Titelinterview mit Hans-Holger Bleß (IGES) und Dr. Thomas Müller (G-BA)

„Zwischen Werteentscheidung und Saldierung“

Dr. Thomas Müller, Leiter der Abteilung Arzneimittel des Gemeinsamen Bundesausschusses, hat Pharmazie und Medizin in Berlin studiert und war zuletzt als Direktor der Zentralapotheke und Leiter der Medizinischen Beschaffung des Universitätsklinikums Rostock AöR auch zuständig für den Einkauf von Arzneimitteln. Müller, der beim Bundesausschuss verantwortlich ist für die interne Vorbereitung aller Arzneimittelbewertungen, kennt Hans-Holger Bleß, beim IGES-Institut zwar eigentlich Bereichsleiter Versorgungsforschung, doch uneigentlich mit sehr vielen Arzneimitteldossiers direkt oder indirekt befasst ist, sehr gut: In ihrer Arbeit treffen die Beiden regelmäßig aufeinander, doch mit recht unterschiedlichen Ansichten rund um das AMNOG-Geschehen, wie im Doppelinterview mit „Monitor Versorgungsforschung“ recht deutlich wird.

>> Vor rund zweieinhalb Jahren hat das IQWiG ein zweiarmiges Pilotprojekt gestartet: Zwei Teams sollten die Einsatzmöglichkeiten zweier ökonomischer Modelle – einerseits den Analytic Hierarchy Process (AHP), andererseits die Conjoint Analysis (CA) – in zwei verschiedenen Indikationen analysieren. Vor rund einem Jahr wurde durch das IQWiG das Arbeitspapier zur AHP von Prof. Jizerman, Universität Twente, vor wenigen Wochen nun das von Prof. Mühlbacher, Hochschule Neubrandenburg, zur CA veröffentlicht. Bei der Veröffentlichung des CA-Papiers schrieb das IQWiG: „Die Methode der Conjoint Analysis ist prinzipiell geeignet, um herauszufinden, welche Präferenzen Patientinnen und Patienten bei Therapiezielen haben.“ Herr Müller, unter welchen Prämissen können Sie sich vorstellen, dass der G-BA diese beiden oder eines der beiden Modelle anerkennt, um Patientenpräferenzen in den AMNOG-Prozess stärker als bisher berücksichtigen zu können?

Müller: Hierzu muss man wissen, dass das ein methodischer Vorschlag ist, der aus dem Bereich der Kosten-Nutzen-Bewertung kommt, in dem der QALY sowie die Effizienzgrenze die beiden wesentlichen Ansätze sind. In beiden Ansätzen benötigt man ein aggregiertes Nutzenmaß und dazu wiederum eine Methodik, welche die verschiedenen Endpunkte und Effekte von Wirkstoffen und Therapien zusammenführt. Die Conjoint Analysis und der Analytic Hierarchy Process sind Befragungsansätze, aus deren Ergebnissen mit Hilfe von statistischen Methoden ein Gesamtmaß gebildet wird.

Bleß: Ergänzend sollte man aber erwähnen, dass der G-BA im Zuge des AMNOG-Verfahrens die Kosten-Nutzen-Bewertung noch nicht ein einziges Mal praktiziert hat.

Müller: Der Gesetzgeber hat die Kosten-Nutzen-Bewertung ans Ende des AMNOG-Verfahrens gesetzt. Sie kann durchgeführt werden, ist es aber noch nicht, weil weder ein pharmazeutischer Unternehmer noch der GKV-SV eine solche beantragt haben.

„Diese individuelle Werteentscheidung trifft ein Gremium von Leuten, die von der Erkrankung nicht direkt betroffen sind und die Patientensicht nicht repräsentieren.“
Hans-Holger Bleß

Was halten Sie denn nun von der Methodik an sich?

Müller: Ich stehe diesen statistischen Methoden eher skeptisch gegenüber, da jedwede Patientenpräferenz immer eine individuelle Wertentscheidung ist. Wir sollten aufpassen, dass individuelles Krankheitserleben nicht mit statistischen Methoden in eine Art – aus meiner Sicht nicht sachgerechte – Mehrheitsentscheidung überführt wird.

Bleß: Dennoch sind gerade solche Wertentscheidungen verfahrensimmanent, da Ergebnisse mit unterschiedlichen Effektrichtungen und in unterschiedlichen Kategorien wie Überleben, Morbidität, Lebensqualität und unerwünschte Wirkungen zu einem Gesamtergebnis saldiert werden müssen.

Müller: Im G-BA werden keine abstufenden Wertungen vorgenommen oder gar Hierarchien von Endpunkten gebildet.

Würden Sie ein Beispiel nennen?

Müller: Bei der Beschlussvorbereitung stellen wir Endpunkte nebeneinander, wie es der Gesetzgeber in der Arzneimittelnutzenverordnung gefordert hat. Es kann zum Beispiel durchaus sein, dass bei der einen Erkrankung für den Patienten ein Monat längeres Überleben und bei der anderen dagegen die Vermeidung oder Linderung von Juckreiz entscheidend ist. Das lässt sich nicht gegeneinander aufrechnen.

Das heißt doch im Umkehrschluss, dass es dem G-BA nicht um die Frage geht, ob in einer Erkrankung das Überleben wichtiger ist als die Lebensqualität.

Müller: Nur für den Fall, dass wir erkennen, dass ein Wirkstoff bezogen auf einen Endpunkt positive Effekte und bei einem anderen negative, also gegenläufige Effekte hat, prüft der G-BA, ob das auch im Gesamtergebnis abgebildet werden muss. Der G-BA berücksichtigt dies dann, indem saldiert wird. Die positiven und negativen Effekte insgesamt werden betrachtet und im Gesamtergebnis abschließend bewertet.

