

Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden 4.2

Version 1.0 vom 22.04.2015

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8 (KölnTurm)
50670 Köln

Tel.: +49 (0)221 – 35685-0

Fax: +49 (0)221 – 35685-1

E-Mail: methoden@iqwig.de

Internet: www.iqwig.de

Inhaltsverzeichnis

	Seite
Abkürzungsverzeichnis.....	iv
1 Dokumentation der Anhörung	1
2 Würdigung der Anhörung	2
2.1 Würdigung der Stellungnahmen zu Kapitel 1 „Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen“.....	2
2.2 Würdigung der Stellungnahmen zu Abschnitt 2.1 „Produktspezifische Verfahrensabläufe“.....	2
2.3 Würdigung der Stellungnahmen zu Abschnitt 2.1.4 „Kosten-Nutzen- Bewertung nach § 35b SGB V“.....	3
2.4 Würdigung der Stellungnahmen zu Abschnitt 2.1.5 „Potenzialbewertung“.....	3
2.5 Würdigung der Stellungnahmen zu Abschnitt 2.1.7 „Gesundheitsinformationen“.....	4
2.6 Würdigung der Stellungnahmen zu Abschnitt 2.2.1 „Auswahl externer Sachverständiger“.....	5
2.7 Würdigung der Stellungnahmen zu Abschnitt 3.1.1 „Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens“ und Abschnitt 3.2.4 „Patientenberichtete Endpunkte“.....	6
2.8 Würdigung der Stellungnahmen zu Abschnitt 3.1.5 „Zusammenfassende Bewertung“.....	6
2.9 Würdigung der Stellungnahmen zu Abschnitt 3.5 „Diagnostische Verfahren“.....	7
2.10 Würdigung der Stellungnahmen zu Abschnitt 3.8 „Potenzialbewertung“.....	10
2.11 Würdigung allgemeiner Stellungnahmen zu Kapitel 4 „Kosten-Nutzen- Bewertung medizinischer Interventionen“.....	10
2.12 Würdigung der Stellungnahmen zu Abschnitt 4.1 „Einführung“.....	11
2.13 Würdigung der Stellungnahmen zu Abschnitt 4.1.1 „Rechtliche Grundlagen für die Kosten-Nutzen-Bewertung nach SGB V“.....	11
2.14 Würdigung der Stellungnahmen zu Abschnitt 4.1.2 „Perspektive“.....	12
2.15 Würdigung der Stellungnahmen zu Abschnitt 4.1.3 „Zeithorizont“.....	13
2.16 Würdigung der Stellungnahmen zu Abschnitt 4.1.4 „Auswahl der Komparatoren“.....	13
2.17 Würdigung der Stellungnahmen zu Abschnitt 4.1.5 „Versorgungspfad“.....	13
2.18 Würdigung der Stellungnahmen zu Abschnitt 4.1.7 „Datengrundlage“.....	14
2.19 Würdigung der Stellungnahmen zu Abschnitt 4.1.9 „Interpretation der Ergebnisse“.....	14

2.20	Würdigung der Stellungnahmen zu Abschnitt 4.2 „Modellierung“.....	15
2.21	Würdigung der Stellungnahmen zu Abschnitt 4.3 „Nutzen“	15
2.22	Würdigung der Stellungnahmen zu Abschnitt 4.4 „Kosten“.....	19
2.23	Würdigung der Stellungnahmen zu Abschnitt 4.5 „Epidemiologische Daten“ ..	21
2.24	Würdigung der Stellungnahmen zu Abschnitt 4.7 „Unsicherheit (Sensitivitätsanalysen)“	21
2.25	Würdigung der Stellungnahmen zu Abschnitt 4.8 „Ausgaben-Einfluss- Analyse (Budget-Impact-Analyse)“.....	22
2.26	Würdigung der Stellungnahmen zu Abschnitt 4.9 „Besonderheiten der Kosten-Nutzen-Bewertung nach § 35b SGB V“.....	22
2.27	Würdigung der Stellungnahmen zu Kapitel 6 „Evidenzbasierte Gesundheitsinformation für Bürgerinnen und Bürger“	23
2.28	Würdigung der Stellungnahmen zu Abschnitt 6.3.4 „Aktualisierung“	25
2.29	Würdigung der Stellungnahmen zu Kapitel 7 „Informationsbeschaffung“.....	25
2.30	Würdigung der Stellungnahmen zu Abschnitt 7.2 „Überprüfung der Informationsbeschaffung“.....	27
2.31	Würdigung der Stellungnahmen zu Abschnitt 8.1.3 „Rangordnung verschiedener Studienarten/ Evidenzgrade“	27
2.32	Würdigung der Stellungnahmen zu Abschnitt 8.3.3 „Beurteilung klinischer Relevanz“	27
3	Literaturverzeichnis	29
4	Offenlegung potenzieller Interessenkonflikte	33
4.1	Potenzielle Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen, Firmen und Privatpersonen	33
4.2	Potenzielle Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (wissenschaftlicher Beirat)	35
5	Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll.....	38
5.1	Teilnehmerliste der wissenschaftlichen Erörterung.....	38
5.2	Liste der Stellungnahmen, zu denen kein Vertreter an der wissenschaftlichen Erörterung teilgenommen hat	39
5.3	Tagesordnung der wissenschaftlichen Erörterung	40
5.4	Protokoll der wissenschaftlichen Erörterung.....	40
	Anhang A – Dokumentation der Stellungnahmen	107

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AHP	Analytic Hierarchy Process (analytischer Hierarchieprozess)
AMNOG	Arzneimittelmarktneuordnungsgesetz
AOK	Allgemeine Ortskrankenkasse
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V.
BMG	Bundesministerium für Gesundheit
CA	Conjoint-Analyse
DCE	Discrete-Choice-Experiment
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
EDSS	Expanded Disability Status Scale
G-BA	Gemeinsamer Bundesausschuss
GKV	gesetzliche Krankenversicherung
HTA	Health Technology Assessment
IQR	Interquartilsregion
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
KNB	Kosten-Nutzen-Bewertung
NHB	Net Health Benefit
NICE	National Institute for Health and Care Excellence
NMB	Net Monetary Benefit
PKV	private Krankenversicherung
pU	pharmazeutischer Unternehmer
QALY	Quality-Adjusted Life Year (qualitätsadjustiertes Lebensjahr)
RCT	Randomized controlled trial (randomisierte kontrollierte Studie)
RR	relatives Risiko
SGB	Sozialgesetzbuch
SSRI	Selective Serotonin Reuptake Inhibitor (selektive Serotonin-Wiederaufnahmehemmer)
WHO	World Health Organization (Weltgesundheitsorganisation)
WiDO	Wissenschaftliches Institut der AOK

1 Dokumentation der Anhörung

Am 18.06.2014 wurde der Entwurf der Allgemeinen Methoden Version 4.2 veröffentlicht und zur Anhörung gestellt. Bis zum 07.08.2014 konnten schriftliche Stellungnahmen eingereicht werden. Insgesamt wurden 31 Stellungnahmen form- und fristgerecht abgegeben. Diese Stellungnahmen sind im Anhang abgebildet.

Unklare Aspekte in den schriftlichen Stellungnahmen wurden in einer wissenschaftlichen Erörterung am 01.10.2014 im Institut diskutiert. Das Wortprotokoll der Erörterung befindet sich in Kapitel 5.

2 Würdigung der Anhörung

Im Stellungnahmeverfahren zum Dokument „Allgemeine Methoden: Entwurf für Version 4.2“ vom 18.06.2014 gingen bis zum 07.08.2014 Stellungnahmen von 31 Organisationen, Institutionen, Firmen und Einzelpersonen ein. Alle Stellungnahmen wurden in der Überarbeitung des Methodenpapiers berücksichtigt. In diesem Dokument werden ausschließlich die Kommentare aus den Stellungnahmen gewürdigt, die sich auf die aktualisierten bzw. neuen Abschnitte der Allgemeinen Methoden beziehen. Sofern sich aus den Stellungnahmen Änderungen in diesen Abschnitten ergaben, ist dies im vorliegenden Dokument explizit beschrieben.

2.1 Würdigung der Stellungnahmen zu Kapitel 1 „Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen“

In einigen Stellungnahmen wird vorgeschlagen, in Kapitel 1 nach den allgemeinen Ausführungen zur evidenzbasierten Medizin auch allgemeine Ausführungen zur Gesundheitsökonomie einzufügen, da beide auch im Sozialgesetzbuch (§ 35b SGB V) erwähnt werden.

Diese Anregung wurde aufgenommen, und in Kapitel 1 wurden allgemeine Ausführungen zu den internationalen Standards der Gesundheitsökonomie ergänzt.

2.2 Würdigung der Stellungnahmen zu Abschnitt 2.1 „Produktspezifische Verfahrensabläufe“

Insbesondere bezüglich der neuen IQWiG-Produkte Kosten-Nutzen-Bewertung (KNB) und Potenzialbewertung, aber auch übergreifend wird in einigen Stellungnahmen die nicht einheitliche Praxis der Veröffentlichung von IQWiG-Produkten kritisch kommentiert. Entsprechend, d. h. für verschiedene Produkte, wurde außerdem die Angemessenheit der Beteiligungsmöglichkeiten der (Fach-)Öffentlichkeit, darunter die Terminierung von Anhörungsverfahren sowie die Frage der Einbindung externer Sachverständiger in die Produkterstellung, hinterfragt.

Grundsätzlich wird die schnellstmögliche Veröffentlichung aller IQWiG-Produkte bzw. zur Anhörung zu stellender Vorläufer (z. B. Vorbericht und vorläufiger Berichtsplan) angestrebt. Produktspezifische Verfahrensvarianten ergeben sich aus verbindlichen Vorgaben zum Bewertungsverfahren gemäß Verfahrensordnung des G-BA, denen das IQWiG unterworfen ist. Dies kann z. B. zu verzögerter Veröffentlichung (Addendum, Dossierbewertung) bis hin zu einem Verzicht auf Veröffentlichung (Potenzialbewertung) führen. Weiterhin kann aus denselben Gründen die Entscheidung gefällt werden, die Terminierung von Stellungnahmefristen produktspezifisch abweichend gegenüber denen bei Berichten zu verkürzen (Kosten-Nutzen-Bewertung nach § 35b SGB V; siehe hierzu auch folgenden Abschnitt 2.3) oder auf ein IQWiG-seitiges Anhörungsverfahren zu verzichten (das Anhörungsverfahren bezüglich Dossierbewertungen wird vom G-BA durchgeführt).

Ebenfalls gesetzliche Anforderungen begründen die Möglichkeiten, ob bzw. wie intensiv externe Sachverständige in den Erstellungsprozess von IQWiG-Produkten eingebunden werden können. Dies trifft insbesondere auf Potenzial- und Dossierbewertungen zu, in denen regelhaft (hoch) vertrauliche, als „Betriebs- und Geschäftsgeheimnisse“ gekennzeichnete Dokumente übermittelt werden. Abhängig vom Erstellungsprozess des jeweiligen Produkts kann ein sinnvoller Einbezug externer Sachverständiger nur teilweise durch die Vergabe von Forschungsaufträgen realisiert werden. Zuweilen ist dies im Rahmen der engen Fristen und Vorgaben nur durch zielgerichtete Beratungsanfragen zu medizinisch-fachlichen Fragestellungen möglich. Es ergibt sich daher kein Änderungsbedarf.

2.3 Würdigung der Stellungnahmen zu Abschnitt 2.1.4 „Kosten-Nutzen-Bewertung nach § 35b SGB V“

In einigen Stellungnahmen wird die Stellungnahmefrist von 3 Wochen bei Kosten-Nutzen-Bewertungen nach § 35b SGB V kritisiert.

Für Kosten-Nutzen-Bewertungen gilt gemäß § 35b Absatz 1 Satz 3 SGB V der § 35a Absatz 1 Satz 3 und Absatz 2 Satz 3 SGB V entsprechend. In § 35a SGB V wird eine Frist von 3 Monaten bis zum Abschluss der Nutzenbewertung vorgegeben. Übertragen auf die Kosten-Nutzen-Bewertung bedeutet dies, dass zwischen der Veröffentlichung des Vorberichts und des Abschlussberichts von Kosten-Nutzen-Bewertungen nach § 35b SGB V – dies gilt wohlbemerkt nicht für Kosten-Nutzen-Bewertungen, die vom G-BA nicht im Rahmen des § 35b SGB V, sondern als Aufträge nach § 139a Nr. 5 SGB V an das IQWiG gegeben werden – 3 Monate liegen. Diese kurze Bearbeitungsfrist erfordert auch eine kürzere Stellungnahmefrist. Für Kosten-Nutzen-Bewertungen, die nicht dem § 35b SGB V folgen, gelten die sonst üblichen Stellungnahmefristen für Berichte. Zu den Fristen einer Kosten-Nutzen-Bewertung nach § 35b SGB V kann im Übrigen die Verfahrensordnung des G-BA [13] eingesehen werden. An der entsprechenden Stelle wurde im Methodenpapier ein Verweis auf das Gesetz eingefügt.

In einigen Stellungnahmen wird kritisiert, dass eine mündliche wissenschaftliche Erörterung nur optional vorgesehen sei.

Wie schon ausgeführt, handelt es sich bei den Kosten-Nutzen-Bewertungen nach § 35b SGB V um ein Verfahren, das zwischen Vorbericht und Abschlussbericht nur eine Frist von 3 Monaten erlaubt (siehe 5. Kapitel § 31 Verfahrensordnung des G-BA [13]). Ein entsprechender Verweis auf die Verfahrensordnung wurde eingefügt.

2.4 Würdigung der Stellungnahmen zu Abschnitt 2.1.5 „Potenzialbewertung“

In einer Stellungnahme wird kritisiert, dass unklar bleibt, in welchen Fällen tatsächlich eine eigene Recherche durch das IQWiG vorgesehen ist und wie weit diese geht. Hierbei fordert eine Stellungnahme, dass „eine Recherche des medizinischen Hintergrundes nicht nur optional durchzuführen“ sei. Zweitens sei nicht geklärt, wie das Verhältnis zwischen eigener Recherche und Recherche des Antragstellers zu sehen ist und welche Grenzen hierbei

gezogen werden bzw. gezogen werden müssen. Dies gelte umso mehr, als das Methodenpapier selbst betont, dass Unterschiede zur Nutzenbewertung bestehen, sodass fraglich ist, ob auf die allgemeinen Grundsätze der Nutzenbewertung zurückgegriffen werden kann.

Der Hinweis auf mögliche eigene Recherchen erfolgte im Methodenpapier mit der expliziten Erläuterung, dass diese optional zur Unterstützung der Bewertung erfolgen. Dies kann beispielsweise der Fall sein, um sich in das Antragsthema einzuarbeiten oder wenn konkrete Fragen zur Operationalisierung von Endpunkten etc. bestehen. Eine Recherche zum medizinischen Hintergrund ist jedoch nicht regelhaft erforderlich unter anderem, weil es vorkommen kann, dass sich Anträge auf bereits vom Institut bewertete Themen erstrecken.

Gemäß § 137e SGB V ist es die Pflicht des Antragstellers, „aussagekräftige Unterlagen vorzulegen, aus denen hervorgeht, dass die Methode hinreichendes Potenzial für eine Erprobung bietet“. Hieraus ergibt sich, dass Anträge nach § 137e SGB V dahin gehend zu prüfen sind, ob die vorgelegten Unterlagen eine hinreichend vollständige Datengrundlage für eine Potenzialbewertung bieten können. Auch für eine solche Prüfung kann es erforderlich sein, eigene Recherchen zu spezifischen Aspekten durchzuführen. Es ist jedoch nicht Aufgabe des Institutes, bei unzureichender Aussagekraft der Unterlagen die vorhandenen Mängel nachzubessern.

Es ergibt sich daher kein Änderungsbedarf.

2.5 Würdigung der Stellungnahmen zu Abschnitt 2.1.7 „Gesundheitsinformationen“

In einer Stellungnahme wird gewünscht, klare Regeln für das Einbeziehen von externen Reviews zu bestimmen und offen darzulegen.

In Abschnitt 2.1.7 ist die Regelung für neu erstellte Gesundheitsinformationen beschrieben. Mit Ausnahme der Begleitinformationen zu IQWiG-Produkten werden die Textentwürfe nach der ressortinternen Qualitätssicherung in ein externes Review gegeben. Die Begleitinformationen, die ausschließlich auf IQWiG-Produkten basieren, werden durch die erstellenden Ressorts begutachtet. Diese Regelung ist eindeutig und transparent. Dies wurde im Methodenpapier verdeutlicht.

Die Auswahl der externen Reviewer folgt den in Abschnitt 2.2.1 beschriebenen Grundsätzen. Daher wurde die Aufzählung dort um Gesundheitsinformationen erweitert.

In einer Stellungnahme wird kritisiert, dass die Möglichkeit zur offiziellen Stellungnahme zu den Gesundheitsinformationen auf Auftraggeber, Kuratorium und weitere Gremien des IQWiG beschränkt sei. Es fehle die Möglichkeit, dass auch Personen und Institutionen in der Wissenschaft, Hersteller pharmazeutischer Produkte und Medizinprodukte und sonstige Interessierte Stellung nehmen können.

