellen Öffnung von Gesundheits- und Pflegeinstitutionen, die ihre Angebote auf die zunehmende kulturelle und ethnische Vielfalt der Bevölkerung ausrichten müssen. Dies hat Folgen sowohl für die Aus- und Weiterbildung von Fachpersonal als auch für Abläufe und Denkprozesse in den Institutionen.

Diesen und ähnlichen Fragen widmet die Friedrich-Ebert-Stiftung (Hiroshimastraße 17, Berlin-Tiergarten) die Fachtagung „Gesundheitsversorgung und Pflege in der Einwanderungsgesellschaft", die am 21. April von 11.30 bis 17.30 Uhr stattfinden wird. Anmeldung: http://fes.de/sets/s_ver.htm <<

Placebo-Studie

>> Patienten erwarten von einem teuren Medikament eine bessere Wirkung als von einem billigen. Schon diese Erwartung steigert den Behandlungserfolg. Diesen Zusammenhang belegt eine Pilotstudie von Espay (2015) mit Parkinson-Patienten, bei der keine echten Medikamente verabreicht wurden. <<

„Mehr Versorgungsforschungung"

>> Anlässlich des Weltkrebstags 2015 plädiert die Deutsche Krebsgesellschaft für eine Stärkung der Krebsforschung in Deutschland, besonders im Hinblick auf die Versorgung der Patienten im medizinischen Versorgungsalltag.

„Medizinische Innovationen müssen ihre Tauglichkeit nicht nur an ausgewählten Patienten in kontrollierten klinischen Studien unter Beweis stellen, sondern auch im medizinischen Routinebetrieb", betont Prof. Andreas Hochhaus, Vertreter der Landeskrebsgesellschaften im Vorstand der Deutschen Krebsgesellschaft. „Entscheidend sind außerdem stationäre und ambulante Versorgungsstrukturen, die möglichst allen Krebspatienten den raschen Zugang zu einer qualitativ hochwertigen Behandlung ermöglichen", ergänzt Prof. Christian Jackisch, ebenfalls Vertreter der Landeskrebsgesellschaften im Vorstand der Deutschen Krebsgesellschaft.

„Um die Versorgung Krebskranker weiter zu verbessern, brauchen wir, neben der Grundlagenforschung und der klassischen klinischen Forschung, mehr Versorgungsforschung. Sie identifiziert Stärken und Schwächen der real praktizierten Medizin – und liefert auf diese Weise oft wertvolle Anhaltspunkte für eine bessere Versorgungsqualität", sagt Dr. Johannes Bruns, Generalsekretär der Deutschen Krebsgesellschaft. Bislang werde die Bedeutung der Versorgungsforschung meist unterschätzt. Solche Studien sollten ausreichend finanziert werden, argumentiert Bruns und begrüßt den angekündigten Innovationsfonds der Bundesregierung. <<