Dieses Saldieren hört sich wie eine exakte Wissenschaft an. Doch in Wirklichkeit wird man wohl Äpfel und Birnen „saldieren“.

Müller: Genau aus diesem Grunde kann es keinen Algorithmus geben, weder im Sinne einer Rechenmethodik noch einer empirischen Befragung, die das IQWIG mit diesen beiden Modellversuchen erprobt hat. Für den G-BA ist und bleibt dieses Procedere auch eine individuelle Werteentscheidung, die auf der Grundlage der Evidenz und der Stellungnahmen getroffen werden muss.

Bleß: Wobei diese individuelle Werteentscheidung ein Gremium von Leuten trifft, die von der Erkrankung nicht direkt betroffen sind und die Patientensicht nicht repräsentieren. Im Gegenzug sind zumindest die entsendenden Einrichtungen von den Folgen ihrer Entscheidungen betroffen.

Das verstehe einmal einer: Auf der einen Seite hält das IQWiG und auch der G-BA die höchste Evidenzlatte in Form von RCT-Studien hoch und auf der anderen Seite werden sozusagen eminenzbasierte Indivi-

dualentscheidungen getroffen. Das ist schon eine arg große Spreizung.

Müller: Das Plenum des G-BA hat die Aufgabe, neue Wirkstoffe auf einer abstrakten Populationsebene zu bewerten, um Preisverhandlungen auf einer Evidenzbasis vorzubereiten. Das muss man abgrenzen von einer individuellen Entscheidung auf Arzt-Patienten-Ebene – diese wird vom G-BA in keiner Weise eingeschränkt. Jeder Patient überlegt sich höchst individuell gemeinsam mit seinem Arzt, welche Therapie für ihn in seiner ganz speziellen Situation infrage kommt.

Bleß: Damit trifft aber der Patient gemeinsam mit dem Arzt lediglich für sich eine Werteentscheidung.


Dr. Thomas Müller, G-BA

Müller: Genau. Deshalb will ja der G-BA dem Arzt und dem Patienten möglichst gute und entscheidungsrelevante Informationen über alle zu erwartenden Effekte zur Verfügung stellen. Bei einer onkologischen Therapie kann es eine Chance geben, dass ein Patient mit einem speziellen Medikament beispielsweise sieben Monate länger zu leben hat, womit aber auch bestimmte Nebenwirkungen einhergehen. Diese Abwägungs-Entscheidung muss in jeder solchen Therapiesituation vom Arzt und Patienten immer wieder aufs Neue getroffen werden.

Bleß: Das wäre das Ideal. Doch so kommt man nie zu einer Gruppenentscheidung.

Müller: Der Bundesausschuss steht vor der Herausforderung, Produkte zu bewerten, um auf Basis dieser Beschlüsse Preise zu definieren. Fertigarzneimittel in der GKV haben eben nur einen Preis für alle Patienten. Hier sehen wir durchaus die Schwierigkeit, die verschiedenen Effekte zu aggregieren, um zu einem einheitlichen Bewertungsergebnis und darauf aufbauend zu einem einheitlichen Erstattungsbetrag zu kommen. Dieses Bewertungsergebnis ersetzt aber in keinem Fall die individuelle Therapieentscheidung.

„Wir haben in Deutschland noch keinen Algorithmus nach dem NICE-Vorbild mit einem ICER und einer QALY-Berechnung."
Dr. Thomas Müller

Bleß: Aber genau das ist doch der Ansatz des IQWiG mit der Conjoint-Analyse oder dem AHP. Hierin stecken gerade die bislang ungenutzten Möglichkeiten der Präferenzmessung.

Müller: Und der Ansatz des NICE ist eben der QALY. Wir haben in Deutschland noch keinen Algorithmus nach dem NICE-Vorbild mit einem ICER und einer QALY-Berechnung. Wie wir hier verfahren, ist eine Art Gegenrechnen: Wenn die Evidenz sowohl zum Schaden als auch zum Nutzen signifikant ist, wird saldiert.

Jetzt rechnen Sie auf einmal doch gegen, wo Sie gerade gesagt haben, es gibt keinen Algorithmus.

Müller: Das liegt an den Kategorien der Nutzenverordnung. Der Gesetzgeber hat drei positive Kategorien vorgegeben, wobei „Heilung" in der höchsten Kategorie „erheblicher Zusatznutzen" genannt wird. Es gibt hier keine Skala von 1 bis 100, sondern ein semiquantitatives Vorgehen. Qualitative Merkmale wie „nachhaltige", „besonders ausgeprägte", „moderate" oder „geringe" Effekte werden saldiert. Insgesamt gibt es im G-BA eine sehr gefestigte Position, dass empirische Befragungen die so getroffenen Werteentscheidung nicht ersetzen können. Ich bin auch sehr skeptisch, ob man die mögliche Breite der individuellen Werteentscheidungen überhaupt so statistisch verarbeiten kann wie naturwissenschaftliche Evidenz.

Bleß: Der G-BA trifft normative Entscheidungen, die auf eine Kohorte replizieren, die im Prinzip Gesamtdeutschland umfasst. Darum müsste sich der Bundesausschuss doch im Endeffekt auch entsprechender Instrumente bedienen, die diese Mehrheit statistisch erfassen könnten. Das Treffen reiner Individualentscheidungen ist mir hier ein bisschen zu wenig.

Müller: Hier muss ich deutlich widersprechen. Diese individuellen Werteentscheidungen können nicht einer Mehrheit folgen. Eine Werteentscheidung kann der G-BA dem einzelnen Patienten eben nicht abnehmen.

Aber der G-BA macht das doch paternalistisch für alle Patienten. Wäre es da nicht besser, einen Mehrheitsentschluss über eine Conjoint-Analyse oder eine andere statistische Methode zumindest versuchen zu wollen?