In § 8 Abs. 4 der Stiftungssatzung ist die Möglichkeit zur Anhörung zu insbesondere Gesundheitsinformationen wie folgt geregelt:

„Vor der Abgabe von Empfehlungen des Instituts zu Angelegenheiten, welche unmittelbar die Patientenbelange tangieren, insbesondere zur Bereitstellung verständlicher allgemeiner Informationen zur Qualität und Effizienz in der Gesundheitsversorgung, ist den im Kuratorium vertretenen Repräsentanten der Patientinnen und Patienten sowie den Beauftragten der Bundesregierung für die Belange der Patientinnen und Patienten Gelegenheit zur Stellungnahme im Rahmen einer Anhörung zu geben.“

Diese Möglichkeit der schriftlichen Anhörung wurde auf alle Mitglieder des Kuratoriums erweitert. Im Kuratorium sind die Selbstverwaltungsorgane der Trägerorganisationen des G-BA, maßgebliche nicht im G-BA vertretene Organisationen von Leistungserbringern und den Sozialpartnern und sonstige für das Gesundheitswesen relevante Organisationen repräsentiert, darunter auch Verbände der Pharmaindustrie. Es liegt in der Verantwortung der Vertreter, die Stellungnahme der von ihnen vertretenen Organisationen oder Institutionen zu organisieren oder diese zur Abgabe einer Stellungnahme anzuregen. So organisiert beispielsweise die AWMF die Anhörung der medizinischen Fachgesellschaften.

Es besteht daher kein Änderungs- oder Ergänzungsbedarf.

In einer Stellungnahme wird nach weiteren Kanälen gefragt, über die die Gesundheitsinformationen des IQWiG der Öffentlichkeit zur Verfügung gestellt werden.

Die Gesundheitsinformationen des IQWiG werden der Öffentlichkeit primär über die Website www.gesundheitsinformation.de zur Verfügung gestellt, im Rahmen von Kooperationsverträgen, aber auch über die Websites der Kooperationspartner. Zudem werden ausgesuchte Gesundheitsinformationen auch in Form von Broschüren und Flyern aufgelegt. Es besteht kein Änderungs- oder Ergänzungsbedarf.

2.6 Würdigung der Stellungnahmen zu Abschnitt 2.2.1 „Auswahl externer Sachverständiger“

Die Art der Ausschreibung bei Beteiligung externer Sachverständiger in die Produkterstellung wird in verschiedenen Stellungnahmen kommentiert.

Das IQWiG strebt die Beteiligung externer Sachverständiger an der wissenschaftlichen Produkterstellung an. Diesbezüglich ist bereits unter Abschnitt 2.2 erläutert, dass die produktspezifischen Möglichkeiten einer sinnvollen Einbindung auch von gesetzlichen Rahmenbedingungen sowie ggf. auch von Vorgaben durch die Verfahrensordnung des G-BA abhängen.

Die Darstellung zur Vergabe wissenschaftlicher Forschungsaufträge wurde als Reaktion auf die Stellungnahmen noch einmal angepasst, indem klargestellt wird, dass (weiterhin) Auftragsbekanntmachungen für Forschungsaufträge nach § 139b Abs. 3 SGB V auf der Website veröffentlicht werden, es jedoch begründete Ausnahmefälle geben kann, z. B. bei einem besonders eilbedürftigen Auftrag. Daneben besteht weiterhin das auf der Website

des IQWiG publizierte Angebot zur Beteiligung bei Beratungen zu medizinisch-fachlichen Fragestellungen über die Sachverständigen-Datenbank.

2.7 Würdigung der Stellungnahmen zu Abschnitt 3.1.1 „Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens“ und Abschnitt 3.2.4 „Patientenberichtete Endpunkte“

In mehreren Stellungnahmen wird kritisiert, dass im Entwurf zu den Methoden 4.2 der Endpunkt Patientenzufriedenheit aus der Liste der ggf. ergänzend darzustellenden Endpunkte gestrichen wurde. Die Patientenzufriedenheit könne einen Hinweis auf die subjektive Akzeptanz der zu bewertenden Intervention und auch auf eine möglicherweise erhöhte Lebensqualität geben. In einer Stellungnahme wird diese Streichung hingegen explizit begrüßt, da dadurch eine mögliche Stärkung zur verbesserten Implementierung der Erfassung der Lebensqualität erreicht werden könne mit dem Ziel, dass Studien den Endpunkt Lebensqualität in Zukunft stärker berücksichtigen.

Als Konsequenz aus den Stellungnahmen und der Diskussion in der wissenschaftlichen Erörterung zu diesem Endpunkt wird der Endpunkt Patientenzufriedenheit auch in den Methoden 4.2 als ggf. ergänzend zu bewertender Endpunkt beschrieben. Allerdings wird klargestellt, dass dies nur dann möglich ist, wenn durch die Patientenzufriedenheit im konkreten Bewertungsfall gesundheitsbezogene Aspekte abgebildet werden. Ergänzend ist darauf hinzuweisen, dass Patientenzufriedenheit und Lebensqualität unterschiedliche Konzepte sind [2]. Dies wird in Studienpublikationen wiederholt nicht beachtet, insbesondere werden Fragebögen zur Patientenzufriedenheit teilweise unter der Rubrik Lebensqualität beschrieben. Aus Veränderungen in der Patientenzufriedenheit ergeben sich aber nicht zwangsläufig Veränderungen in der Lebensqualität.

Zudem wird in einigen Stellungnahmen kritisiert, dass im Entwurf der Methoden 4.2 die Ausführungen aus § 35b SGB V zu patientenrelevanten Endpunkten aus Abschnitt 3.1.1 gelöscht wurden.

Die Angaben in § 35b SGB V führen nicht zu einer inhaltlichen Änderung der Methoden (und damit auch nicht die im Entwurf der Methoden 4.2 vorgenommene Streichung dieser Angaben). Im Gegenteil bestätigen die Ausführungen in § 35b SGB V das in den Methoden beschriebene Vorgehen. Aufgrund der Stellungnahmen wird der zuvor gestrichene Absatz jedoch wieder ergänzt.

2.8 Würdigung der Stellungnahmen zu Abschnitt 3.1.5 „Zusammenfassende Bewertung“

Es gab einige Stellungnahmen zum Abschnitt 3.1.5, die auch den Abschnitt 4.3 betreffen. Diese werden in Abschnitt 2.21 dieses Dokuments gewürdigt.

2.9 Würdigung der Stellungnahmen zu Abschnitt 3.5 „Diagnostische Verfahren“

Der Grundsatz, dass diagnostische (oder prognostische) Informationen nur dann nutzbringend sind, wenn sich therapeutische oder präventive Konsequenzen ergeben, wird in mehreren Stellungnahmen kritisch hinterfragt. Spezifisch wurde vorgeschlagen, diagnostischen oder prognostischen Interventionen auch dann einen Nutzen zuzusprechen,

- 1) wenn sie es dem Patienten ermöglichen, „seine Lebensplanung dementsprechend“ anzupassen,
- 2) wenn sie für die Definition, die Klassifikation oder die Kodierung von Erkrankungen erforderlich sind oder
- 3) wenn sie den Bedarf für weitere diagnostische Interventionen reduzieren.

Wie in der mündlichen Erörterung am 01.10.2014 erläutert, sieht das Sozialrecht eine GKV-Leistungspflicht nur dann vor, wenn die fragliche Intervention einen Effekt auf Mortalität, Morbidität oder gesundheitsbezogene Lebensqualität hat. Daneben ist – wie von den Stellungnehmenden angemerkt – die Wirtschaftlichkeit der Behandlung zu berücksichtigen. Der unter a) skizzierte Effekt einer Intervention bezieht sich jedoch auf einen allgemeinen, nicht gesundheitlichen Vorteil. Daher erscheint dieses Argument nicht tragfähig.

Die unter b) genannten Konzepte verbessern vermutlich die Beschreibung von Krankheitszuständen, besitzen aber nur dann eine Patientenrelevanz, wenn diese genauere Beschreibung wiederum geeignet ist, über eine stärker zielgerichtete Behandlung Mortalität, Morbidität oder Lebensqualität zu verbessern.

Bei dem unter c) angeführten Argument liegt vermutlich ein Missverständnis vor. Der Nutzen eines diagnostischen Tests kann in der Tat darin bestehen, dass der Bedarf für weitere diagnostische Interventionen reduziert wird. Dieser Sachverhalt ist im Methodenpapier bereits adäquat abgebildet, wenn erläutert wird, dass „direkte patientenrelevante Vorteile“ eines neuen Tests im Rahmen der Nutzenbewertung zu berücksichtigen sind. Auf diese Weise kann ein neuer diagnostischer Test einen oder mehrere etablierte Tests ersetzen, die ihrerseits jedoch wieder einen Einfluss auf Therapie (oder Prävention) und damit auf patientenrelevante Endpunkt haben müssen.

Um besser zu verdeutlichen, dass diagnostische und prognostische Tests nicht nur über die Steuerung von Therapie und Prävention, sondern auch über die Steuerung weiterer Diagnostik einen Nutzen erbringen können, wurde die Formulierung angepasst. Die Forderung, dass Diagnostik „therapeutische oder präventive Konsequenzen“ haben müsse, wurde erweitert auf „medizinische Konsequenzen“ allgemein.

In einer Stellungnahme wird angemerkt, dass „prospektiv geplante, retrospektiv durchgeführte Studien“ eine hohe Bedeutung und Validität in der Bewertung diagnostischer und prognostischer Marker hätten. Solche Studien verwendeten meist „molekularbiologische Analysen an asserviertem Tumorgewebe“, die im Rahmen eines RCT gesammelt wurden. Die Stellungnehmenden schlagen eine explizite Aufnahme dieses Studientyps in das Methodenpapier sowie spezielle Qualitätskriterien vor.

Wie in der Stellungnahme angeregt wurde das sogenannte „prospektiv-retrospektive“ Design explizit im Methodenpapier aufgenommen und referenziert. Ebenfalls umgesetzt wurde der Vorschlag, weitere Qualitätsaspekte (hohe Vollständigkeit der Daten, Aktualität der durchgeführten Therapie) in das Methodenpapier aufzunehmen.

In derselben Stellungnahme wird gefordert, dass bei Vorliegen von RCT-Ergebnissen zum Nutzen einer diagnostischen Methode zusätzlich zu prüfen sei, ob der diagnostische Test „qualitätsgesichert“ sei. Ähnlich sieht auch eine zweite Stellungnahme RCT-Ergebnisse als bedeutungslos an, sofern nicht vorher die Genauigkeit des diagnostischen Tests evaluiert worden sei. In einer dritten Stellungnahme wird konträr dazu ausgeführt, dass manche RCT-Designs auch dann positive Ergebnisse aufweisen könnten, wenn der diagnostische Test keinerlei Diskriminationsfähigkeit besitze.

Es ist unstrittig, dass ein diagnostischer Test in aller Regel eine hinreichend hohe Diskriminationsfähigkeit aufweisen muss, um Patienten sinnvoll der einen oder anderen Behandlung zuordnen zu können. Daher wird bei der Entwicklung neuer diagnostischer Tests meist zuerst die Testgüte untersucht, bevor ein RCT geplant wird. Dennoch ist die Kenntnis der Testgüte für die Bewertung von RCT-Ergebnissen nicht notwendigerweise erforderlich. Sofern sich in einem RCT im Interaktionsdesign tatsächlich eine Interaktion zwischen Testergebnis und Therapieeffekt nachweisen lässt, ist gleichzeitig auch der Nachweis erbracht, dass der diagnostische Test über eine offensichtlich hinreichende (wenn auch vielleicht nicht quantifizierbare) Testgüte verfügt. Lediglich bei einem RCT im Strategiedesign ist es theoretisch denkbar, dass sich ein scheinbarer Nutzen eines diagnostischen Tests zeigt, obwohl der Test lediglich eine Ratewahrscheinlichkeit besitzt und die Patienten zufällig in Testpositive und -negative unterteilt: Sofern nämlich im Vergleichsarm ohne Testdurchführung immer eine Standardtherapie erfolgt, im Interventionsarm aber ein Teil der Patienten aufgrund des (zufälligen) Testergebnisses eine effektivere Alternativtherapie erhalten, kann der Fehlschluss gezogen werden, der Effekt zugunsten der neuen Behandlungsstrategie würde auf der diagnostischen Intervention beruhen. Als zweite Besonderheit des Strategiedesigns muss bedacht werden, dass die Durchführung des Tests selbst einen direkten Einfluss auf den beobachteten Endpunkt haben kann, sodass sich die Wirkungen der diagnostischen und der therapeutischen Intervention addieren. Dennoch sind diese Situationen vermutlich selten, sodass es korrekt ist, im Methodenpapier zu beschreiben, dass in der Regel beide Interventionen (Diagnostik und Therapie) ihre Ziele (Diskrimination und Wirksamkeit) erfüllen müssen, um einen Nutzen nachweisen zu können. Die Kenntnis der Testgüte bleibt damit nachrangig.

Auch wenn die Stellungnahmen zahlreiche hilfreiche Argumente enthalten, in welchen klinischen Situationen welches randomisierte Design Vor- oder Nachteile bei der Evaluation einer diagnostischen Intervention bietet, so sind diese Ausführungen doch relevanter bei der Planung als bei der Interpretation einer Studie. Auch würde es den Rahmen des Methodenpapiers sprengen, alle spezifischen Vor- und Nachteile der verschiedenen Designs auszuführen. Für das Methodenpapier ergibt sich daher kein Änderungsbedarf.

In einer Stellungnahme wird argumentiert, dass die gemeinsame Betrachtung von Diagnostik und Therapie sinnvoll sei, dass daher aber jede Nutzenbewertung einer diagnostischen Intervention in Wahrheit eine Nutzenbewertung einer diagnostisch-therapeutischen Kombination sei, sodass das Ergebnis der Bewertung allein für die untersuchte Kombination Gültigkeit besitze.

Die Argumentation des Stellungnehmenden ist nachvollziehbar. Feste Kopplungen zwischen diagnostischen und therapeutischen Interventionen finden sich bereits in einigen Nutzenbewertungen [16,18]. Auch bei vielen weiteren beauftragten Themen ist unstrittig, welche therapeutischen Konsequenzen sich aus einem positiven oder negativen Ergebnis einer diagnostischen Intervention ergeben. Dennoch ist es medizinisch in vielen Fällen geboten, die Bewertung diagnostischer Interventionen ohne Eingrenzung der medizinischen Konsequenzen durchzuführen, insbesondere wenn diese Konsequenzen vielfältig sein können (beispielsweise Chirurgie, Radiotherapie, Chemotherapie oder palliative Therapie als mögliche Konsequenzen einer onkologischen Staging-Diagnostik). Ferner erscheint es unrealistisch, regelmäßig in der Behandlung einzelner Patienten (und damit auch in der Erstattung der Leistungen) eine feste Kombination aus Diagnostik und Therapie vorzunehmen, weil es meist sinnvolle und gängige Praxis ist, das therapeutische Vorgehen nach Durchführung des diagnostischen Tests neu zu planen. Aus diesem Argument ergibt sich demnach keine Notwendigkeit einer Änderung.

Die Aussagekraft von Anreicherungsdesigns wird in einer Stellungnahme hervorgehoben. Es wird argumentiert, dass ein Nutzen einer diagnostischen Maßnahme bereits darin gesehen werden könne, wenn die Diagnostik zumindest einen Teil der Patienten einer wirksameren Therapie zuführe – unabhängig davon, ob vielleicht auch der übrige Teil der Patienten einen vergleichbaren Therapienutzen hätte. Entscheidend sei der „Vergleich zur bisherigen Praxis“.

Wie im Methodenpapier bereits dargelegt, hängt die Aussagekraft des Anreicherungsdesigns entscheidend damit zusammen, welche Annahmen bezüglich des möglichen Therapieeffektes in der nicht randomisierten Gruppe plausibel sind. Theoretisch ist es vorstellbar, dass ein Therapienutzen im nicht randomisierten Anteil der Patienten ohnehin realiter nicht umsetzbar ist, weil klare Gründe dagegen sprechen. Eine solche Situation liegt beispielsweise dann vor, wenn eine Diagnostik auch der Lokalisation einer Erkrankung dient: Bei testnegativen Personen könnten bestimmte Therapien (z. B. Resektion) aufgrund der fehlenden Lokalisation gar nicht angewendet werden. In anderen Situationen jedoch wird man nicht davon ausgehen können, dass sich eine Therapie auch bei Testnegativen verbietet. Daher ist weniger der Vergleich zur bisherigen Praxis, sondern vielmehr die grundsätzliche Möglichkeit der Therapie wesentlich. Insgesamt erscheinen die spezifischen Prämissen, die bei der Interpretation von RCTs im Anreicherungsdesign zu machen sind, im Methodenpapier ausreichend beschrieben.

2.10 Würdigung der Stellungnahmen zu Abschnitt 3.8 „Potenzialbewertung“

In mehreren Stellungnahmen wird kritisiert, dass die neu hinzugefügten Kategorien der Ergebnissicherheit (gering, sehr gering und minimal) einer „sprachlichen Herabwürdigung“ oder „Stigmatisierung der vorliegenden Daten“ gleichkämen.