Müller: Diese Forschungsrichtung ist sicher wichtig, um informierte Entscheidungen treffen zu können. Empirische Forschung, wie Patienten Krankheiten werten, halte ich für essenziell. Dieses Vorgehen ist gerade auch als valide Basis für die Patientenvertretung im Bundesausschuss wertvoll. Dennoch: Wenn man mit statistischen Methoden herausbekommt, dass beispielsweise 70 Prozent der Befragten der Meinung sind, dieser oder jener Endpunkt sei besonders wichtig, kann das nicht gleichzeitig bedeuten, dass deswegen die anderen Endpunkte, die der Gesetzgeber fordert, zu vernachlässigen sind. Denn für 30 Prozent der Patienten können genau die anderen Endpunkte wertvoll sein. Das ist dann keine Entscheidung, die der G-BA paternalistisch trifft.

Herr Bleß, Sie haben beim Stichwort „paternalistisch" gezuckt.

Bleß: Ja, weil die im Verfahren so hochgehaltene Patientenrelevanz durchgehend paternalistisch entschieden wird, anstatt den Patienten einzubeziehen und zu befragen. Wir erstellen im IGES Institut seit Jahren eine ganze Reihe von solchen Präferenzmessungen, wozu wir meistens die Discrete-Choice-Analyse nutzen, die wiederum eine Variante der Conjoint-Analyse ist. Dabei machen wir im Prinzip genau das, was auch der G-BA macht: Wir messen Saldierungs- und Abwägungsentscheidungen, aber eben aus der Patientenperspektive. Dabei wird

erfasst, welche Nebenwirkungen in Kauf genommen werden, um welchen positiven Effekt zu erreichen und natürlich werden dazu negative und positive Effekte in den Saldo genommen. Diese Präferenzmessung beantwortet genau die Frage, die Herr Müller ausgeführt und die sich der G-BA ständig zu stellen hat.


Hans-Holger Bleß, IGES-Institut

Welche denn?

Bleß: Es gibt immer unterschiedliche Nutzenaspekte bei der Bewertung eines neuen Arzneimittels, beispielsweise einen geringfügigen Effekt in der Lebensqualität, in der Mortalität oder in der Morbidität. Daher lautet die Frage: Was wird aus dreimal geringfügig? Wird daraus ein beträchtlich oder ein erheblich?

Die ganz ehrliche Antwort?

Bleß: Wir wissen es nicht, wenn wir die Patienten nicht befragen. Auch bei anderen Fragen müssen wir passen, wenn wir auf die Sicht der Patienten verzichten, beispielsweise wie Wirkung und Nebenwirkung zu verrechnen sind. Oder wie man die krankheitsbezogene Lebensqualität – die ja ihrerseits bereits Dimensionen beinhaltet, die die Morbidität beschreiben – ohne Redundanzen in die Saldierung einbezieht. Aber es ist im Verfahren nun einmal ein aggregiertes Gesamtmaß vorgesehen und sei es für eine Teilpopulation. Daher muss immer eine solche Abwägungsentscheidung getroffen werden, aber sie wird zur Zeit halt paternalistisch ausgeübt.

Es wird übrigens auch paternalistisch darüber entschieden, was denn patientenrelevant ist und was nicht.

Bleß: Exakt. Das gilt insbesondere für die sogenannten Surrogate wie etwa die Zeit des progressionsfreien Überlebens bei Krebserkrankungen oder das dauerhafte virologische Ansprechen (SVR) bei Viruserkrankungen.

Da erklärt das IQWiG immer wieder, dass dies methodisch nicht validiert und deswegen nicht patientenrelevant sei.

Bleß: Mich stört diese auf die Methodik reduzierte Blickweise. Der Patient wird auf dem Papier ins Zentrum des Verfahrens gestellt, doch es sind Dritte, die darüber entscheiden, was dem Patienten wichtig zu sein hat und was nicht, und wie er sich zu entscheiden und was wie zu bewerten hat.

Ihre Meinung zu AHP und CA?

Bleß: Ich denke, dass Analysen, die mit diesen Instrumenten gemacht werden können, sehr wohl auch in der frühen Nutzenbewertung genutzt werden können und schon jetzt genutzt werden sollten, um genau jene Saldierungsentscheidungen wesentlich valider treffen zu können. Denn bisher gibt es, wie Herr Müller ausgeführt hat, eben keine erkennbare Methodik für die Saldierung, die obendrein und abweichend vom ansonsten sehr transparenten Verhalten des G-BA absolut intransparent erfolgt.

Wären denn die Lösungsansätze, die das IQWiG in Spiel gebracht hat, ein probater Ansatz?

Bleß: Das IQWiG hat zumindest einen Rechenalgorithmus für den Nutzen, doch bleibt auch hier die Saldierung eine im Verborgenen stattfindende Abwägungsentscheidung. Doch der G-BA hat sich diesem Rechenalgorithmus sowieso nicht genähert, sondern sich den Spielraum eröffnet, den der Wortlaut der Rechtsverordnung lässt. Wobei ich die Begrifflichkeiten nicht als semiquantitativ bezeichnen würde, es handelt sich um semantische Tautologien. Was heißt denn genau wenn „gering" als „moderat aber nicht geringfügig" definiert wird, was hilft es, wenn „beträchtlich" mit „deutlich" erläutert wird oder „erheblich" mit „großer Verbesserung"? Diese Ungenauigkeit gibt eigentlich jedweden Spielraum, ob ein Effektmaß nun moderat, gering, geringfügig oder was auch immer ist. Da ist der Spielraum doch recht groß, was aber per se nicht schlecht sein muss.

Warum?

Bleß: Weil der G-BA dadurch noch andere Aspekte mit einbeziehen kann, als das ein purer Rechenalgorithmus leisten kann. Doch da, wo saldiert wird im Namen des Patienten, würde ich mir schon wünschen, dass der Patient auch mit einbezogen und dessen Stimme gehört wird, indem man ihn befragt – und genau dafür ist die Methodik ausgereift genug.

Müller: Der Gesetzgeber hat für den G-BA eine sehr intensive Patientenbeteiligung vorgesehen. Es gibt sowohl die themenbezogenen Patientenvertreter als auch die ständigen, alle arbeiten sehr intensiv im AMNOG-Verfahren mit. Doch muss ich bezüglich der hier diskutierten Modelle warnen und bremsen: Der Gesetzgeber fordert im AMNOG eine evidenzbasierte Entscheidung. Dabei bezieht er sich auf naturwissenschaftliche Evidenz und nicht auf die Patienten-Präferenz. Das, was sie machen, ist doch zum Teil reine Meinungsforschung. Und damit etwas ganz anderes als evidenzbasierte Medizin mit aus validen Studien ablesbaren Effekten.

Das nennt man einen Konflikt.

Müller: Genau. Der G-BA muss eine evidenzbasierte Entscheidung treffen, bei der die Patientenpräferenz eine Rolle spielen kann, aber eben nicht die ausschlaggebende. Eben weil sie eher ein Meinungsbild und kein naturwissenschaftliches Ergebnis ist. Ich bin als Mediziner skeptisch, ob man hier Mehrheitsentscheidungen tatsächlich treffen kann. Ich stimme Ihnen zu, dass man mit der Patientenpräferenz bestimmte Effekte besser erfassen kann, wie beispielsweise in der MS-Therapie die Spastik, die für den Patienten im Alltag sehr einschränkend sein kann. Wenn mit Hilfe solcher Analysen derartige Effekte reproduzierbar und damit valider erfasst werden können, ist das auch für den Bundesausschuss relevant. Derartige Erkenntnisse darf man, um das Wort zu gebrauchen, nicht paternalistisch abtun. Doch noch einmal: Die Wertung, die Meinungsbildung an sich erfüllt nicht die harten Anforderungen der evidenzbasierten Medizin. Da muss der G-BA einer klaren Linie folgen.

BERLINER PFLEGEKONFERENZ

AUS DER PRAXIS – FÜR DIE PRAXIS

am 11. und 12. November 2014 im dbb forum berlin

Erfahren Sie mehr und diskutieren Sie mit namhaften Vertretern aus der Gesundheits- und Familienpolitik sowie mit Experten und unterschiedlichen Teilnehmergruppen aus dem Bereich Pflege.



Manuela Schwesig
Bundesministerin für Familie, Senioren, Frauen u. Jugend (angefragt)



Dr. Norbert Blüm
Bundesminister a. D.



Karl-Josef Laumann
Beauftragter der Bundesregierung für die Belange der Patientinnen u. Patienten sowie Bevollmächtigter für Pflege



Cornelia Prüfer-Storcks
Senatorin für Gesundheit und Verbraucherschutz der Hansestadt Hamburg (angefragt)



Mario Czaja
Senator für Gesundheit und Soziales von Berlin



Prof. Dr. Ursula Lehr
Vorsitzende, Bundesarbeitsgemeinschaft der Senioren-Organisationen BAGSO e. V.



Dr. Gerd Landsberg
Geschäftsführendes Präsidialmitglied Deutscher Städte- und Gemeindebund



Jens Spahn
Gesundheitspolischer Sprecher der CDU/CSU-Fraktion im dt. Bundestag (angefragt)



Thomas Isenberg
Gesundheitspolischer Sprecher der SPD-Fraktion im Berliner Abgeordnetenhaus



Andreas Westerfellhaus
Präsident deutscher Pflegerat DPR



Franz Knieps
Vorstand BKK Dachverband



Jürgen Hohnl
Geschäftsführer IKK e.V.

VERLEIHUNG MARIE-SIMON-PFLEGEPREIS

MARIE SIMON PFLEGEPREIS *in Kooperation mit* DStGB Deutscher Städte- und Gemeindebund

Auszeichnung innovativer Pflegeprojekte

Veranstalter
spectrumK GmbH
Spittelmarkt 12
10117 Berlin

Veranstaltungsort
dbb forum berlin
Friedrichstr. 169/170
10117 Berlin

Organisationsbüro
Tel. +49 (0)30 21 23 36-153 / -154
Fax +49 (0)30 21 23 36 39 157
info@berliner-pflegekonferenz.de

Weitere Informationen unter
www.berliner-pflegekonferenz.de
www.marie-simon-pflegepreis.de

Dann muss das viel beschworene „lernende System“ AMNOG eben lernen, mit derartigen ökonomischen Modellen, die schon mit hoher wissenschaftlicher Professionalität daherkommen, entsprechend umzugehen. Wenn ich die mit Pilotstudien zu AHP und CA vom IQWiG beauftragten Professoren mit dem von Ihnen, Herr Müller, verwandten Begriff „Meinungsforschung“ konfrontieren würde, springen die im Karree.

Müller: Wenn denn validierte Verfahren wie Conjoint-Analysen, Discret-Choice-Experimente oder der Analytic Hierarchic Process Hinweise geben können, die denen der evidenzbasierter Medizin entsprechen, sollte man sie verwenden. Der G-BA nimmt in den beschriebenen Fällen, in denen sich positive und negative Effekte gegenüberstehen, auf der Grundlage von Werteabschätzungen eine Saldierung vor. Das kann auch zu Mehrheitsentscheidungen führen, auch wenn diese bisher ausgesprochen selten sind – der weitaus größte Teil der Beschlüsse sind bisher einstimmig und mit Zustimmung der Patientenvertretung gefasst worden. Diese Entscheidung gilt dann für die jeweilige Patientengruppe und ist Basis der Preisverhandlung. Der einzelne Arzt und sein Patient sind aber nicht daran gebunden, sondern müssen ihre eigene Entscheidung treffen.