Die Bezeichnungen der Ergebnissicherheit nicht randomisierter Studien waren Thema der wissenschaftlichen Erörterung am 01.10.2014. Weder aus den Stellungnahmen noch in der Erörterung ergaben sich konkrete Vorschläge zu einer andersartigen Benennung. Auch zeigte sich, dass im Bereich der klinischen Forschung unterhalb von Fallserien und Fallberichten keine Studiendesigns eine noch geringere interne Validität aufweisen, sodass die Kategoriebezeichnung „minimale Ergebnissicherheit“ offensichtlich zutreffend ist. Diese Bezeichnung stellt überdies keine Herabwürdigung dar, weil nicht vergleichenden Studien außerhalb von Potenzial- und Nutzenbewertung durchaus eine im Einzelfall große Bedeutung zugemessen werden kann [1]. Für das Methodenpapier ergibt sich daher kein Änderungsbedarf.

Im Zusammenhang mit der Ergebnissicherheit werden auch die vorgeschlagenen Effektstärken von 0,5 und 0,8 (als Grenzwerte für das RR) hinterfragt. Während eine Stellungnahme es als „nicht angemessen“ beschreibt, „generelle Grenzen für das relative Risiko anzusetzen“, empfindet ein anderer Stellungnehmender die Grenzen als „nicht geeignet“, weil „genaue und klare Abgrenzungen“ erforderlich seien.

Neben den widersprüchlichen Vorschlägen der Stellungnehmenden für und wider Grenzwerte enthalten die Stellungnahmen keinen konkreten Ansatzpunkt oder Vorschlag, wie stattdessen vorgegangen werden könnte. Es ergibt sich daher kein Änderungsbedarf.

In einer Stellungnahme wird die Frage aufgeworfen, wie sich aus Ergebnissen zu verschiedenen Endpunkten ein Gesamtmaß des Potenzials bestimmen lasse. Auch in einer zweiten Stellungnahme vermissten die Stellungnehmenden eine „hinreichende Bestimmtheit und Klarheit“ des Potenzialbegriffs.

Die Festlegung eines Gesamtpotenzialmaßes erscheint zurzeit weder möglich noch erforderlich. Auch in den Stellungnahmen werden keine Vorschläge dazu unterbreitet, wie der Potenzialbegriff präziser festzulegen wäre. Daher ergeben sich hieraus keine Änderungen für das Methodenpapier.

2.11 Würdigung allgemeiner Stellungnahmen zu Kapitel 4 „Kosten-Nutzen-Bewertung medizinischer Interventionen“

In mehreren Stellungnahmen gibt es Äußerungen zur Effizienzgrenze allgemein und Verweise auf frühere Stellungnahmen zu den Methoden und zur Kosten-Nutzen-Bewertung der Antidepressiva (G09-01).

Alle Argumente zum Konzept der Effizienzgrenze wurden in den Dokumentationen und Würdigungen zu den Stellungnahmen zu den Methoden von 2009 ausreichend erörtert [20].

In einigen Stellungnahmen wird gefordert, dass neben der Bewertung der wissenschaftlichen Evidenz zur Kosten-Nutzen-Bewertung (Assessment) mehrere Vorschläge zur Interpretation (Appraisal) abgegeben werden.

In einer Kosten-Nutzen-Bewertung nach § 35b SGB V werden Informationen zusammengestellt, die dem Entscheidungsträger helfen können, über Angemessenheit und Zumutbarkeit von Preisen zu urteilen. Eine abschließende Einschätzung, wie die Ergebnisse einer Kosten-Nutzen-Bewertung in eine Entscheidung münden können, ist dabei nicht Bestandteil des Auftrags. In Kosten-Nutzen-Bewertungen, die nach § 139a SGB V beauftragt werden, können je nach Auftrag andere Informationen präsentiert werden, wenn diese KNB nicht an ein Verfahren der frühen Nutzenbewertung nach AMNOG gekoppelt sind.

2.12 Würdigung der Stellungnahmen zu Abschnitt 4.1 „Einführung“

In einer Stellungnahme wird bemängelt, dass die Forderung, Abweichungen von den in Kapitel 4 vorgelegten Methoden zu begründen, zu einem unbegrenzten Spielraum für das Institut führe.

Die Methoden 4.2 gelten für alle Beteiligten gleichermaßen, also sowohl für das Institut bei der Erstellung einer Kosten-Nutzen-Bewertung als auch bei der Bewertung einer beispielsweise von einem pU erstellten und eingereichten Kosten-Nutzen-Bewertung. Somit eröffnet der vom Stellungnehmer zitierte Satz auch den pharmazeutischen Unternehmern die Möglichkeit, mit einer entsprechenden Begründung von den Methoden 4.2 im Bereich der Kosten-Nutzen-Bewertung abzuweichen. Dies zieht keine Änderung der Methoden nach sich.

2.13 Würdigung der Stellungnahmen zu Abschnitt 4.1.1 „Rechtliche Grundlagen für die Kosten-Nutzen-Bewertung nach SGB V“

In einer Stellungnahme wird ausgeführt, dass es keine Rechtsgrundlage für zwingend indikationsspezifische Wirtschaftlichkeitskriterien gäbe.

Der Stellungnehmer verweist richtigerweise darauf, dass das ausgewählte Zitat in Abschnitt 4.1.1 in der Stellungnahme des BMG aus dem Jahr 2008 im Abschnitt „Indikationsbezogene Nutzenbewertung“ auftaucht [6]. Um die Kontroverse um die Rechtsexegese aus den Jahren 2009 und 2010 nicht erneut zu entfachen, wurde der Abschnitt im Methodenpapier gestrichen.

Die Kosten-Nutzen-Bewertung nach § 35b SGB V soll Informationen dafür bereitstellen, dass anschließend Preise für ein zu bewertendes Arzneimittel ausgehandelt werden können. In diesem Verwertungszusammenhang geht es also zunächst um einen Vergleich in der Indikation. Dies wurde mit dem vom Stellungnehmer zitierten Satz ausgedrückt. Darüber hinaus kann eine KNB, wenn dies vom Entscheidungsträger gewünscht wird, auch zur Beantwortung einer indikationsübergreifenden Priorisierung im Gesundheitswesen herangezogen werden. Welches Maß des Gesamtnutzens dann allerdings einzusetzen wäre, um Ungerechtigkeiten gegenüber betroffenen Patientengruppen zu vermeiden, müsste Gegenstand von gesell-

schaftlichen Diskussionen und einer wissenschaftlich gestützten Entscheidungsfindung beim G-BA sein.

In einigen Stellungnahmen wird darüber hinaus eine indikationsübergreifende Bewertung in Verbindung mit indikationsübergreifenden Wirtschaftlichkeitskriterien gefordert. In einer Stellungnahme wird ergänzt, dass die indikationsübergreifende Bewertung der Allokation von Ressourcen in diejenigen Bereichen diene, in denen sie den größten gesamtgesellschaftlichen Nutzen entfalten.

Vorab gilt es festzustellen, dass weltweit alle Kosten-Nutzen-Bewertungen zunächst indikationsspezifisch durchgeführt werden. So gibt es keine Kosten-Nutzen-Bewertung, in der Ergebnisse eines Modells zur koronaren Herzkrankheit zugleich mit denen eines Modells zum Asthma bspw. abgetragen würden. Allgemeiner gesprochen: Es muss immer erst eine Kosten-Nutzen-Bewertung innerhalb einer Indikation oder gar Subindikation (2. oder 3. Therapielinie etc.) vorgenommen werden.

Anschließend kann indikationsübergreifend bewertet werden, wobei es einen veröffentlichten Schwellenwert nur in ganz wenigen Ländern gibt, z. B. in Großbritannien. Die Mehrheit der Länder hat keinen festen Schwellenwert.

Eine Zusammenfassung der Ergebnisse über Indikationsgebiete hinweg ist nach einer indikationsbasierten Kosten-Nutzen-Bewertung sicherlich möglich. Ob es aber das Ziel einer KNB im Rahmen einer solidarisch finanzierten Versorgung mit dem medizinisch Notwendigen sein kann, dass der größte gesamtgesellschaftliche Nutzen erzielt werden soll, müsste in einer gesellschaftlichen Debatte geklärt werden. Es ist durchaus denkbar im Sinne eines bedarfsorientierten oder an der Verwirklichungschance gemessenen Ansatzes, den auch z. B. der Nobelpreisträger Amartya Sen teilt [8,32], dass die Verteilung nach den Vorstellungen der Stellungnehmenden keine gesellschaftliche Mehrheit in Deutschland findet (vgl. auch Votum des dt. Ethikrats 2010 zur Kosten-Nutzen-Bewertung [10]).

Außerdem wären konkrete Vorschläge wünschenswert. Wenn es um das QALY als Maß des Gesamtnutzens geht, dann wird dies in Abschnitt 2.21 gewürdigt. Wenn es um einen festen Schwellenwert geht, an dem alle neuen Technologien gemessen werden, sollte es dazu eine gesellschaftliche Debatte geben. Zumindest handelt es sich hier nicht um eine methodische Frage der internationalen Standards der Gesundheitsökonomie, sondern um eine Frage der Wertentscheidung.

2.14 Würdigung der Stellungnahmen zu Abschnitt 4.1.2 „Perspektive“

In 2 Stellungnahmen wird eine Präzisierung des Unterschieds der GKV-Perspektive zur GKV-Versichertenperspektive gewünscht.

In Abschnitt 4.4.1 der Allgemeinen Methoden wird dies im Detail erläutert. Um eine Doppelung zu vermeiden, wurde dies in Abschnitt 4.1.2 der Allgemeinen Methoden nur durch eine kleine Ergänzung verdeutlicht. Dazu wurde ein Verweis gesetzt.

2.15 Würdigung der Stellungnahmen zu Abschnitt 4.1.3 „Zeithorizont“

In einer Stellungnahme wird eine Verdeutlichung gewünscht, was mit dem Begriff „Zeithorizont der Nutzenbestimmung“ gemeint ist.

Der Text wurde präzisiert.

2.16 Würdigung der Stellungnahmen zu Abschnitt 4.1.4 „Auswahl der Komparatoren“

In einer Stellungnahme wird die Änderung des Begriffs „Substanzklasse“ in „Wirkstoffklasse“ gefordert.

Der Begriff „Substanzklasse“ wurde in den Allgemeinen Methoden in „Wirkstoffklasse“ geändert.

In weiteren Stellungnahmen wird angemerkt, dass die Frage der fehlenden Heterogenität bzw. umgekehrt der Begriff der ausreichenden Homogenität geklärt werden müsse.

In Abschnitt 4.1.4 der Allgemeinen Methoden wird allgemein davon gesprochen, dass Wirkstoffe zu Klassen zusammengefasst werden können, wenn ausreichend Homogenität vorliegt. Hier wurde die weitere Bedingung eingefügt, „wenn dies medizinisch sinnvoll erscheint“. Schon in der Kosten-Nutzen-Bewertung der Antidepressiva wurde auch aufgenommen, dass Homogenität nicht in allen Endpunkten gegeben sein muss [17]. Wie dort auf Seite 188 für die SSRI ausgeführt, war wegen „Heterogenität für den Endpunkt „Therapieabbruch aufgrund unerwünschter Ereignisse“ [...] keine Betrachtung der Klasse möglich“. Ob eine ausreichende Homogenität vorliegt, wird mit den üblichen meta-analytischen Verfahren überprüft (siehe Abschnitt 8.3.8).

2.17 Würdigung der Stellungnahmen zu Abschnitt 4.1.5 „Versorgungspfad“

In zwei gleichlautenden Stellungnahmen wird eine Definition des Begriffs Versorgungspfad gefordert. Zugleich wird eine Abgrenzung zum Begriff Versorgungskontext gewünscht.

Der Begriff wird nun in Abschnitt 4.1.5 der Allgemeinen Methoden näher erläutert. Der Begriff Versorgungskontext umfasst allgemein die Versorgung für eine Indikation in Deutschland. Dies wurde im Methodenpapier umformuliert.

In einer Stellungnahme wird die Erklärung einer Piggyback-Studie gefordert.

Zur Verdeutlichung wurde die entsprechende Passage in den Allgemeinen Methoden umformuliert.

In einer Stellungnahme wird vorgeschlagen, zunächst zu beschreiben, was ein Modell ist, und dann auf den Begriff Versorgungspfad einzugehen.

Der Versorgungspfad wird zuerst dargestellt, da er als Grundlage für die Erstellung des Modelles dient. In Abschnitt 4.1.5 des Methodenpapiers wurde zum besseren Verständnis der

Begriff des entscheidungsanalytischen Modells eingeführt und entsprechend auf die Abschnitte 4.1.6 und 4.2 des Methodenpapiers verwiesen.

2.18 Würdigung der Stellungnahmen zu Abschnitt 4.1.7 „Datengrundlage“

In einigen Stellungnahmen wird die Gefahr gesehen, dass nationale Leitlinien überholte Empfehlungen enthalten könnten, Empfehlungen in ihrer Aussagekraft wegen fehlender Empfehlungseinstufung eingeschränkt sind oder sogar ganz fehlen.

Die jahrelange Erfahrung aus der Bewertung von Leitlinien zeigt, dass deutsche Leitlinien nicht generell veraltet sind und eine eingeschränkte Aussagekraft haben. Auch wenn Leitlinien aus anderen Ländern im Einzelfall aktueller sein sollten, bedeutet dies noch nicht, dass sie auch im deutschen Gesundheitssystem anwendbar sind. So können Aspekte z. B. aus dem US-amerikanischen Versorgungsgeschehen nicht einfach in das deutsche Versorgungsgeschehen übertragen werden, da ambulante und stationäre Behandlung ganz anders ausgestaltet sein können und es ggf. eine dort genannte Berufsgruppe in Deutschland gar nicht gibt.

Für die Entwicklung des Versorgungspfads an sich können daher nur deutsche Leitlinien genutzt werden. Allenfalls für Teilaspekte in der Versorgung können auch internationale Leitlinien herangezogen werden. Dafür gelten aber dieselben Bedingungen wie bei der Bewertung von Leitlinien, d. h., die Leitlinien müssen aus Ländern des WHO-Stratums 1 kommen.

In einer Stellungnahme wird gefragt, wie die Experten bei Befragungen, durchgeführt durch externe Ersteller einer Kosten-Nutzen-Bewertung oder durch das IQWiG, ausgewählt werden.

Zur Verdeutlichung wurde ein Hinweis in Abschnitt 4.1.7 auf die Methoden der quantitativen Sozialforschung eingefügt.

2.19 Würdigung der Stellungnahmen zu Abschnitt 4.1.9 „Interpretation der Ergebnisse“

In einer Stellungnahme wird vorgeschlagen, in Abschnitt 4.1.9 zunächst die Effizienzgrenze zu beschreiben und dann erst in die Interpretation einzusteigen.

In Abschnitt 4.1.9 der Allgemeinen Methoden und an anderen Stellen wurde nochmals am Aufbau des Kapitels 4 des Methodenpapiers gearbeitet. Dadurch wurden Passagen zur besseren Verständlichkeit umgestellt bzw. umgearbeitet. Die inhaltlichen Aussagen bleiben davon unberührt.

In einer Stellungnahme wird gefordert, für den Entscheidungsträger eine klare Vorgehensweise zu definieren, wie zu gewichten ist.

Eine Vorgabe für den Entscheidungsträger, wie zu gewichten ist, ist nicht vorgesehen.

In einigen Stellungnahmen wird darauf verwiesen, dass die Effizienzgrenze bzw. ihre Extrapolation nicht alleiniges Kriterium der Angemessenheit eines Preises für eine zu bewertende Intervention sein kann.

Selbst wenn man diese Einschätzung teilt, bleibt unbestreitbar, dass die Kosten-Nutzen-Bewertung mit dem Ansatz der Effizienzgrenze und die Ausgaben-Einfluss-Analyse wichtige Informationen zur Beurteilung der Angemessenheit und Zumutbarkeit einer Intervention liefern. Die Bewertung des IQWiG unterstützt den G-BA bei seiner Entscheidungsfindung. Der G-BA kann dabei ggf. weitere Kriterien zur Beurteilung der Angemessenheit eines Preises berücksichtigen.

2.20 Würdigung der Stellungnahmen zu Abschnitt 4.2 „Modellierung“

In einigen Stellungnahmen wird der Publikation der Modelle der Kosten-Nutzen-Bewertungen widersprochen.

Entsprechend der bisherigen Praxis wird die Veröffentlichung aller für die Bewertung einer Technologie relevanten Informationen als unabdingbar angesehen. Nicht veröffentlichte Informationen können daher für eine Bewertung auch nicht berücksichtigt werden.

In einer Stellungnahme wird eine Umformulierung gefordert, was Sensitivitätsanalysen leisten sollen.

Der betreffende Satz in Abschnitt 4.2.2 des Methodenpapiers wurde leicht umformuliert, um darauf hinzuweisen, dass insbesondere die Effekte von Annahmen und Modell-Inputs auf die Sicherheit des Ergebnisses durch Sensitivitätsanalysen überprüft werden.