Bleß: Die zentrale methodische Herausforderung ist immer die Auswahl einer repräsentativen Befragungsgruppe. Das können je nach Aufgabenstellung von einer bestimmten Krankheit betroffene Patienten sein, oder eine Versichertengemeinschaft, weil das diejenigen sind, die die Kosten zu tragen haben. Auch muss die Auswahl der Fragen so erfolgen, dass keine Übergewichtung einzelner Aspekte einen Bias erzeugt, indem sie redundant sind und praktisch mehrfach einfließen. Das erfordert ein feines Gefühl und viel Erfahrung. Die Methodik an sich ist nicht das Problem, denn die ist inzwischen sehr ausgereift, weil sie seit langem auch jenseits des Gesundheitswesens verwendet wird, um letztlich eine Zahlungsbereitschaft von Menschen bezüglich einzelner Attribute von Produkten zu ermitteln.

„Die Herausforderung an den G-BA ist nicht nur die Patientenperspektive, sondern auch die Versichertenperspektive einzunehmen.“
Dr. Thomas Müller

Typischerweise bei Autos.

Bleß: Oder für Ökosiegel auf der Waschmaschine, selbst dafür gibt es Discrete-Choice-Experimente. Was aber nichts anderes heißt, als dass diese Instrument seit langem validiert sind und inzwischen Goldstandard sind. Nur eben noch nicht im Gesundheitsbereich.

Doch wie stellen Sie sicher, dass nicht eine sozusagen theoretische Meinung abgefragt wird.

Bleß: Bei einer Erkrankung werden zunächst die Betroffenen ausgewählt, weil nur sie in der Lage sind, eine Werteentscheidung zu treffen. Das ist viel schwieriger für einen Gesunden, der dies abstrakt tun muss. Ebenso werden die Fragen intern validiert und genau darauf geachtet, dass es keine Redundanzen gibt, damit hinterher die Gewichtung stimmt. Ebenso gibt es Pilotvorstufen, bei denen kleine Panels befragt werden, um zu sehen, ob die ausgewählten Fragen für die angewandte Methode auch geeignet sind, um verwertbare Aussagen zu bekommen. Danach erst wird die Studie durchgeführt und die Ergebnisse statistisch ausgewertet. Das ist so gestaltet, dass den einzelnen Attributen eine Bewertung zukommt, und so auch das Gesamtmaß von verschiedenen Aspekten bewertet werden kann.

Müller: Die Herausforderung an den G-BA ist, nicht nur die Patientenperspektive sondern auch die Versichertenperspektive einzunehmen. Diesen Konflikt erleben wir in fast jeder Arzneimittelbewertung, indem die themenbezogenen Patientenvertreter bestimmte Effekte aus ihrer ganz eigenen Sicht betrachten und diese Meinung – auch sehr nachvollziehbar – deutlich einbringen, beispielsweise zu leichten Hyperglykämien bei Diabetes. Dann ist es tatsächlich keine leichte Aufgabe, in einem auf Pluralität basierenden Gremium wie dem G-BA eine Gesamtperspektive zu finden. Insofern sind die Kriterien, die der Gesetzgeber in der Arzneimittelnutzenverordnung nennt, sehr hilfreich. Die Einteilung in leichte Nebenwirkungen, mittelmäßige und schwere mag etwas grob sein, aber man kann nach differenzierter Betrachtung der ersten 80 AMNOG-Ergebnisse deutlich erkennen, dass es bei den Produkten große Unterschiede gibt. Und die werden eben auch abgebildet.

Sie sagten, es gebe keine einzige richtige Kosten-Nutzen-Bewertung und schon gar keine Gesamtnutzen-Bewertung über verschiedene Arzneimitteltherapien hinweg. Wie will man denn da je hinkommen?

Müller: Das ist eine gute Frage. Meine Meinung ist, dass wir dazu noch eine umfangreiche Diskussion brauchen.

Die wollen wir ja gerade mit anstoßen.

Müller: Die ist auch nötig, wenn wir uns die aktuelle Entwicklung bei Arzneimittelpreisen vor Augen halten, bei denen einige Orphans Jahrestherapiekosten-„Schallgrenzen“ von bis zu einer Million Euro schon erreicht haben. Ich bin davon überzeugt, dass wir auch in Deutschland mittelfristig einen methodischen Ansatz brauchen, um Kosten und Nutzen und Preise in einen fairen Ausgleich zu bringen. Aus meiner Sicht müssten hierzu allerdings mehr Ökonomen – und nicht nur Gesundheitsökonomen – und deren Instrumente hinzugezogen werden.

Warum nicht nur Gesundheitsökonomen?

Müller: Weil es bei der Frage nach einem fairen Preis nicht nur um QALYs und Effizienzgrenzen geht, sondern auch um volkswirtschaftliche Fragen. Denn hier müssten künftig auch Diskussionen geführt werden zur Schaffung einer Balance zwischen Preis und Aufwendungen für Studien, aber auch Anreize gegeben werden für Forschung und Entwicklung. Und gleichzeitig dürfte dem GKV-System nicht zu viel Geld entzogen werden, damit es nicht instabil wird.

Wäre da nicht das QALY-Konzept ein probater Ansatz?