In einer Stellungnahme wird geäußert, grundsätzlich akzeptierte Modellierungsumgebungen (z. B. Excel) anzugeben.

Vorzugsweise sollten die Modelle in den in der Gesundheitsökonomie gebräuchlichen Softwarepaketen programmiert und eingereicht werden, wobei man sich an anderen HTA-Einrichtungen wie NICE orientieren sollte [35].

In zwei identischen Stellungnahmen wird vorgeschlagen, das Einflussdiagramm durch ein Beispiel zu veranschaulichen.

Ein Beispiel liegt in Chapman und Sonnenberg [7] vor, das aus einer früheren Publikation von Owens [27] adaptiert wurde. Auch in der Kosten-Nutzen-Bewertung der Antidepressiva, Seite 83, wird ein Einflussdiagramm gezeigt [17]. In den Allgemeinen Methoden wird darauf verzichtet, lehrbuchartige Inhalte darzustellen.

2.21 Würdigung der Stellungnahmen zu Abschnitt 4.3 „Nutzen“

Aus einigen Stellungnahmen entsteht der Eindruck, dass Lebensqualität als übergeordneter Endpunkt angesehen werden muss.

In § 35b Absatz 1 Satz 4 SGB V wird die gesundheitsbezogene Lebensqualität als ein Endpunkt neben der Mortalität und der Morbidität genannt. Damit sind alle 3 Endpunkte gleichermaßen zu berücksichtigen, wenn es um ein Maß des Gesamtnutzens geht.

In einer Stellungnahme wird ganz grundsätzlich angezweifelt, dass der Nutzen in sehr vielen Krankheitsgebieten überhaupt kardinalskaliert gemessen werden kann.

Wenn dies grundsätzlich bezweifelt wird, wäre jede Form der Kosten-Nutzen-Bewertung nicht möglich. Kardinalität ist eine Grundannahme in der Nutzen- bzw. Kosten-Nutzen-Bewertung und insbesondere für die Bestimmung von Nutzwerten [15].

In einigen Stellungnahmen wird gefordert, dass Methoden bzw. Instrumente der Messung von Patientenpräferenzen, ausdrücklich AHP und CA genannt, auch in der Nutzenbewertung herangezogen werden sollten. Daher sollte die Darstellung der Methoden an anderer Stelle im Methodenpapier verortet werden.

Prinzipiell könnten Patientenpräferenzen auch bei der Bewertung der Relevanz der unterschiedlichen Nutzen- und Schadenaspekte herangezogen werden. Hierzu ist allerdings noch weitere Forschung nötig. Die Verortung der Darstellung der Erhebung von Patientenpräferenzen in Kapitel 4 ergibt mehr Sinn, da beide methodischen Ansätze im Rahmen des Maßes des Gesamtnutzens nach § 35b SGB V vorgestellt werden sollten. In der Version 4.1 des Methodenpapiers war die Kosten-Nutzen-Bewertung noch nicht integriert [19]. Durch die Integrierung der Kosten-Nutzen-Bewertung in das Methodenpapier 4.2 konnte die Darstellung der Patientenpräferenzen in das Kapitel 4 Kosten-Nutzen-Bewertung verschoben werden. Durch entsprechende Verweise von Kapitel 4 auf Kapitel 3 wird deutlich, dass die Methoden der Erhebung von Patientenpräferenzen auch in Nutzenbewertungen zum Einsatz kommen können. Eine Doppelung der Darstellung ist aber zu vermeiden. Auf Grundlage der Hinweise wurden im Methodenpapier die Verweise von Kapitel 4 auf Kapitel 3 nochmals geprüft und zum Teil verändert.

In einer Stellungnahme wird empfohlen, Fragen zur Messung der Lebensqualität, wie z. B. die zu befragende Population, die Verfahren und Möglichkeiten des Mappings, zu beantworten und konkrete Anweisungen zu geben.

Die Empfehlung wurde aufgegriffen und an den entsprechenden Stellen im Methodenpapier umgesetzt bzw. verdeutlicht. Es wird darauf hingewiesen, dass die Frage, bei wem die in das QALY einfließenden Werte erhoben worden sein sollten, nämlich bei den Betroffenen, an der entsprechenden Stelle in der zur Stellungnahme stehenden Fassung des Methodenpapiers 4.2 schon beantwortet wurde [20]. Präferiert werden insbesondere in der Nutzenbewertung die Erhebung der gesundheitsbezogenen Lebensqualität bei Betroffenen (Patient[inn]en), am besten in den RCTs, in denen auch andere patientenrelevante Endpunkte untersucht werden, bzw. die Nutzung von Tarifen, die unter Anwendung von generischen Instrumenten der Erhebung von gesundheitsbezogener Lebensqualität entwickelt wurden, die ebenfalls über Betroffene generiert wurden.

Eine Empfehlung zu den Verfahren abzugeben ist wissenschaftlich wenig sinnvoll, da – wie aktuelle wissenschaftliche Literatur zeigt [28,29] – alle Instrumente und Vorgehensweisen ihre Stärken und Schwächen haben, die sich zum Teil krankheitsspezifisch zeigen (Stichwort Reagibilität).

In diesem Zusammenhang wurde auch nach der Zulässigkeit des Mappings gefragt. Mit Mapping ist hier der Transfer von Ergebnissen, die mit einem Instrument gewonnen wurden, z. B. SF-36 oder SF-12, auf Nutzwerte eines anderen Instruments, z. B. EQ-5D, gemeint. Mapping ist als problematisch anzusehen. In Ausnahmefällen ist das Mapping möglich; allerdings sollte es dann validierte Mappingfunktionen geben (siehe britische Leitlinie [24], norwegische Leitlinie [25] sowie Table A20 der EUnetHTA Draft Guideline zu gesundheitsökonomischen Evaluationen [12]). Von einem Mapping von krankheitsspezifischen auf generische Instrumente ist abzuraten. In solchen Fällen sollten die krankheitsspezifischen Instrumente dann zur Darstellung der Lebensqualität herangezogen werden.

In einer Stellungnahme wird um Klärung einer Passage aus dem Buch von Drummond et al. [11] zu Präferenzen, Nutzwerten und Werten gebeten. Außerdem wird eine umfangreiche Methodendiskussion vorgeschlagen.

Im Methodenpapier werden zunächst die im Zusammenhang mit dem Maß des Gesamtnutzens wichtigsten Begriffe Präferenz (Preference), Nutzwert (Utility) und Wert (Value) genannt, ohne darauf einzugehen, ob es sich dabei um Synonyme handelt. Im Buch von Drummond et al. [11] heißt es auf Seite 142 gleich zu Beginn des Abschnitts „Utility, value, and preference“ dazu: „Preferences is the umbrella term that describes the overall concept; utilities and values are different types of preferences.“ Auf Deutsch heißt das also: Präferenzen sind der übergeordnete Begriff, der das umfassende Konzept beschreibt und mit Nutzwerten und Werten als unterschiedlichen Formen der Präferenzen gleichgesetzt werden kann. Auch in anderen Publikationen werden diese Begriffe als Synonyme gebraucht, z. B. Richardson et al. [28]: „If people have a preference (utility) for happiness [...].“ Andere wiederum trennen die Begriffe Präferenz und Nutzwert, der dann als die Repräsentation einer Präferenz(-struktur), als Nutzwertfunktion einer Präferenz aufgefasst wird. Die sich daraus ergebende Debatte zur Sprachregelung zu führen, kann nicht Gegenstand eines Methodenpapiers sein. Allerdings ist es wichtig, die grundlegenden Begriffe zu erwähnen. Gerade weil diese Begriffe nach methodischem Ansatz und nach Verortung in einer ökonomischen Theorie etwas anderes „bezeichnen“, wurde dies qualifizierend in Abschnitt 4.3.3 des Methodenpapiers angemerkt. Im Übrigen gilt auch an dieser Stelle, dass in den Allgemeinen Methoden darauf verzichtet wird, Lehrbuchwissen darzustellen.

In einigen Stellungnahmen wird bemängelt, dass keine Stellung zur Auswahl des jeweils am besten geeignet erscheinenden Verfahrens bei der Erhebung von Patientenpräferenzen genommen wird. Insgesamt wird zum Teil eine weitergehende Erläuterung gewünscht, wo das Institut die Verfahren AHP und CA, die Einsatzmöglichkeiten und die Vergleichbarkeit der Ergebnisse sieht.

Wie im Entwurf der Allgemeinen Methoden 4.2 vom 18.06.2014 auf Seite 89 angedeutet, gibt es idealtypische Situationen, in denen entweder nur ein AHP oder nur eine CA zur Anwendung kommen sollten. Im Gesundheitswesen geben einige Wissenschaftler/-innen dennoch eindeutig der einen oder der anderen Methode den Vorzug. Sie begründen dies mit theoretischen Überlegungen und verweisen z. B. darauf, dass die CA in der mikroökonomischen Theorie begründet sei.

Diese Begründungen reichen für das Für oder Wider einer der Methoden nicht aus. Zum einen gibt es viele Fälle, in denen eine theoretisch gut begründete Herleitung eines (medizinisch plausiblen) Zusammenhangs sich in der Praxis als nicht verifizierbar erwiesen hat [33]. Zum anderen muss die theoretische Herleitung mit dem Grundgedanken der Entscheidungsfindung auf Basis von HTA verträglich sein, was nicht für alle theoretischen Herleitungen zutreffen muss. Daher bleibt das IQWiG derzeit dabei, keiner Methode den Vorzug zu geben, und listet exemplarisch ein paar Kriterien, die bei der Wahl einer der beiden Methoden den Ausschlag geben können.

Zudem wurde in der Erörterung am 1.10.2014 hinlänglich deutlich, dass es zur Reliabilität, zur Validität und zur Repräsentativität noch ungeklärte Fragen gibt, siehe z. B. Lancsar und Swait [23]. Ob es überhaupt Sinn macht, diese Kriterien der Sozialempirie hier anzuwenden, müsste darüber hinaus diskutiert werden [23].

Zur Übereinstimmung von verschiedenen Methoden bei der Erhebung von Präferenzen liegen bisher nur vereinzelt Publikationen aus dem Bereich Gesundheit vor. Forschungen aus anderen Bereichen wie Verkehr, aber insbesondere aus dem Bereich der Konsumforschung zu übertragen, wird kontrovers diskutiert. Auch die Stellungnehmenden konnten ihre Einschätzungen nicht auf empirische Literatur stützen. Dieser Umstand verdeutlicht nochmals, dass ein Desiderat der Forschung besteht, z. B. zu klären, was die maximale Anzahl an Attributen in einem AHP und in einer CA ist, die Befragte bewältigen können.

Die hier aufgeworfenen Fragen und Probleme werden ebenfalls in der Dissertation von de Bekker-Grob [9] benannt. Z. B. äußert sich die Autorin dazu, dass die externe Validität, dass also Befragte sich tatsächlich nachher so entscheiden, wie sie sich in einer CA geäußert haben, bisher nicht untersucht wurde (Seiten 186 und 190f). Auch hat sie in ihren eigenen Untersuchungen beobachtet, dass die Antwortrate von 31 bis 33 % in der allgemeinen Bevölkerung darauf zurückzuführen sein könnte, dass eine wahlbasierte CA, also ein Discrete-Choice-Experiment (DCE), wegen der Komplexität (Seiten 187 und 190) zu einer Verzerrung führen kann, da Personen je nach Bildung und Sprachfähigkeit („language skills“) nicht partizipieren. Dies wird dann auch in die Empfehlungen für zukünftige Forschung (Seite 193) aufgenommen.

Neben der Zielsetzung könnte das Krankheitsbild die Wahl des Verfahrens bestimmen. Wie oben aufgeführt werden viele Aspekte der Verfahren noch weiterhin erforscht, sodass eine Priorisierung eines Verfahrens auf der jetzigen Informationsbasis nicht möglich erscheint.

Eine tiefer gehende methodische Diskussion nicht nur im Rahmen von Leitlinien oder Methodenhandbüchern wird als sinnvoll angesehen. In den Methoden bleibt es dabei, die Anforderungen an eine gute Praxis der Durchführung von Untersuchungen zu Patientenpräferenzen zu beschreiben.

In einigen Stellungnahmen wird eine Bewertung der Verfahren AHP und CA anhand der Kriterien Objektivität, Reliabilität, Validität und Repräsentativität gewünscht.

Die aktuelle Diskussion hierzu ist in dem gerade erschienenen Artikel von Lancsar und Swait [23] zusammengefasst. Dort finden sich neben den Desideraten der Forschung auch Hinweise auf die weitere Literatur zu diesen Fragestellungen.

Im Übrigen entsprechen die im Methodenpapier angeführten Gütekriterien, die die Qualität von Studien zur Erhebung von Patientenpräferenzen sicherstellen sollen und zu einer Minimierung des Einflusses von Verzerrung beitragen sollen, den Konsensuskriterien zu CA der entsprechenden Arbeitsgruppe der ISPOR [4]. In diese eher prozesshaften Kriterien gehen auch Aspekte der Objektivität und Reliabilität ein.

In einer Stellungnahme wird eine vertiefende Darstellung der frequentistischen und Bayes'schen Statistik vorgeschlagen.

An der entsprechenden Stelle in Abschnitt 4.3.4 der Allgemeinen Methoden wurde ein Verweis auf den Abschnitt 8.3.2 gesetzt, in dem auch auf Bayes'sche Methoden eingegangen wird.

2.22 Würdigung der Stellungnahmen zu Abschnitt 4.4 „Kosten“

In einer Stellungnahme wird darauf verwiesen, dass erstattungsfähige direkte nicht medizinische Kosten auch aus der Perspektive der Sozialversicherung eingeschlossen werden sollten.

Die entsprechenden Passagen wurden angepasst. Wichtig ist zu betonen, dass dies aber nicht dieselben direkten erstattungsfähigen nicht medizinischen Kosten sein müssen wie aus der Perspektive der GKV.

In einer Stellungnahme wird vorgeschlagen, dass der Ausschluss von PKV-Versicherten erwähnt werden soll.

Für die Bestimmung und Bewertung der Ressourcen aus der Perspektive der GKV oder der GKV-Versichertengemeinschaft sind die PKV-Versicherten unerheblich. Richtig wäre es, bei der Ressourcenbewertung aus Sicht der Gesellschaft zu berücksichtigen, dass die Leistungen der PKV-Vollversicherten teurer sind als die Leistungen in der Versorgung der GKV-Versicherten. Dazu gibt es entsprechende Darstellungen, z. B. Bock et al. [3].

In einer Stellungnahme wird die Verwendung des Humankapitalansatzes für die indirekten Kosten gefordert.

Die Anwendung des Friktionskostenansatzes entspricht dem Hannoveraner Konsens [14]. Daher ergibt sich keine Änderung.

In einer Stellungnahme wird darauf hingewiesen, dass der pharmazeutischen Industrie der Zugang zu GKV-Routinedaten nicht gegeben sei.

GKV-Routinedaten stellen nicht die einzige Quelle für die Ermittlung der Kosten dar. Andere Daten können von niedriger Aussagekraft sein, was im Einzelfall Auswirkungen auf die Belastbarkeit der Ergebnisse einer Kosten-Nutzen-Bewertung haben kann.

Im Übrigen sind nunmehr über das DIMDI Daten der GKV verfügbar (§§ 303a bis 303e SGB V). Zudem zeigt die Dossierbewertung, dass pharmazeutische Hersteller über Auftragsinstitute auf Auswertungen von GKV-Daten zurückgreifen können.

In einer Stellungnahme wird vorgeschlagen, die Berücksichtigung der Produktivitätsausfälle auf der Kostenseite sprachlich so anzupassen, dass auch eine Berücksichtigung auf der Nutzenseite in einer Kosten-Nutzen-Bewertung möglich ist.

Dem Vorschlag wurde gefolgt. An der entsprechenden Stelle wurde der Text geändert.

In einer Stellungnahme wird gefordert, die Lebensqualität bei Angehörigen nicht generell von der Berücksichtigung auszuschließen. Dabei wird auf Seite 93 in Abschnitt 4.4.2 verwiesen.

Dort heißt es im Kontext: „Die Lebensqualität von Angehörigen wird im Allgemeinen auf der Nutzenseite nicht berücksichtigt. Sollten deren Freizeitverluste betrachtet werden, sollten sie auch auf der Kostenseite bewertet werden [5,21,26,36].“ In diesem Passus wird die Berücksichtigung der Lebensqualität der Angehörigen nicht generell abgelehnt. Allerdings wird dies in Kosten-Nutzen-Bewertungen fast immer nicht gemacht. Stellvertretend können dann eben die Freizeitverluste auf der Kostenseite in einer KNB einbezogen werden.

In einer Stellungnahme wird vorgeschlagen, anstelle des günstigsten Vertreters eines Wirkstoffs bzw. einer Wirkstoffklasse einen Durchschnittspreis anzusetzen. Der Stellungnehmende begründet, dies würde nicht die tatsächlichen durchschnittlichen Preise reflektieren.

Folgt man der vorgetragenen Begründung, müsste man alle kassenspezifischen Rabatte einbeziehen, um die tatsächlichen durchschnittlichen Preise zu berücksichtigen. Dies ist praktisch nicht umsetzbar, würde aber die realen Kosten der GKV widerspiegeln. Diese realen Kosten der GKV würden in den meisten Fällen noch unter den Preisen der günstigsten Vertreter liegen. Daher soll weiterhin – auch aus Umsetzungsgründen – nach dem Wirtschaftlichkeitsprinzip der günstigste Vertreter eines Wirkstoffs- bzw. einer Wirkstoffklasse gewählt werden [22].

In einer Stellungnahme werden Hinweise darauf gewünscht, dass international in der gesellschaftlichen Perspektive zumeist keine andere Ressourcenbewertung als in der Perspektive des jeweiligen Gesundheitssystems vorgenommen wird.

Im EUnetHTA-Projekt [12] zur Zusammenstellung von Leitlinien konnte in keiner einzigen nationalen Leitlinie eine Ausführung dazu gefunden werden, dass Ressourcen aus gesellschaftlicher Sicht anders bewertet werden sollten. Wenn die gesellschaftliche Perspektive gewählt wird, könnte dies auch nur den Einschluss von Ressourcen für Produktivitätsausfälle bedeuten.

2.23 Würdigung der Stellungnahmen zu Abschnitt 4.5 „Epidemiologische Daten“

In einer Stellungnahme wird gewünscht, Näheres zur Ermittlung von Basiswahrscheinlichkeiten zu erläutern.

In Abschnitt 4.5.2 der Allgemeinen Methoden wird erläutert, woher die Daten kommen sollten und welche Anforderungen an die Qualität der Daten zu stellen ist, wenn sie als Basiswahrscheinlichkeiten in ein Modell einfließen.

In einer Stellungnahme wird gefordert, explizit Daten kommerzieller Anbieter zu erwähnen.

An der entsprechenden Stelle im Dokument „Allgemeine Methoden: Entwurf für Version 4.2“ vom 18.06.2014 und in Abschnitt 4.5.2 der Allgemeinen Methoden 4.2 wird keine abschließende Nennung von Quellen gemacht. Im Übrigen geht das Institut davon aus, dass die Daten kommerzieller Anbieter, wenn sie in Kosten-Nutzen-Bewertungen eingesetzt werden sollen, so aufbereitet werden, wie es die Kriterien einer wissenschaftlichen Publikation erfordern. Nur so können die Auswertungen nachvollziehbar und transparent gemacht werden.

2.24 Würdigung der Stellungnahmen zu Abschnitt 4.7 „Unsicherheit (Sensitivitätsanalysen)“

In mehreren Stellungnahmen wird kritisiert, der Net Health Benefit an sich bzw. in Verbindung mit dem Interquartilsabstand wirke glättend.

Der Net Health Benefit (NHB) und der Net Monetary Benefit (NMB) sind übliche Darstellungen der Unsicherheit. (17, 18). Beide Konzepte zielen darauf ab, einen Quotienten als Ergebnis einer Kosten-Nutzen-Bewertung zu vermeiden. Beim NMB werden QALYs als Ergebnis auf der Nutzenseite in monetäre Einheiten umgewandelt; die neue Technologie wird angenommen, wenn gilt: $(\lambda * \text{QALYs}) - \text{Kostendifferenz} > 0$, wobei λ dem Schwellenwert entspricht, der in der Darstellung der Unsicherheit von 0 bis unendlich verändert wird. Bei der alternativen Darstellung über den NHB gilt die folgende Gleichung: $\text{QALYs} - (\text{Kosten} / \lambda) > 0$, wobei sich in dem Quotienten $(\text{Kosten} / \lambda)$ die Geldeinheiten wegekürzen. Eine Technologie wird dann angenommen, wenn sie mehr als 0 QALYs stiftet. Beide Darstellungsformen können demnach nicht glättend wirken, da sie den ganzen Umfang der Unsicherheit darstellen. Wie im Methodenpapier dargestellt, werden die Verteilungen der Effekte und Kosten im Einzelfall geprüft und ihre Auswirkungen auf die Ergebnisse der Unsicherheits-

analysen abgeschätzt. Weitere Ergebnisdarstellungen sind nicht per se erforderlich (siehe Abschnitt 2.26).

Eine Interquartilsregion (IQR) stellt die Unsicherheit zwischen der 25. und 75. Perzentile einer Verteilung dar. Wie der Median ist sie unempfindlich gegen Ausreißer. Mit der IQR kann ein Korridor für einen zusatznutzenbasierten Erstattungspreis ausgewiesen werden. Darüber hinaus können auch andere Intervalle herangezogen werden, um das Ausmaß der Unsicherheit noch weiter zu bestimmen. Die IQR stellt jedoch einen Kompromiss dar zwischen einer umfassenden Darstellung der Unsicherheit und der Angabe eines vertretbaren Verhandlungsspielraums für einen Erstattungsbetrag.

Weiterhin wird in einer Stellungnahme gefordert, Unsicherheit in Form von Scatter Plots, Vertrauensintervallen, Konturplots, Kosten-Effektivitäts-Akzeptanz-Kurven etc. darzustellen.

Zum einen ist die Darstellung der Unsicherheit in vielfacher Variation nicht leistbar, weder für die Hersteller, von denen dies auch gefordert werden müsste, noch für das Institut. Zum anderen sind die vom Stellungnehmer genannten Verfahren ungeeignet, die Unsicherheit im Rahmen des Konzepts der Effizienzgrenze darzustellen [34]. Wie dort gezeigt wurde, bevorzugt das Institut daher weiterhin die Darstellung des NHB oder von sog. Preis-Kosten-Akzeptanz-Kurven. Weitere Verfahren sind auch für den Entscheidungsträger wenig hilfreich.

2.25 Würdigung der Stellungnahmen zu Abschnitt 4.8 „Ausgaben-Einfluss-Analyse (Budget-Impact-Analyse)“

In einer Stellungnahme wird angeregt, einen Satz zur Funktion der Ausgaben-Einfluss-Analyse als Kann-Regelung zu formulieren.

Der Satz wurde entsprechend angepasst.

2.26 Würdigung der Stellungnahmen zu Abschnitt 4.9 „Besonderheiten der Kosten-Nutzen-Bewertung nach § 35b SGB V“

In einer Stellungnahme wird vorgeschlagen, den Begriff zusatznutzenbereinigt durch zusatznutzenbasiert zu ersetzen.

Aus Sicht des Instituts passen beide Begriffe. Da der Begriff zusatznutzenbasiert möglicherweise besser verständlich ist, wurde der Begriff zusatznutzenbereinigt wie vorgeschlagen ersetzt.

In weiteren Stellungnahmen wird die Verwendung der Interquartilsregion (IQR) kritisiert.

Für Preisverhandlungen müssen alle Parteien ausreichend über die Unsicherheit einer KNB informiert werden, aber zugleich auch ein Korridor für einen zusatznutzenbasierten Erstattungspreis ausgewiesen werden. Dies leistet die Interquartilsregion. Darüber hinaus können auch andere Intervalle herangezogen werden, um das Ausmaß der Unsicherheit noch weiter zu bestimmen.

2.27 Würdigung der Stellungnahmen zu Kapitel 6 „Evidenzbasierte Gesundheitsinformation für Bürgerinnen und Bürger“

In einer Stellungnahme wird das Kriterium einer Prävalenz von 1 % für die Erstellung von Gesundheitsinformationen als fragwürdig eingeschätzt. Die rasch zunehmenden Erkenntnisse der biomedizinischen Forschung führten zu einer Differenzierung vormals homogen erscheinender Krankheitsbilder in klinisch und biologisch distinkte Subgruppen. Diese unterschieden sich zunehmend in Diagnose, Therapie und Management. Daher seien offenere Kriterien für die Themenwahl angemessen. Zudem seien die Änderungen im Stand des Wissens in der Hämatologie, Onkologie und der Hämostaseologie nicht linear, sondern wellenförmig, sodass variabelere Aktualisierungsintervalle als das beschriebene dreijährige Intervall nötig seien.

Kein Anbieter von Gesundheitsinformationen ist allein in der Lage, die Medizin in Breite und Tiefe abzudecken. Die Begrenzung der Themen ist immer eine Abwägung zwischen Wichtigkeit und den eigenen Möglichkeiten. Die Festlegung schafft eine transparente Definition, mit der das IQWiG seinem gesetzlichen Auftrag gerecht wird. Wie in Absatz 6.3.1 des vorgelegten Methodenentwurfs ausgeführt, wird zur Erstellung des Themenkatalogs primär auf den Versorgungsreport des WIdO der AOK zurückgegriffen. Nach Berechnungen des WIdO umfassen die auf Basis der 1-Prozent-Grenze ausgewählten Diagnosen (ICD10-Dreisteller) etwa 83 % der ambulant oder stationär abgerechneten Diagnosen der AOK-Patienten.

Das beschriebene Vorgehen schließt jedoch nicht aus, dass auch weitere Quellen bei der Auswahl der Themen berücksichtigt werden können. Ergänzende Quellen sind im gleichen Absatz aufgeführt. Es sei noch angemerkt, dass kein Informationsportal valide Gesundheitsinformationen zu sämtlichen Erkrankungen bereitstellen kann. Eine gewisse Beschränkung des einzelnen Anbieters und eine Vernetzung mit weiteren Erstellern evidenzbasierter Gesundheitsinformationen sind unumgänglich.

Wie in Absatz 6.3.4 des vorgelegten Methodenentwurfs ausgeführt, kann die Halbwertszeit der Evidenz themenspezifisch deutlich über, aber auch deutlich unter dem Zeitraum von 3 Jahren liegen. Das Aktualisierungsintervall von 3 Jahren beschreibt lediglich, wann der Aktualisierungsbedarf der Texte spätestens überprüft wird. Wie im gleichen Absatz beschrieben, werden im Rahmen eines regelmäßigen Evidenzscannings identifizierte Studien, systematische Übersichten oder Meldungen im Hinblick auf den Aktualisierungsbedarf der Gesundheitsinformationen bewertet. Diese Bewertung kann auch zu einer Aktualisierung vor Ablauf des 3-jährigen Intervalls führen. Nicht zuletzt können auch evidenzgestützte Hinweise von beispielsweise Fachgesellschaften auf potenziellen Aktualisierungsbedarf der Informationen eine Prüfung und gegebenenfalls frühere Aktualisierung auslösen.

Es besteht kein Änderungs- oder Ergänzungsbedarf.

In einer Stellungnahme wird gefragt, ob für die Gesundheitsinformationen auch ein Testlesen durch die Zielgruppen vorgesehen sei und ob auch geplant sei, Experten aus Fachgesellschaften beim Testlesen einzubeziehen.

Wie in Absatz 2.1.7 und 6.4.1 des vorgelegten Methodenentwurfs beschrieben, wird im Rahmen der Texterstellung die Gesundheitsinformation zunächst einer klinischen Expertin / einem klinischen Experten des jeweiligen Fachgebietes zur Begutachtung vorgelegt. Im anschließenden begrenzten Stellungnahmeverfahren organisiert die AWMF als Mitglied des Kuratoriums themenspezifisch die Weiterleitung des Textentwurfs der Gesundheitsinformation an die Fachgesellschaften. Diese haben dann die Möglichkeit, eine Stellungnahme einzureichen. Auch die dem Kuratorium angehörigen Patientenvertreter und der Patientenbeauftragte der Bundesregierung haben im Rahmen des Stellungnahmeverfahrens die Möglichkeit, Rückmeldungen zum Textentwurf zu geben. Parallel zum Stellungnahmeverfahren wird der Textentwurf durch externe Auftragnehmer in Form von Fokusgruppen von potenziellen Nutzerinnen und Nutzern nach festgelegten Kriterien bewertet. Überdies werden auch die Patientinnen und Patienten, die zur Erstellung der Erfahrungsberichte interviewt wurden, eingeladen, sich zu äußern.

Es besteht kein Änderungs- oder Ergänzungsbedarf.

In der gleichen Stellungnahme wird eine Veröffentlichung des Themenkatalogs und dessen Überprüfung durch die Fachgesellschaften angeregt. Auch hinsichtlich notwendiger Aktualisierungen sollten Vertreter der entsprechenden Fächer kontaktiert werden.

Im Methodenpapier ist in Abschnitt 6.1.3 dargestellt, wie sich die der Themenkatalog für die Gesundheitsinformationen des IQWiG ableitet. Die Grundlage, der Versorgungsreport des Wissenschaftlichen Instituts der AOK (WiDO), ist öffentlich zugänglich. Es steht den Fachgesellschaften darüber hinaus jederzeit frei, auf den möglichen Aktualisierungsbedarf der Gesundheitsinformationen auf Basis entsprechender Evidenz aufmerksam zu machen.

Der Abschnitt 6.1.3 der Allgemeinen Methoden wurde um einen Hinweis auf die Website des WiDO, auf der der Versorgungsreport zu finden ist, ergänzt.

In einer Stellungnahme wird die Änderung der Vorgehensweise zur Themenwahl kritisiert. Es wird vorgeschlagen, bei der Themenidentifizierung weiterhin auf öffentlich verfügbare Quellen zurückzugreifen und diese im Methodenpapier explizit zu benennen. Beispielhaft wird die „Gesundheitsberichterstattung Bund“ [sic] genannt.

Wie in Absatz 6.3.1 des vorgelegten Methodenentwurfs ausgeführt, wird zur Erstellung des Themenkatalogs primär auf den Versorgungsreport des WiDO der AOK zurückgegriffen. Dieser Report ist öffentlich verfügbar, sowohl in einer Printversion als auch im Internet. Der Versorgungsreport beinhaltet, basierend auf etwa 24 Millionen AOK-Versicherten, Angaben zu Prävalenzen und Hospitalisierungsraten für die 1500 häufigsten Erkrankungen. Dadurch werden, wie in Absatz 6.3.1 beschrieben, wichtige Aspekte epidemiologischer Bedeutung wie Häufigkeit der Erkrankung und Häufigkeit der Inanspruchnahme medizinischer Leistungen ausreichend berücksichtigt. Der Themenkatalog deckt auch die häufigen nicht onkologischen Todesursachen ab. Im Bereich Onkologie sind nur relativ wenige Krebsarten im Themenkatalog. Da es mit dem Krebsinformationsdienst aber ein spezialisiertes Angebot gibt, ist hier eine Arbeitsteilung und Kooperation für sinnvoll. Darüber hinaus werden im gleichen Absatz

ergänzende Quellen und Quellkategorien aufgeführt. Die hinter den Kategorien stehenden optionalen Quellen sämtlich explizit zu benennen würde den Rahmen des Methodenpapiers sprengen. Es besteht kein Änderungs- oder Ergänzungsbedarf.

2.28 Würdigung der Stellungnahmen zu Abschnitt 6.3.4 „Aktualisierung“

In einer Stellungnahme wird das reguläre dreijährige Aktualisierungsintervall für in aller Regel illusorisch gehalten und die Berücksichtigung der deutlich längeren Überarbeitungsintervalle der Verfahrensstandards der stellungnehmenden Fachgesellschaft angeregt.

Das reguläre 3-jährige Aktualisierungsintervall beschreibt, nach welchem Zeitraum der Aktualisierungsbedarf der Gesundheitsinformationen spätestens geprüft wird. Dieses Intervall ist zunächst einmal unabhängig von den Aktualisierungsintervallen der Leitlinien verschiedener Fachgesellschaften. Eine Synchronisierung ist weder sinnvoll noch möglich. Sollte aus der Aktualisierung einer Leitlinie ein Aktualisierungsbedarf für eine Gesundheitsinformation resultieren, kann diese auch außerhalb des 3-Jahres-Intervalls überarbeitet werden. Es besteht kein Änderungs- oder Ergänzungsbedarf.

2.29 Würdigung der Stellungnahmen zu Kapitel 7 „Informationsbeschaffung“

Eine Stellungnahme weist daraufhin, dass in der Auflistung der Quellen für die Suchen in Studienregistern das Studienregister Pharmnet.BUND fehlt. Laut Verfahrensordnung des G-BA zur frühen Nutzenbewertung von Arzneimitteln sei eine solche Suche Pflicht.

Der Hinweis ist richtig. Bei der Bewertung von Arzneimitteln ist Pharmnet.BUND als Quelle enthalten; diese fehlende Angabe wurde ergänzt.

Des Weiteren weist diese Stellungnahme daraufhin, dass für das Register Pharmnet.BUND eine Exportfunktion fehlt, und bittet darum, dies zu erörtern.

Hierzu ist festzustellen, dass solche Hinweise an das Deutsche Institut für Medizinische Dokumentation und Information (DIMDI) zu senden sind, da das DIMDI die IT-Plattform für dieses Register stellt.

In mehreren Stellungnahmen wurde kritisiert, dass unaufgefordert übermittelte Studiendaten inhaltlich nicht mehr berücksichtigt würden. Aus Sicht der Stellungnehmenden sei dieses Vorgehen nicht mit dem Grundprinzip der evidenzbasierten Medizin vereinbar, nach dem die gesamte Evidenz zu berücksichtigen sei, und könne dazu führen, dass relevante Informationen bzw. aktuelle Evidenz unberücksichtigt bleiben. Dies könne selbst zu einer Verzerrung des Berichtsergebnisses führen. Weiterhin sei es nicht mehr möglich, fehlerhafte oder unvollständige Darstellungen bzw. Veröffentlichungen des Instituts zu korrigieren. Ebenso könne es die Gefahr bergen, dass Informationen, die auf eine Patientengefährdung hinweisen, nicht berücksichtigt werden.