Müller: Ich halte weder das QALY-Konzept noch die Effizienzgrenze für ausreichend. Denn bei beiden fehlt die volkswirtschaftliche Gesamtperspektive, außerdem nehmen beide Methoden dem Gesetzgeber nicht die Werteentscheidung ab für die Frage, was ein fairer und akzeptabler Preis ist. Dazu müsste der Gesetzgeber erst einmal die nächsten, viel weiterführenden Schritte machen, was aber nicht abzusehen ist. Im Moment kann ich nicht einmal erkennen, dass wir in die Kosten-Nutzen-Bewertung einsteigen. Wir hatten immerhin gerade die Meldung, dass die 50. Erstattungsbetragsverhandlung erfolgreich abgeschlossen wurde. In dieser Hinsicht ist der Ansatz des Gesetzgebers zur Zeit jedenfalls recht erfolgreich.

Bleß: Ich würde auch vorhersagen, dass der Gesetzgeber erst mal weiter abwartet, was die Selbstverwaltung aus den bisherigen Rahmenvorgaben macht. Das war schon immer so, und es gibt derzeit auch keinen dringlichen Grund, das mit dem AMNOG installierte Prinzip



Caring and Curing

Leben retten und Gesundheit verbessern – das ist unser Ziel

Die Entwicklung bahnbrechender neuer Medikamente steht für Novartis an erster Stelle. Sie schaffen neue Behandlungsmöglichkeiten für bislang unerfüllte medizinische Bedürfnisse der Patienten.

Patienten und ihre Bedürfnisse können jedoch sehr unterschiedlich sein. Deshalb bietet Novartis neben innovativen Medikamenten auch Möglichkeiten zur Krankheitsvorbeugung sowie Generika an und verbessert den Zugang zu medizinischer Versorgung.

NOVARTIS
caring and curing

Novartis Pharma GmbH · Roonstraße 25 · 90429 Nürnberg · www.novartis.de

wieder zu verlassen. Ob man wirklich ein indikationsübergreifendes Gesamtmaß des Nutzens braucht, möchte ich auch hinterfragen. Aus der reinen Nutzenperspektive heraus muss ich die Heilung einer Mittelohrentzündung nun mal nicht mit der Lebensqualität von onkologischen Patienten vergleichen. Denn diese Art von Information ist aus der Nutzendimension heraus sinnbefreit, nicht aber aus jener der Preisfindung. Auch möchte ich hinterfragen, ob wir uns nach dem QALY-Konzept sehnen sollten, weil wir ganz schnell bei der Frage anlangen, was ein Lebensjahr kosten darf.

Das ist eine Frage, die in Deutschland sicherlich ohnehin sehr schwer zu führen sein wird.

Bleß: Aber wenn, sollte sie in der Gesellschaft geführt werden und ganz sicher nicht von HTA-Behörden. Wir sehen schon im Methodenpapier der Kosten-Nutzen-Bewertung des IQWiG, dass da ein kleines bemerkenswertes Kapitelchen drin ist, das sich „Zumutbarkeit der Kostenübernahme" nennt. Das ist noch sehr zurückhaltend befüllt, aber letztlich wird damit bei jeder Errechnung eines Preises die letzte Konsequenz adressiert: Was ist, wenn das Rechenergebnis dazu führt, dass die GKV die Ausgaben nicht stemmen kann? Wenn diese letzte Konsequenz einmal eintritt, steigen wir in eine neue Ära der Allokationsentscheidungen ein. Dann geht eine ethische Wertedebatte los, die ich ungern beim IQWiG oder beim G-BA verortet sehen, sondern immer in der Gesellschaft verankern würde. Von daher ist die bislang ausgeübte Verhandlungslösung wahrscheinlich nicht das schlechteste, weil vom G-BA die Evidenz im Verhältnis zum Nutzen eingespielt wird. Dabei können die Verhandlungspartner natürlich alle möglichen anderen Aspekte mit einbeziehen, sei es das Versorgungsangebot, der Versorgungsbedarf und andere Effekte.

„Gesamt betrachtet ist erkennbar, dass die Nutzenbewertung des NICE weniger rigide ist als die des IQWiG. Auch wird entspannter mit dem Thema der Surrogate umgegangen."
Hans-Holger Bleß

Es gibt aber auch ganz andere Aspekte, die den Preis bilden und die Einfluss auf Kosten haben.

Bleß: Sicher. Bei der Preisbildung ist zuallererst die Anzahl der Patienten wichtig, die versorgt werden. Dabei gehe ich einmal davon aus, dass Entwicklungskosten eines Arzneimittels relativ vergleichbar sein werden, egal, ob nun 100 Patienten oder 100.000 Patienten damit versorgt werden.

Aber was ist die Folge?

Bleß: Bei einer kleinen Patientengruppe hat das neue Arzneimittel hohe inkrementelle Kosten und somit eine schlechte Kosteneffizienz. In der Logik der QALY kann das dazu führen, dass Menschen mit seltenen Erkrankungen nicht mehr versorgt werden, weil die Pharmaunternehmen die Kosten nicht einpreisen können, die sie eigentlich brauchen. Das hat das NICE erkannt und verlässt darum bei Orphans die indikationsübergreifende Bewertung.

Müller: Diese Problematik des QALY und die Entwicklung in UK sehe ich genauso.

Bleß: Was die Kosten anbetrifft, haben wir zudem im Moment die reine Kostenträgerperspektive und nur die sogenannten direkten (Krankheits-)Kosten betrachtet.

Aber was ist mit den gesellschaftlichen Kosten?