Zunächst soll noch einmal darauf hingewiesen werden, dass das in Kapitel 7 der Allgemeinen Methoden beschriebene Vorgehen allein Studiendaten betrifft, die außerhalb jeglicher

bestehender Regelungen (z. B. innerhalb eines Stellungnahmeverfahrens) unaufgefordert übermittelt werden. Diese Regelungen bleiben davon unberührt.

Aufgrund der bestehenden Regelungen bieten sich mehrere Möglichkeiten der Einreichung von Studiendaten. Im Rahmen von Stellungnahmeverfahren können immer Daten mit der Stellungnahme eingereicht werden. Das Stellungnahmeverfahren erfüllt darüber hinaus auch genau den Zweck, auf unvollständige bzw. fehlerhafte Darstellungen hinzuweisen. Mit Unterzeichnung der Vertraulichkeitsvereinbarung zur Übermittlung von Studiendaten im Rahmen von Herstelleranfragen besteht darüber hinaus immer die explizite Aufforderung zur Übermittlung von Studiendaten, z. B. sobald ein Studienbericht einer zunächst noch laufenden Studie fertiggestellt ist. Die Befürchtung eines Stellungnehmenden, dass ein Hersteller ggf. Jahre bis zum nächsten Stellungnahmeverfahren für einen Bericht warten muss, bis er verfügbare Studiendaten einreichen darf, ist vor diesem Hintergrund somit nicht nachvollziehbar. Insgesamt werden die bestehenden Regelungen zur Einreichung von Studiendaten als ausreichend angesehen, um eine Bewertung auf Basis der gesamten relevanten Evidenz durchzuführen. Insgesamt ergibt sich daraus kein Änderungsbedarf an den Allgemeinen Methoden.

In mehreren Stellungnahme wird die Nichtberücksichtigung von unaufgefordert übermittelten Daten insbesondere im Zusammenhang mit den Ausführungen in den Abschnitten 1.2.4 und 7.2 der Allgemeinen Methoden 4.2 kritisch gesehen, nach denen eine zentrale Strategie der evidenzbasierten Medizin darin besteht, alle nach der Qualität ihres Designs und ihrer Durchführung angemessenen Studien zu einer Frage zu identifizieren und so den Stand des zuverlässigen Wissens zusammenzufassen (Abschnitt 1.2.4), bzw. die Überprüfung der eingereichten Daten zur etablierten Arbeitsweise des Instituts gehöre (Abschnitt 7.2).

Wesentlich für eine Bewertung nach den Grundsätzen der evidenzbasierten Medizin ist, dass der Einschluss von Evidenz nicht selektiv geschieht, da es ansonsten zu einer Verzerrung der Ergebnisse kommen kann. Durch den Einschluss von unaufgefordert übermittelten Daten ist zudem nicht gewährleistet, dass die gesamte relevante Evidenz in die Bewertung einfließt. Die Gefahr des selektiven Einschlusses von Evidenz ist hingegen womöglich eher erhöht. Somit steht das Vorgehen nicht im Widerspruch zu den Ausführungen in Abschnitt 1.2.4 des Methodenpapiers. Hinsichtlich des Verweises der Stellungnehmenden auf den Abschnitt 7.2 sei darauf hingewiesen, dass dort nicht die Überprüfung von eingereichten Daten, sondern vielmehr die Überprüfung der Informationsbeschaffung im Rahmen von Potenzial- oder Dossierbewertungen beschrieben wird, die vom pharmazeutischen Unternehmer bzw. Antragsteller vorgelegt wurden. Bei diesen Bewertungen ist eine Übermittlung von Daten außerhalb des Dossiers bzw. des Antrags, aufgefordert oder unaufgefordert, aufgrund des Verfahrens grundsätzlich nicht vorgesehen. Es ergibt sich kein Änderungsbedarf an den allgemeinen Methoden.

2.30 Würdigung der Stellungnahmen zu Abschnitt 7.2 „Überprüfung der Informationsbeschaffung“

Für einen Stellungnehmenden ist es unklar, wie die formale Überprüfung der Literaturrecherche im Rahmen eines Antrags zur Erprobung von Untersuchungs- und Behandlungsmethode konkret erfolgt. Zudem weist der Stellungnehmende darauf hin, dass die Beurteilung der Validität des Antrags eine entscheidende Rolle spielt. Eine zentrale Funktion kommt dabei der Information zu, ob der für die jeweilige Fragestellung relevante Studienpool durch die Antragsunterlagen vollständig erfasst wird oder aber eine selektive Berichterstattung durch den Antragsteller erfolgt. Der Stellungnehmende wünscht sich daher eine konkrete Beschreibung der einzelnen Schritte, wie die Überprüfung der Informationsbeschaffung bei der Potenzialbewertung erfolgt, um den Unsicherheitsfaktor der Validität von Erprobungsanträgen besser einschätzen zu können.

Die Überprüfung der Informationsbeschaffung erfolgt wie beschrieben: Es erfolgt ein formaler Abgleich danach, ob alle formalen Inhalte der Antragsvorlage umgesetzt wurden. Zusätzlich wird – wie in Abschnitt 7.1.1. beschrieben – eine inhaltliche Qualitätsprüfung der Informationsbeschaffung unter Berücksichtigung der Übersichtsarbeit von Sampson [30,31] durchgeführt sowie falls möglich eine Qualitätssicherung anhand eines Testsets vorgenommen. Eine weitergehende Überprüfung der Informationsbeschaffung ist in Potenzialanträgen derzeit nicht vorgesehen. Jedoch kann im Rahmen eines durch den G-BA in Auftrag gegebenen Addendums eine systematische Recherche durchgeführt werden. Inhaltlich ergibt sich kein Änderungsbedarf.

2.31 Würdigung der Stellungnahmen zu Abschnitt 8.1.3 „Rangordnung verschiedener Studienarten/ Evidenzgrade“

In mehreren Stellungnahmen wird gefordert, indirekte Vergleiche und Netzwerk-Meta-Analysen in die allgemeine Evidenzhierarchie einzuordnen.

In Abschnitt 8.3.9 der Allgemeinen Methoden wird erläutert, dass indirekte Vergleiche und Netzwerk-Meta-Analysen in der Regel eine geringere Ergebnissicherheit aufweisen als adäquate direkt vergleichende Studien. Eine exakte Einordnung in eine starre Evidenzhierarchie ist nicht sinnvoll; daher besteht hier kein Änderungsbedarf.

2.32 Würdigung der Stellungnahmen zu Abschnitt 8.3.3 „Beurteilung klinischer Relevanz“

In mehreren Stellungnahmen wird kritisiert, dass aus den in Abschnitt 8.3.3 ergänzten Sätzen nicht hervorgeht, inwiefern bestimmte Instrumente als validiert oder etabliert einzuschätzen sind. Insbesondere der Begriff „etabliertes Instrument“ sei genauer zu definieren.

Die Einschätzung, ob ein in einer Studie eingesetztes Instrument validiert und infolgedessen als hinreichend valide anzusehen ist, wird bei komplexen Skalen regelhaft anhand der üblichen (Test-)Gütekriterien Validität, Reliabilität und Objektivität vorgenommen. Messinstrumente, die patientenrelevante Konstrukte abbilden und diese Anforderungen erfüllen,

werden in der Nutzenbewertung berücksichtigt. Über solche validierten bzw. validen Instrumente hinaus werden in Studien aber auch Messinstrumente auf Basis von (komplexen) Skalen eingesetzt, bei denen zwar ggf. nicht hinreichend untersucht ist, ob die Anforderungen an eine Validierung erfüllt sind, die sich aber aufgrund des bereits jahrelangen Einsatzes als Instrument bewährt haben. Ein Beispiel ist die Expanded Disability Status Scale (EDSS) im Bereich Multiple Sklerose. Die Einschätzung, ob ein bestimmtes Instrument als etabliert zu bezeichnen ist, ist jedoch stark von der jeweiligen Fragestellung und dem jeweiligen Instrument abhängig. Aus diesem Grund wird dies jeweils eine Einzelfallentscheidung notwendig machen. Eine Ergänzung der allgemeinen Methoden um eine genaue und trennscharfe Definition des Begriffes „etabliertes Instrument“ ist daher nicht möglich. Daraus ergibt sich insgesamt kein Änderungsbedarf an den allgemeinen Methoden.

In mehreren Stellungnahmen wird hinterfragt, warum die Verwendung von validierten bzw. etablierten Instrumenten für die Berücksichtigung von mittels (komplexen) Skalen erhobenen Endpunkten lediglich als „eine“ Voraussetzung (von mehreren) beschrieben ist und welches die übrigen Voraussetzungen seien.

Für Endpunkte, die mittels (komplexen) Skalen erhoben werden, gelten zunächst dieselben Voraussetzungen wie für alle übrigen in der Nutzenbewertung berücksichtigten Endpunkte. Dazu gehört beispielsweise, dass es sich um einen patientenrelevanten Endpunkt handelt bzw. dass ein patientenrelevantes Konstrukt gemessen wird. Für Endpunkte, die mit Skalen erhoben werden, kommt als zusätzliche Voraussetzung hinzu, dass entsprechend validierte bzw. etablierte Instrumente eingesetzt werden. „Eine Voraussetzung“ ist somit im Sinne von „eine im Vergleich zu anderen Endpunkten zusätzliche Voraussetzung“ zu verstehen. Somit ergibt sich kein Änderungsbedarf in den Allgemeinen Methoden.

In einer Stellungnahme wird für die Verwendung von (komplexen) Skalen neben der Validität als zusätzliche Voraussetzung gefordert, dass es sich um kultursensible Instrumente handelt.

Ob es sich bei einem Instrument um ein kultursensibles Instrument handelt, ist Teil der Prüfung der Validität dieses Instrumentes. Liegen aus dieser Prüfung für ein Instrument Hinweise vor, dass kulturbedingte Unterschiede in der Validität vorliegen bzw. dass es für bestimmte Patientengruppen nicht validiert ist, kann dies dazu führen, dass das entsprechende Instrument in der Nutzenbewertung nicht berücksichtigt wird. Es ergibt sich kein Änderungsbedarf in den Allgemeinen Methoden.

3 Literaturverzeichnis

1. Albrecht J, Meves A, Bigby M. Case reports and case series from Lancet had significant impact on medical literature. *J Clin Epidemiol* 2005; 58(12): 1227-1232.
2. Asadi-Lari M, Tamburini M, Gray D. Patients' needs, satisfaction, and health related quality of life: towards a comprehensive model. *Health Qual Life Outcomes* 2004; 2: 32.
3. Bock JO, Brettschneider C, Seidl H, Bowles D, Holle R, Greiner W et al. Ermittlung standardisierter Bewertungssätze aus gesellschaftlicher Perspektive für die gesundheitsökonomische Evaluation. *Gesundheitswesen* 2015; 77(1): 53-61.
4. Bridges JF, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA et al. Conjoint analysis applications in health: a checklist; a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Health* 2011; 14(4): 403-413.
5. Brouwer W, Rutten F, Koopmanschap M. Costing in economic evaluations. In: Drummond M, McGuire A (Ed). *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001. S. 68-93.
6. Bundesministerium für Gesundheit. Stellungnahme zur Methodik der Kosten-Nutzen-Bewertung von Arzneimitteln [online]. 06.08.2008 [Zugriff: 14.11.2014]. URL: <http://www.pa-gesundheit.de/pdf/KNB/03/8.3-BMG-Stellungnahme-KNB-Methodik.pdf>.
7. Chapman GB, Sonnenberg FA. *Decision making in health care: theory, psychology, and applications*. Cambridge: Cambridge University Press; 2000.
8. Clark DA. The capability approach: its development, critiques and recent advances [online]. 11.2005 [Zugriff: 14.11.2014]. (GPRG Working Paper Series; Band 32). URL: <http://www.gprg.org/pubs/workingpapers/pdfs/gprg-wps-032.pdf>.
9. De Bekker-Grob EW. *Discret choice experiments in health care: theory and applications*. Enschede: Ipskamp Drukkers; 2009. URL: http://repub.eur.nl/pub/21908/091125_Bekker-Grob,%20Esther%20Wilhelmina%20de.pdf.
10. Deutscher Ethikrat (Ed). *Nutzen und Kosten im Gesundheitswesen: Zur normativen Funktion ihrer Bewertung - Stellungnahme*. Berlin: Deutscher Ethikrat; 2011. URL: <http://www.ethikrat.org/dateien/pdf/stellungnahme-nutzen-und-kosten-im-gesundheitswesen.pdf>.
11. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2005.
12. European network for Health Technology Assessment. *Methods for health economic evaluations: a guideline based on current practices in Europe; draft guideline* [online]. 29.10.2014 [Zugriff: 16.01.2015]. URL: http://www.eunetha.eu/sites/5026.fedimbo.belgium.be/files/news-attachments/methodological_guideline_methods_for_health_economic_evaluations_20141101.pdf.

13. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. 08.05.2014 [Zugriff: 22.05.2014]. URL: https://www.g-ba.de/downloads/62-492-873/VerfO_2014-03-20.pdf.
14. Graf von der Schulenburg JM, Greiner W, Jost F, Klusen N, Kubin M, Leidl R et al. Deutsche Empfehlungen zur gesundheitsökonomischen Evaluation: dritte und aktualisierte Fassung des Hannoveraner Konsens. Gesundheitsökonomie & Qualitätsmanagement 2007; 12(5): 285-290.
15. Greiner W. Lebensqualitätsbewertung und Utilities in der Gesundheitsökonomie [online]. In: IQWiG Symposium 2013; 29.11.2013; Köln, Deutschland. [Zugriff: 14.11.2014]. URL: <https://www.iqwig.de/de/veranstaltungen/herbst-symposium/symposium-2013/2013-11-29-prof-dr-rer-pol-wolfgang-greiner.5354.html>.
16. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Biomarkerbasierte Tests zur Entscheidung für oder gegen eine adjuvante systemische Chemotherapie beim primären Mammakarzinom: vorläufiger Berichtsplan; Auftrag D14-01 [online]. 05.09.2014 [Zugriff: 24.11.2014]. URL: https://www.iqwig.de/download/D14-01_Vorlaeufiger-Berichtsplan_Biomarker-bei-Mammakarzinom.pdf.
17. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medikamentösen Behandlungen: Dokumentation und Würdigung der Anhörung zum Vorbericht; Auftrag G09-01 [online]. 03.09.2013 [Zugriff: 06.11.2013]. URL: https://www.iqwig.de/download/G09-01_DWAV_Kosten-Nutzen-Bewertung-von-Venlafaxin-Duloxetin-Bupropion-u....pdf.
18. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Spermogrammparameter für eine Indikation zur Intracytoplasmatischen Spermieninjektion (ICSI) statt In-vitro-Fertilisation: Abschlussbericht; Auftrag N12-02 [online]. 10.09.2014 [Zugriff: 24.11.2014]. (IQWiG-Berichte; Band 242). URL: https://www.iqwig.de/download/N12-02_Abschlussbericht_Spermogrammparameter-fuer-eine-Indikation-zur-ICSI-statt-IVF.pdf.
19. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.1. Köln: IQWiG; 2013. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-1.pdf.
20. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Dokumentation und Würdigung der Stellungnahmen zur „Aktualisierung einiger Abschnitte der Allgemeinen Methoden Version 4.0 sowie neue Abschnitte zur Erstellung der Allgemeinen Methoden Version 4.1“ [online]. 28.11.2013 [Zugriff: 14.11.2014]. URL: https://www.iqwig.de/download/Dokumentation_und_Wuerdigung_der_Stellungnahmen_IQWiG_Methoden_4-1.pdf.
21. Johannesson M, Karlsson G. The friction cost method: a comment. J Health Econ 1997; 16(2): 249-255.