Bleß: Es gibt Indikationen, ich nenne mal als Beispiel das Rheuma, bei denen ein Großteil der finanziellen Krankheitslast auf der Gesellschaft oder anderen Kostenträgern lastet. Und zwar aufgrund von Arbeitsunfähigkeiten, Verrentungen und Pflege. Es ist eine ganz schwierige Situation, wenn die gesetzliche Krankenversicherung einen fairen Preis bilden soll, wobei sie eigentlich von dem finanziellen Nutzen gar nicht profitiert, weil die Einsparung ganz woanders verortet ist. Doch genau das wäre in der Kosten-Nutzen-Bewertung vorgesehen, dass die Perspektive der indirekten Kosten auf Antrag des G-BA mit einbezogen werden kann.

Wurde das schon mal gemacht?

Bleß: Nein.

Müller: Noch nicht. Das ist letztlich vom Gesetzgeber zwar so angelegt, aber die allermeisten Verfahren enden mit der Preisverhandlung. Die Stufe, von der wir hier reden, würde nach der Preisverhandlung kommen. Doch das hat bisher weder ein Unternehmer noch der GKV-SV beantragt.

Wie steht es denn in England mit der Zufriedenheit mit dem QALY?

Müller: Da gibt es paradoxe Effekte. Beispielsweise ermöglichen neue Produkte in der Onkologie oft einige Monate mehr Lebenserwartung, teilweise verbunden mit schlechterer Lebensqualität. Das führt dann dazu, dass die Produkte im QALY-Konzept so hoch bepreist werden, dass die finanzielle Grenze, die in England gesetzt worden ist, überschritten wird. Das wiederum hat dazu geführt, dass die Politik in England einen ergänzenden „Onkologie-Fonds" aufgesetzt hat. Zusätzlich dazu wurde bereits vor längerer Zeit eine „End-of-Life-Regelung" implementiert, welche die QALY-Grenze im onkologischen Bereich prolongiert. Doch generell muss man ganz klar feststellen, dass das QALY-Konzept dazu führt, dass einige neue Arzneimittel der Bevölkerung nicht zur Verfügung stehen. In Deutschland gibt es in der Onkologie überhaupt keine Einschränkungen. Bei uns werden keine Produkte aus Kostengründen vom Markt befördert, während das in den Ländern, die QALY einsetzen, sehr wohl der Fall ist.

Was aber anscheinend keinen großen gesellschaftlichen Widerstand auszulösen scheint.

Müller: Das kann ich nicht sicher beurteilen, gehe aber davon aus, dass es in England dazu Diskussionen gibt und die neue Regierung genau aus diesem Grund einen Onkologiefonds aufgelegt hat. Es gab gerade eine Veröffentlichung dazu, dass die durchschnittlichen medianen Überlebenszeiten der Krebspatienten in England im europäischen Vergleich schlechter sind. Dass das nun mit dem QALY zusammenhängt, würde ich jetzt nicht behaupten wollen, aber es ist deutlich festzustellen, dass in der Onkologie das QALY-Konzept nun einmal zu Einschränkungen führt, während England in anderen Bereichen wie Antibiotikagabe oder auch Hepatitis-Therapie interessanterweise auch sehr hohe Kosten akzeptiert, wenn es um sehr viele gewonnene Lebensjahre mit einer hohen Lebensqualität geht. Zumindest sind die Bewertungen des NICE recht positiv in diesen Bereichen.

Bleß: Wir haben am IGES Institut gerade intern eine vergleichende Analyse in der Onkologie erstellt. Dabei haben wir verglichen, was die G-BA und was die NICE-Spruchpraxis ausmacht. Ich fand diesen vertiefenden Blick wichtig, weil sich immer wieder einige Leute nach

dem NICE sehnen. Gesamt betrachtet ist erkennbar, dass die Nutzenbewertung des NICE weniger rigide ist als die des IQWiG. Auch wird entspannter mit dem Thema der Surrogate umgegangen. Will heißen: Generell wird in England schneller ein Nutzen anerkannt. Doch dann kommt die harte Schwelle der ökonomischen Betrachtung, durch die einige Produkte nicht in die Versorgung gelangen. Davon sind etliche Produkte betroffen, denen der G-BA einen beträchtlichen Zusatznutzen zuerkannt hat. Im Saldo haben wir gesehen, führt das deutsche Prinzip viel stärker dazu, dass gerade die onkologischen Produkte wirklich beim Patienten ankommen, als das in England der Fall ist.

Was ist mit der gesellschaftlichen Akzeptanz?

Bleß: Ich nehme an, dass fehlende Akzeptanz das NICE dazu gebracht hat, das QALY-Prinzip durch den onkologischen Fonds oder auch im Bereich der Orphan Drugs zu verletzen. Das QALY-Prinzip kennt eigentlich keine Ausnahmen, da es ja gerade den Anspruch hat, indikationsübergreifend gültig zu sein.

Dennoch belohnt das QALY-System Innovationen bis zu einem bestimmten Maß, eben 70.000 oder 100.000 Pfund oder Dollar oder was auch immer. Bei uns wird immer nur einen Abschlag diskutiert.

Müller: Nehmen wir eine Blinddarmoperation: Ein zehnjähriger Patient kommt frühzeitig zu einem guten Operateur und wird 80 Jahre alt. Im QALY-System hätte man dann 70 gewonnene Lebensjahre multipliziert mit einem angenommenen Wert von 40.000 Pfund, die dem Operateur dann zustehen würden. Das Beispiel zeigt gut, dass die QALY-Methodik bei der Preisbildung zu abwegigen Ergebnissen führen kann. Arzneimittel zur Hepatitis-Therapie werden deswegen so positiv von NICE bewertet, weil die Patienten geheilt werden.

Sie sagten: geheilt?