22. Krauth C, Hessel F, Hansmeier T, Wasem J, Seitz R, Schweikert B. Empirische Bewertungssätze in der gesundheitsökonomischen Evaluation: ein Vorschlag der AG Methoden der gesundheitsökonomischen Evaluation (AG MEG). *Gesundheitswesen* 2005; 67(10): 736-746.
23. Lancsar E, Swait J. Reconceptualising the external validity of discrete choice experiments. *Pharmacoeconomics* 2014; 32(10): 951-965.
24. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London: NICE; 2004. URL: http://www.nice.org.uk/niceMedia/pdf/TAP_Methods.pdf.
25. Norwegian Medicines Agency. Guidelines on how to conduct pharmacoeconomic analyses [online]. 01.03.2012 [Zugriff: 14.11.2014]. URL: http://www.legemiddelverket.no/English/price_and_reimbursement/application_for_reimbursement/Lists/PageAttachments/default/NO/Pharmacoeconomic%20guidelines%20-%20Norway.pdf.
26. Oostenbrink JB, Koopmanschap MA, Rutten FF. Standardisation of costs: the Dutch Manual for Costing in economic evaluations. *Pharmacoeconomics* 2002; 20(7): 443-454.
27. Owens DK, Shachter RD, Nease RF Jr. Representation and analysis of medical decision problems with influence diagrams. *Med Decis Making* 1997; 17(3): 241-262.
28. Richardson J, Khan M, Iezzi A, Maxwell A. Subjective wellbeing, utility and quality of life: results from the multi instrument comparison project [online]. 05.2013 [Zugriff: 14.11.2014]. URL: http://www.aqol.com.au/documents/MIC/Subjective_Wellbeing_Brochure_V8.pdf.
29. Richardson J, Iezzi A, Maxwell A. Cross-national comparison of twelve quality of life instruments: MIC paper 1; background, questions, instruments [online]. 09.2012 [Zugriff: 24.11.2014]. (CHE Research Papers; Band 76). URL: <http://www.buseco.monash.edu.au/centres/che/pubs/researchpaper76.pdf>.
30. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. *J Clin Epidemiol* 2009; 62(9): 944-952.
31. Sampson M, McGowan J, Lefebvre C, Moher D, Grimshaw J. PRESS: Peer Review Of Electronic Search Strategies. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2008. URL: http://www.cadth.ca/media/pdf/477_PRESS-Peer-Review-Electronic-Search-Strategies_tr_e.pdf.
32. Sen A. Ökonomie für den Menschen: Wege zu Gerechtigkeit und Solidarität in der Marktwirtschaft. München: Deutscher Taschenbuch Verlag; 2002.
33. Steinbeck G, Paumgartner G (Ed). Therapie innerer Krankheiten. Heidelberg: Springer; 2005.

34. Stollenwerk B, Lhachimi SK, Briggs A, Fenwick E, Caro JJ, Siebert U et al. Communicating the parameter uncertainty in the IQWiG efficiency frontier to decision-makers. Health Econ 2014.
35. Tosh J, Wailoo A. Review of software for decision modelling [online]. 29.07.2008 [Zugriff: 24.11.2014]. URL: <http://www.nicedsu.org.uk/PDFs%20of%20reports/softwarereport-final.pdf>.
36. Weinstein MC, Siegel JE, Garber AM, Lipscomb J, Luce BR, Manning WG et al. Productivity costs, time costs and health-related quality of life: a response to the Erasmus Group. Health Econ 1997; 6(5): 505-510.

4 Offenlegung potenzieller Interessenkonflikte

Im Folgenden sind die potenziellen Interessenkonflikte der Stellungnehmenden sowie weiterer Teilnehmer an der wissenschaftlichen Erörterung (wissenschaftlicher Beirat) zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangabe der einzelnen Personen anhand der Formblätter zur Offenlegung potenzieller Interessenkonflikte (für Stellungnehmende und für Mitglieder des wissenschaftlichen Beirats). Das Formblatt für Stellungnehmende ist unter www.iqwig.de abrufbar. Die in den Formblättern aufgeführten Fragen finden sich im Anschluss an die jeweiligen Zusammenfassungen.

4.1 Potenzielle Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen, Firmen und Privatpersonen

Organisation/ Institution/ Privatperson	Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Allianz Chronischer Seltener Erkrankungen	Andreas L. G. Reimann	ja	ja	nein	ja	ja	ja
Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften	Monika Nothacker	ja	ja	ja	ja	nein	nein
AstraZeneca GmbH	Chantal Merens	ja	nein	nein	nein	nein	nein
	Stefanie Schattling	ja	ja	nein	nein	nein	nein
Boehringer Ingelheim Pharma GmbH & Co. KG	Matthias Pfannkuche	ja	ja	ja	nein	nein	nein
	Justine Rochon	ja	ja	ja	nein	nein	nein
Bundesarbeitsgemeinschaft Selbsthilfe von Menschen mit Behinderung und chronischer Erkrankung und ihren Angehörigen e. V.	Martin Danner	ja	nein	nein	nein	ja	nein
Bundesverband Medizintechnologie e. V.	Gabriela Soskuty	ja	n.a.	n.a.	n.a.	n.a.	n.a.
	Olaf Winkler	ja	n.a.	n.a.	n.a.	n.a.	n.a.
Deutsche Diabetes Gesellschaft und Deutsche Gesellschaft für Kardiologie	Barbara Bitzer	ja	nein	nein	nein	nein	nein
	Stephan Baldus	nein	ja	ja	ja	n.a.	nein
Deutsche Gesellschaft für Hämatologie und medizinische Onkologie	Bernhard Wörmann	nein	nein	nein	nein	nein	nein
Deutsche Gesellschaft für Hygiene und Mikrobiologie	Marianne Abele-Horn	nein	nein	nein	nein	nein	nein

Organisation/ Institution/ Privatperson	Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Deutsche Gesellschaft für Innere Medizin e. V.	Tilman Sauerbruch	nein	ja	ja	nein	nein	nein
Deutsche Gesellschaft für Orthopädie und Unfallchirurgie e. V.	Bertil Bouillon	ja	nein	nein	nein	nein	nein
	Dirk Stengel	ja	ja	ja	ja	nein	nein
Deutsche Krankenhausgesellschaft e. V.	Angela Neumeyer-Gromen	nein	nein	nein	nein	nein	nein
	Christian Ziegler	nein	nein	nein	nein	nein	nein
Deutscher Psoriasis Bund e. V.	Detlev Kunz	nein	ja	ja	ja	ja	nein
Ecker + Ecker GmbH	Birte Muhsal	ja	ja	ja	nein	nein	nein
	Claudia Pütz	ja	ja	ja	nein	nein	nein
GlaxoSmithKline GmbH & Co. KG	Michael Ingenhaag	ja	nein	nein	nein	nein	nein
	Isabelle Schatz	ja	nein	nein	nein	nein	nein
Institut für Gesundheitsökonomie	Aljoscha S. Neubauer	nein	ja	ja	ja	nein	ja
MSD Sharp & Dohme GmbH	Karl J. Krobot	ja	nein	ja	nein	nein	ja
	Stefan Schulz	ja	nein	nein	nein	nein	ja
Novartis Pharma GmbH	Sven Klebs	ja	nein	nein	nein	nein	nein
	Volker Claus	ja	nein	nein	nein	nein	ja
Sanofi-Aventis Deutschland GmbH	Stefan Schinzel	ja	nein	nein	nein	nein	nein
	Franz-Werner Dippel	ja	nein	ja	nein	ja	ja
Verband Forschender Arzneimittelhersteller e. V.	Ch.-Markos Dintsios	ja	ja	nein	nein	nein	nein
	Andrej Rasch	ja	nein	nein	nein	nein	nein
YES Pharmaceutical Development Services GmbH	Johannes Hofmann	ja	ja	ja	nein	nein	nein
Privatperson	Axel Mühlbacher	ja	ja	ja	ja	nein	ja

4.2 Potenzielle Interessenkonflikte von weiteren Teilnehmern der wissenschaftlichen Erörterung (wissenschaftlicher Beirat)

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6
Köbberling, Johannes	nein	nein	nein	nein	nein	nein
Koch, Armin	nein	nein	nein	ja	nein	nein
Meyer, Gabriele	nein	nein	nein	nein	nein	nein

Im Formblatt zur Offenlegung potenzieller Interessenkonflikte (für Stellungnehmende) wurden folgende 6 Fragen gestellt (Version 11/2013):

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband direkt oder indirekt beraten?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Frage 4: Haben Sie und / oder hat die Einrichtung¹, für die Sie tätig sind, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischem Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und / oder hat die Einrichtung¹, für die Sie tätig sind, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?

¹ Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.

Im Formblatt zur Offenlegung potenzieller Interessenkonflikte für Berater des Instituts wurden folgende 6 Fragen gestellt (Version 01/2012):

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor angestellt bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband, direkt oder indirekt beraten?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, Honorare erhalten für Vorträge, Stellungnahmen oder Artikel?

Frage 4: Haben Sie und / oder hat die Einrichtung², die Sie vertreten, abseits einer Anstellung oder Beratungstätigkeit innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband, finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und / oder hat die Einrichtung², bei der Sie angestellt sind bzw. die Sie vertreten, innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren ohne wissenschaftliche Gegenleistung) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband?

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist?

² Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.

5 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll

5.1 Teilnehmerliste der wissenschaftlichen Erörterung

Name	Organisation / Institution / Firma / privat
Abele-Horn, Marianne	Deutsche Gesellschaft für Hygiene und Mikrobiologie e. V. (DGHM)
Baldus, Stephan	Deutsche Gesellschaft für Kardiologie e. V. (DGK)
Bender, Ralf	IQWiG
Bitzer, Barbara	Deutsche Diabetes Gesellschaft e. V. (DDG)
Bouillon, Bertil	Deutsche Gesellschaft für Orthopädie und Unfallchirurgie e. V. (DGOU)
Breuer, Susanne	IQWiG
Claus, Volker	Novartis Pharma GmbH
Danner, Martin	BAG Selbsthilfe e. V.
Dintsios, Ch.-Markos	Verband forschender Arzneimittelhersteller e. V. (vfa)
Dippel, Franz-Werner	Sanofi-Aventis Deutschland GmbH
Ernst, Stefan	Protokollant
Gerber-Grote, Andreas	IQWiG
Großelfinger, Robert	IQWiG
Hofmann, Johannes	YES Pharmaceutical Development Services GmbH
Ingenhaag, Michael	GlaxoSmithKline GmbH & Co. KG (GSK)
Klebs, Sven	Novartis Pharma GmbH
Köbberling, Johannes	Wissenschaftlicher Beirat
Koch, Armin	Wissenschaftlicher Beirat
Krobot, Karl J.	MSD Sharp & Dohme GmbH
Kunz, Hans-Detlev	Deutscher Psoriasis Bund e. V. (DPB)
Lange, Stefan	IQWiG
Merens, Chantal	AstraZeneca GmbH
Meyer, Gabriele	Wissenschaftlicher Beirat
Mostardt, Sarah	IQWiG
Mühlbacher, Axel	Privatperson
Muhsal, Birte	Ecker + Ecker GmbH
Neubauer, Aljoscha S.	Institut für Gesundheitsökonomik (IfG)
Neumeyer-Gromen, Angela	Deutsche Krankenhausgesellschaft e. V. (DKG)
Nothacker, Monika	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF)
Pfannkuche, Matthias	Boehringer Ingelheim Pharma GmbH & Co. KG
Pütz, Claudia	Ecker + Ecker GmbH
Rasch, Andrej	Verband forschender Arzneimittelhersteller e. V. (vfa)
Reimann, Andreas	Allianz Chronischer Seltener Erkrankungen e. V. (ACHSE)
Rochon, Justine	Boehringer Ingelheim Pharma GmbH & Co. KG
Sauerbruch, Tilmann	Deutsche Gesellschaft für Innere Medizin e. V. (DGIM)
Sauerland, Stefan	IQWiG

Name	Organisation / Institution / Firma / privat
Schattling, Stefanie	AstraZeneca GmbH
Schatz, Isabelle	GlaxoSmithKline GmbH & Co. KG (GSK)
Schinzel, Stefan	Sanofi-Aventis Deutschland GmbH
Schulz, Stefan	MSD Sharp & Dohme GmbH
Skipka, Guido	IQWiG
Soskuty, Gabriela	Bundesverband Medizintechnologie e. V. (BVMed)
Stengel, Dirk	Deutsche Gesellschaft für Orthopädie und Unfallchirurgie e. V. (DGOU)
Vervölgyi, Volker	IQWiG
Waltering, Andreas	IQWiG
Wieseler, Beate	IQWiG
Windeler, Jürgen	IQWiG
Winkler, Oliver	Bundesverband Medizintechnologie e. V. (BVMed)
Wörmann, Bernhard	Deutsche Gesellschaft für Hämatologie und medizinische Onkologie e. V. (DGHO)
Ziegler, Christian	Deutsche Krankenhausgesellschaft e. V. (DKG)

5.2 Liste der Stellungnahmen, zu denen kein Vertreter an der wissenschaftlichen Erörterung teilgenommen hat

In der folgenden Tabelle werden Stellungnahmen genannt, zu denen trotz Einladung kein Stellungnehmender oder Vertreter zur wissenschaftlichen Erörterung erschienen ist.

Organisation / Institution / Firma / Privatperson
Bundesverband der Pharmazeutischen Industrie e. V. (BPI)
Deutsche Gesellschaft für Gesundheitsökonomie (dggö)
Deutsche Gesellschaft für Psychoanalyse, Psychotherapie, Psychosomatik und Tiefenpsychologie (DGPT)
Deutsches Netzwerk Versorgungsforschung e. V. (DNVG)
Prof. Dr. Thomas Kohlmann
MARs Market Access & Pricing Strategy GmbH und AHEAD GmbH
Prof. Werner Vach

5.3 Tagesordnung der wissenschaftlichen Erörterung

	Begrüßung und Einleitung
TOP 1	Endpunkt Patientenzufriedenheit
TOP 2	Einbezug von Patientenpräferenzen in (K)NB
TOP 3	Direkter Nutzen diagnostischer Verfahren
TOP 4	Charakterisierung der Evidenz in Potenzialbewertungen
TOP 5	Darstellung von Unsicherheit in KNB
TOP 6	Verschiedenes

5.4 Protokoll der wissenschaftlichen Erörterung

Datum: 01.10.2014, 11:00 bis 15:00 Uhr

Ort: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG),
Im Mediapark 8, 50670 Köln

Moderation: Jürgen Windeler

Moderator Jürgen Windeler: Ich begrüße Sie sehr herzlich zur wissenschaftlichen Erörterung des Methodenpapiers, Version 4.2.

Ich hoffe, dass Sie alle unproblematisch hierher und insbesondere in dieses Gebäude und diesen Raum gefunden haben. Manchmal gibt es das eine oder andere Problem mit der Identifizierung der Hausnummern, haben wir festgestellt. Ab und zu, wenn wir von draußen vom Mittagessen zurückkommen, stoßen wir auf etwas herumirrende Menschen, die die Hausnummer 8 oder 4 oder Ähnliches suchen. Ich hoffe, dass es Ihnen nicht so gegangen ist bzw. dass Sie nicht so lange gesucht haben. Sie scheinen es, jedenfalls die meisten von Ihnen, pünktlich geschafft zu haben.

Ich will die üblichen Vorbemerkungen machen, die die eine oder der andere von Ihnen, die an solchen Erörterungen teilgenommen haben ...

(Störungen in der Mikrofonanlage durch Handygeräusche)

- Ich habe keines mit.

Ich will die üblichen Einführungen machen. Die wichtigste ist, dass Sie bei der Einladung darauf hingewiesen worden sind, dass die Erörterung aufgezeichnet wird. Sie wird sowohl stenografisch protokolliert, als auch auf Band aufgezeichnet. Wir werden das Wortprotokoll der Erörterung veröffentlichen. Dadurch, dass Sie hierhergekommen sind, haben Sie implizit der Veröffentlichung und damit auch der Veröffentlichung Ihrer Beiträge zugestimmt. Falls Sie sich das jetzt nach meiner Darlegung anders überlegen sollten, müsste ich Sie bitten, von

der Teilnahme an dieser Erörterung Abstand zu nehmen. Aber ich denke, dass sich alle darüber im Klaren gewesen sind, als sie hierhergekommen sind.

Damit zusammen hängt auch, dass Sie bitte bei jeder Wortmeldung - nicht nur bei der ersten, wie man das sonst manchmal bei Veranstaltungen macht - Ihren Namen sagen, damit für den Stenografen und auf dem Band die Äußerungen zuverlässig zugeordnet werden können.

Die Erörterung - das kennen diejenigen, die Erörterungen mitgemacht haben - dient ganz vorrangig der Klärung von offenen Fragen, also von Fragen, die uns aus den Stellungnahmen offen geblieben sind. Es geht nicht darum, wie man das aus anderen Erörterungen oder Anhörungen - das ist keine Anhörung - kennt, dass man hier verliest, was man vorher geschrieben hat. Sie können davon ausgehen, dass wir das alles gelesen haben und dass wir auch die Dinge, die wir heute von uns aus hier nicht zur Sprache bringen, glauben, verstanden zu haben. Die Dinge, die wir hier zur Sprache bringen, haben noch Diskussionsbedarf.

Unter TOP 6 besteht noch die Möglichkeit, sich zu anderen Dingen zu äußern. Aber es geht nicht darum, EbM als solche, das Methodenpapier als Ganzes oder Ähnliches zu diskutieren, sondern konkrete Punkte.

Ich weise auch hier noch einmal darauf hin - wir haben das auch bereits im schriftlichen Stellungnahmeverfahren gemacht -, dass wir Änderungen zur Stellungnahme zur Diskussion stellen. Die Erörterung wird sich auf diese geänderten Passagen beziehen. Zu Diskussionen, die man zu anderen Teilen sicherlich immer wieder führen kann und die auch schon wiederholt geführt worden sind, finden es wir an dieser Stelle nicht angebracht, noch einmal zu diskutieren. Wenn wir - die Version 4.2 wird nicht die letzte sein, wie Sie sich vorstellen können - in der nächsten Änderung andere Passagen anpacken, ändern und aktualisieren, wird sich Gelegenheit ergeben, diese zu diskutieren.