Müller: Ich weiß, dass man da verschiedener Auffassung sein kann. Wenn kein Virus mehr nachweisbar ist, sich keine Leberzirrhose entwickelt und keine Lebertransplantation erforderlich ist, kann man von Heilung sprechen. Die Nachhaltigkeit des Effekts muss aber noch abgewartet werden. Wenn man das nun auf gewonnene gute Lebensjahre extrapoliert, belohnt das QALY-Konzept diese Innovation. Anders dagegen in der Onkologie, in der ein Patient oft nur noch eine sehr kurze Lebensspanne bei schlechter Lebensqualität hat. Hier sieht die Rechnung dann ganz anders aus: Etwa 1 Monat mal 40.000 Pfund mal 0.001. Wenn ich diesen Preis nun ins Verhältnis setze zu dem Preis des Arzneimittels, dann schneiden diese Produkte im QALY-System eher schlecht ab. Deswegen kann man pauschal nicht sagen, dass Innovationen in dem einen oder in dem anderen System besser oder schlechter abschneiden. Der Vergleich mit Deutschland ist auch schwierig, weil es bei uns überhaupt noch keine gesellschaftliche Diskussion oder gar einen Konsens dazu gibt, wo eine mögliche Grenze wäre.

Herr Bleß, Sie haben in Ihrem Beitrag „Was gelernt und noch zu lernen ist“ in „Market Access & Health Policy“ (Ausgabe 05/14) geschrieben, dass dem G-BA und dem IQWiG eine Güterabwägung fremd sei. Das AMNOG-Verfahren ist – so heißt es in diesem Beitrag – „vorrangig darauf ausgerichtet, Fehler erster Art – also falsch positive Bewertungen – zu vermeiden. Hierbei werden allerdings Fehler zweiter Art – also zu Unrecht aberkannter Zusatznutzen – in Kauf genommen.“ Hiervon sind insbesondere Hersteller betroffen, die vor der Situation stehen, dass die Studien, die mit der Zulassungsbehörde vereinbart wurden, vor dem IQWiG und dem G-BA keine Berücksichtigung finden. Ein zu Unrecht aberkannter Zusatznutzen kann ja durchaus negative Folgen für Patienten entwickeln und müsste von daher ebenfalls dringlich zu vermeiden sein.

Bleß: Gerade durch die Anwendung der Regelung der evidenzbasierten Medizin wird darauf geachtet, dass keine falsch positiven Aussagen gemacht werden.

Also einem Produkt einen Nutzen zuzusprechen, den es hinterher dann gar nicht hat.

Bleß: Exakt. Wenn ich mir aber die Entscheidungskriterien von Zulassungsbehörden anschaue, dann gibt es da noch mehr in der Waagschale. Die Zulassungsbehörde hat einerseits zwar auch den Wunsch nach Evidenz, auch nach patientenrelevanten Endpunkten, aber sie betrachtet immer auch den vorhandenen Medical Need – also den Bedarf an einer Innovation in einer bestimmten Indikation. Damit stellen die Zulassungsbehörden die Frage, wie es zu bewerten ist, dass der Bevölkerung ein neues Arzneimittel zur Verfügung gestellt wird. Das findet seinen Gipfel in der bedingten Zulassung, die ganz offiziell besagt, dass das Fehlen an Evidenz niedriger zu bewerten ist als der Bedarf der Population an diesem neuen Arzneimittel.

Auch hier findet eine Güterabwägung zwischen zwei Werten ab.

Bleß: Durchaus. Worauf ich aber hinauswill, ist, dass das Zulassungsverfahren im Gegensatz zu G-BA und IQWiG diese Güterabwägung kennt. Im Bereich des Conditional Approval hat der G-BA seinen Spielraum auch genutzt, um keinen Widerspruch zu anderen Behörden herbeizuführen, aber dem Verfahren an sich ist das fremd. Ebenso könnte man einen patientenrelevanten Schaden anrichten, wenn der G-BA einen Zusatznutzen aberkennt und ein Arzneimittel beispielsweise dann vom Markt geht. Ich würde mir manchmal wünschen, dass das Verfahren zugesteht, sich auch irren zu können. Das ist gar nicht anders möglich, was alleine schon an den niedrigen Ergebnissicherheiten liegt. Wenn ich einen Hinweis auf einen Nutzen habe, dann ist die Möglichkeit eines Irrtums doch immanent. Und dazu muss man immer abwägen, was kann ich an Schaden anrichten, wenn ich diesen Zusatznutzen nicht gewähre. Doch diese Abwägungsentscheidung fehlt mir in diesem ganzen Verfahren.

„Dabei muss der G-BA laut Gesetzgeber evidenzbasiert entscheiden und in Hinblick auf das solidarisch finanzierte GKV-System dafür sorgen, dass das Maß des Notwendigen nicht überschritten wird.“
Dr. Thomas Müller

Müller: Das resultiert aus den unterschiedlichen Aufgabenstellungen. Bei der Zulassung geht es primär um die Produktsicherheit und den Konsumentenschutz. Es gilt zu vermeiden, dass Produkte auf den Markt kommen, die nicht sicher sind und keine gute Qualität haben. Wenn es um Wirksamkeitsendpunkte, also um die positiven Effekte auf die Therapie geht, setzt die Zulassungsbehörde teilweise niedrigere Hürden an als der Bundesausschuss bei seiner Nutzenprüfung. Der G-BA prüft, ob neue Produkte besser sind als der Standard. Dabei muss der G-BA laut Gesetzgeber evidenzbasiert entscheiden und in Hinblick auf das solidarisch finanzierte GKV-System dafür sorgen, dass das Maß des Notwendigen nicht überschritten wird. Diese Aufgabe kann die Zulassungsbehörde mit ihren ganz anderen gesetzlichen Rahmenbedingungen nicht erfüllen.

Die Herren Bleß und Müller, vielen Dank für das Gespräch. <<

Das Interview führten MVF-Herausgeber Prof. Dr. Reinhold Roski und MVF-Chefredakteur Peter Stegmaier.