Wir haben die Tagesordnung verschickt mit den Punkten, die angesprochen werden sollen. Wir werden, je nach Ablauf der Diskussion und je nach Zeit, die wir für die einzelnen Punkte brauchen, gegen 13 Uhr, eine kleine Pause machen, damit wir den Rest durchhalten. Ich denke nach allem, was ich mir im Moment vorstellen kann, dass wir vielleicht die Zeit bis 16 Uhr nicht ganz ausschöpfen werden. So viel zunächst zum Ablauf. Gibt es Fragen dazu? - Die sehe ich nicht.

Mein einziger weiterer Hinweis: Sie, besonders diejenigen, die auf der Mauerseite sitzen, müssen sich bitte sehr deutlich zu Wort melden. Die anderen kann ich relativ gut sehen. Aber ich bitte diejenigen auf der ganz rechten und ganz linken Seite, sich pointiert zu Wort zu melden, damit ich das erkennen kann.

Dann steigen wir jetzt in die Tagesordnung ein.

Tagesordnungspunkt 1

Patientenzufriedenheit

Beate Wieseler: In dieser Aktualisierung des Methodenpapiers haben wir in dem Kapitel zur Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens die ergänzende Darstellung der Patientenzufriedenheit herausgenommen.

Der Grund dafür ist, dass die Patientenzufriedenheit im SGB V einschließlich der Arzneimittelnutzenbewertungsverordnung nicht als relevantes Kriterium für die Bewertung von medizinischen Interventionen genannt wird. Gleichzeitig gehen wir davon aus, dass Aspekte oder Auswirkungen von Patientenzufriedenheit, die sich auf die Gesundheit von Patienten beziehen - nur darum geht es im SGB V -, in den genannten Zielgrößen, also in der Mortalität, in der Morbidität oder in der gesundheitsbezogenen Lebensqualität abbildbar sind.

Darauf ist teilweise in den Stellungnahmen eingegangen worden. Es wurden auch solche Beispiele genannt, wie: Wenn das Ergebnis einer Behandlung eine Verbesserung der Symptomatik darstellt, ist natürlich denkbar, dass das zu einer erhöhten Patientenzufriedenheit führt. - Wir sehen in diesem Fall dann das relevante Ergebnis der Behandlung in dieser verbesserten Symptomatik abgebildet.

Also: Wie gesagt, aus unserer Sicht sind gesundheitsbezogene relevante Aspekte in den genannten Endpunkten bereits abgebildet. Unsere Frage an die Stellungnehmenden ist deshalb, welche Aspekte oder Auswirkungen von Patientenzufriedenheit in Bezug auf die Gesundheit des Patienten sehen Sie nicht durch die Zielgrößen Mortalität, Morbidität oder gesundheitsbezogene Lebensqualität abgebildet? Was ist darüber hinaus denkbar?

Moderator Jürgen Windeler: Die Frage richtet sich an niemanden Konkreten, weil in vielen Stellungnahmen das Thema „Patientenzufriedenheit“ in durchaus großer Variationsbreite adressiert worden ist, auch unter den Aspekten, wie wichtig das ist und was man darunter verstehen kann. Aber die Frage von Frau Wieseler war, glaube ich, ausreichend konkret, dass ich in die Runde schaue, bisher aber noch nicht sehe, dass sich jemand dazu positionieren möchte. Wir können die Tagesordnung bis um 13 Uhr abschließen. - Herr Schinzel.

Stefan Schinzel: Ein Punkt ist im Grunde genommen die Annehmlichkeit oder die Unannehmlichkeit, die mit der Einnahme eines Medikamentes verbunden sind, ob es zum Beispiel Tabletten sind, die der Patient auf eine Reise mitnehmen kann, ob er sich dreimal in der Woche etwas subkutan spritzen möchte. Es ist auch denkbar, dass ein Medikament der Patient in eigener Regie anwendet, und zum anderen muss er in die Klinik oder in die Praxis. Sehen Sie das ausreichend beispielsweise unter Quality of Life abgebildet?

Beate Wieseler: Ich denke, der Hinweis auf Lebensqualität ist hier ganz relevant. Wenn diese Annehmlichkeit dazu führt, dass sich die Lebensqualität des Patienten verbessert, ist das durch einen Lebensqualitätsfragebogen abbildbar und fällt dann auch unter die patienten-

relevanten Ergebnisse, die wir berücksichtigen würden. Ich würde es für unzureichend halten, zu postulieren, dass zum Beispiel eine veränderte Einnahmefrequenz auf jeden Fall eine Verbesserung der Lebensqualität bedeutet. Das wäre aus meiner Sicht dann mit entsprechenden Ergebnissen nachzuweisen.

Moderator Jürgen Windeler: Eine Wortmeldung ganz links.

Andreas Reimann: Zunächst: Dass im SGB V explizit die Patientenzufriedenheit bzw. die Behandlungszufriedenheit nicht erwähnt ist, ist kein neuer Fakt. Trotzdem war es bislang möglich, supportive Informationen hierzu im Rahmen der IQWiG-Methodik einzureichen. Insofern überzeugt mich die Begründung noch nicht völlig. Es muss ein Ereignis eingetreten sein, das eine Veränderung dieser Einschätzung seitens des IQWiG verursacht hat. Ich wäre dankbar, wenn Sie dazu Stellung nehmen könnten.

Aber zu Ihrer Frage: Es ist in der Tat so, dass natürlich Patientenzufriedenheit und Behandlungszufriedenheit - hier wäre möglicherweise zu differenzieren - kein Surrogat für Lebensqualität sind. Das ist, glaube ich, völlig unbestritten. Es ist dringend erforderlich, einen Missbrauch dieser Terminologie im Sinne eines Surrogats für Lebensqualität zu vermeiden, weil das gerade Patientenpräferenzen abwertet. Es rückt sie sozusagen in die Nähe von Marketingstatements. Das ist wirklich sehr, sehr unerfreulich. Das halten wir als ACHSE nicht für sinnvoll.

Aber wenn ich ein Arzneimittel oder ein nichtmedikamentöses Verfahren bewerte, möchte ich doch alle mögliche Evidenz - auch wenn sie nachrangig ist und nicht zu den 3 primären Kriterien direkt gehört - bewerten und dann den Gesamtbestand dieser Evidenz im Hinblick auf Hinweise - übrigens auch auf zusätzlichen Schaden - bewerten.

Aber Ihre Frage bezog sich darauf, ob wir hier einen Zusammenhang mit dem Regelungszusammenhang des SGB V, also der Krankenversicherung natürlich, sehen können, denn die Zweckmäßigkeit eines Verfahrens ist, wie bekannt, Voraussetzung - § 12 SGB V - für die Erstattungsfähigkeit. Es kann durchaus ein Verfahren unzweckmäßig sein, wenn die Patientenzufriedenheit abnimmt, wenn sie richtig erhoben worden ist. Wenn man beispielsweise vorher feststellt, dass Patienten gar nicht einmal monatlich eine Therapie bekommen wollen, sondern lieber täglich, weil sie damit besser zurechtkommen oder mit solchen Einnahmeschemata wie zweimal wöchentlich nicht gut zurechtkommen, sie ein Hersteller aber einführt, kann das ein Hinweis sein, dass es gar keine zweckmäßige Therapie ist. Umgekehrt natürlich auch - das ist überhaupt kein Frage. Das kann auch im Positiven sein.

Deswegen ist mein Petition, supportive Informationen immer aufzunehmen und nicht von vornherein methodisch abzulehnen. - Vielen Dank.

Moderator Jürgen Windeler: Wollen wir Prof. Sauerbruch noch dazu nehmen?

Tilmann Sauerbruch: Wir haben diesen Punkt auch kurz adressiert. Mir ist nicht ganz klar, warum Sie die Patientenzufriedenheit gestrichen haben. Vielleicht habe ich auch nicht richtig aufgepasst. Könnten Sie das noch einmal klarstellen? Das ist in Ihrem Paper auf Seite 39.

Moderator Jürgen Windeler: Frau Wieseler.

Beate Wieseler: Wir haben im Grunde genommen eine gewisse Schärfung der Definition auf Basis des SGB V vorgenommen. Das Ganze ist in der Arzneimittelnutzenbewertungsverordnung durch die Definition des Nutzens und die Definition von therapeutischen Effekten vorgenommen worden. In der Folge haben wir diese Anpassung vorgenommen.

Ich komme auf Ihren Punkt zurück, Herr Reimann, mit der Zweckmäßigkeit der Therapie. Auch die Zweckmäßigkeit der Therapie bezieht sich im SGB V aus meiner Sicht auf den Gesundheitszustand des Patienten. Im Grunde genommen postulieren Sie jetzt, dass die Zufriedenheit den Gesundheitszustand beeinflusse - im positiven wie im negativen Sinne, wenn Sie sagen, es könne ein Nutzen durch die Patientenzufriedenheit abgebildet werden, aber auch ein Schaden, weil der Patient unzufrieden ist und deshalb die Therapie gegebenenfalls nicht adäquat wahrnimmt.

Ich würde nicht diese Verbindung postulieren wollen, sondern ich würde tatsächlich den Einfluss auf die Gesundheit des Patienten abgebildet sehen wollen. Deshalb unser Vorschlag, das Ganze nicht über die Patientenzufriedenheit abzubilden, sondern tatsächlich über den konkreten Einfluss gegebenenfalls dieser Zufriedenheit auf das Ergebnis bezüglich der Gesundheit, auf die Mortalität, die Morbidität oder die Lebensqualität.

Moderator Jürgen Windeler: Herr Mühlbacher.

Axel Mühlbacher: Ich würde gern auf 2 Punkte eingehen. Das war einmal das Thema, dass die Anwendung für den Patienten eine Rolle spielen könnte. Hier war es noch mal die Frage, wenn man es streicht, was ist passiert, dass man es heute, nach 5 Jahren, streicht.

Noch einmal zur Vertiefung: Theoretisch ist es sicherlich sinnvoll, darauf zu verweisen, dass die Form der Anwendung oder die Zufriedenheit Auswirkungen auf die Lebensqualität haben kann. Das kann dann wieder durch ein Tool abgebildet werden. Pragmatisch stellt sich die Frage: Welches Lebensqualitätsinstrument ist in der Lage, so sensitiv diese Dinge aufzunehmen? Zum Zweiten stellt sich die Frage: Wird die Lebensqualität als Instrument, als Tool zur Erfassung oder Messung dieser Größe, jeweils wieder neu entworfen?

Beate Wieseler: Ich denke, zur Begründung habe ich bereits Stellung genommen. Was Ihre Frage zur Lebensqualität angeht, denke ich schon, dass es, wenn es eine Auswirkung der Patientenzufriedenheit auf die Lebensqualität gibt, mit vorhandenen Instrumenten abbildbar ist. Ich würde nicht sehen, warum die vorhandenen Lebensqualitätsinstrumente nicht in der Lage wären, das abzubilden. Da gibt es auf der einen Seite generische Instrumente, die Sie erkrankungsübergreifend einsetzen können. Wenn das spezifische Aspekte sind, sich also eine

Zufriedenheit in einer krankheitsspezifischen Lebensqualität niederschlagen würde, gibt es auch, denke ich, Fragebögen, die das abbilden können.

Moderator Jürgen Windeler: Herr Mühlbacher.

Axel Mühlbacher: Dürfte ich ganz kurz - nur zum Verständnis - nachhaken: Gehen Sie davon aus, dass ein EQ-5D und ein SF-6D in der Lage sind, zum Beispiel unterschiedliche Formen der Anwendung - um es ganz konkret zu machen: Einnahme von Tabletten oder Spritze - in Form der Lebensqualität abzuzeichnen?

Beate Wieseler: Der EQ-5D ist sicherlich nicht der beste Fragebogen, um Lebensqualität zu messen. Ich glaube, da sind wir uns vielleicht einig. Die Frage ist: Was wollen wir abbilden? Was ist Lebensqualität?

Lebensqualität ist der Einfluss der Erkrankung oder der Behandlung auf das physische, psychische oder soziale Wohlbefinden des Patienten. Wenn die Tatsache, dass ich 1 Tablette statt 2 Tabletten nehme, einen Einfluss auf diese Faktoren hat, denke ich: Die Fragebögen können das abbilden. Wenn das keinen Einfluss auf diese Faktoren hat, bilden die Fragebögen das natürlich auch nicht ab. Dann wäre es für mich aber eben auch kein Outcome, kein Effekt bezüglich der Lebensqualität.

Moderator Jürgen Windeler: Herr Ingenhaag.

Michael Ingenhaag: Grundsätzlich sind wir der gleichen Meinung wie Herr Reimann, dass die gesamte Evidenz angeschaut und auch bewertet werden soll, damit man diese Transparenz hat.

Wir stellen uns weiterhin 2 Fragen. Zum einen: Welchen Effekt hat das Streichen der Patienten- bzw. Behandlungszufriedenheit auf die Versorgungssicht und die Patientensicht, wenn es danach zum G-BA geht, bei dem die Versorgungsfrage liegt, wenn in dem Dossier darüber keine Informationen mehr vorhanden sind?

Zum Zweiten: Wo geht die Patientensicht ein, wenn die Patientenzufriedenheit und die Behandlungszufriedenheit gestrichen werden?

Moderator Jürgen Windeler: Dann nehmen wir noch Herrn Dippel und Herrn Dintsios dazu.

Franz-Werner Dippel: Ich möchte den Punkt von Herrn Mühlbacher unterstützen. Es gibt in zahlreichen Indikationen sicher keine wirklich guten krankheitsspezifischen Messinstrumente. Es gibt aber sehr wohl gute Messinstrumente für die Patientenzufriedenheit. Heißt das denn, dass sozusagen mit dem Wegfall der Patientenzufriedenheit auch automatisch diese Messinstrumente nicht mehr akzeptiert werden?

Moderator Jürgen Windeler: Herr Dintsios.

Ch.-Markos Dintsios: Unabhängig von der Sozialrechtsexegese, die das IQWiG betreibt, finde ich, dass das Sozialgesetzbuch kein enges Korsett ist. Das IQWiG schreibt sich auf seine Fahnen, dass es ein wissenschaftliches Institut ist. Das heißt: Es könnte darüber hinaus durchaus Endpunkte einbeziehen, die nicht so explizit in entsprechenden Paragrafen des SGB V enthalten sind.

Zweitens. Der Endpunkt wurde früher genannt - schon seit der Gründung des IQWiG. In Wahrheit war er im ersten Methodenpapier drin, und er hat niemanden gestört in dieser Hinsicht. Ich sehe auch keinen Konflikt.

Dritter Punkt. Frau Wieseler hat eine Kausalität aufgebaut. Wenn die Patientenzufriedenheit relevant ist und Einfluss auf die gesundheitsbezogene Lebensqualität nimmt, wird das dort auch abgegriffen. Die Frage ist: Ich kann mir die Kausalität genauso umgekehrt vorstellen. Wenn die Lebensqualität tangiert ist, kann auch das eine Auswirkung auf die Patientenzufriedenheit haben. Wieso gibt es einen monokausalen Zusammenhang in die eine Richtung und nicht in die andere Richtung? Darüber lässt sich trefflich streiten, auch philosophisch ein bisschen.

Um das Ganze hier jetzt abzuschließen: Dieser Endpunkt hat niemanden gestört. Es gibt einige Fälle, in denen er auch mit bedient wurde - nicht so oft übrigens. Er war bis jetzt nicht entscheidungsrelevant. Aber was macht man in einer Situation, in der die vorliegende Evidenz über alle anderen Endpunkte in ein Patt führt? Wieso sollte man nicht die gesamte verfügbare Evidenz noch in die Bewertung einbeziehen. Wenn dieser Endpunkt bedient würde in so einem Fall - was spricht dagegen, ihn zu berücksichtigen?

Ich halte den Akt des IQWiG, diesen Endpunkt aus dem Methodenpapier herauszunehmen, nicht für eine glückliche Entscheidung.

Beate Wieseler: Zur Frage: Wo geht die Patientensicht in die Bewertung ein? Zum einen ist es natürlich wünschenswert, dass die Patientensicht direkt in den Studien in die Bewertungen eingeht, indem ich Patienten zu ihrer Symptomatik und zu ihrer Lebensqualität befrage. Die Patientensicht geht natürlich auch im G-BA in das Verfahren mit ein, weil - ich denke, wie Sie wissen - es eine Patientenbeteiligung beim G-BA gibt und auch solche Aspekte, wie wir sie gerade angesprochen haben, von den Patienten in die Diskussion eingebracht werden können. Da sehe ich jetzt nicht, dass die Patientensicht im Verfahren relevant dadurch geschwächt würde, dass diese Patientenzufriedenheitsfragebögen, die wir im Übrigen sehr selten sehen und wo es natürlich sehr viel weniger gute Instrumente gibt, als es bei der Lebensqualität der Fall ist ... Das adressiert, glaube ich, auch Ihre Frage schon.

Die Erfassung von Lebensqualität ist sehr viel weiter entwickelt als die Erfassung von Patientenzufriedenheit. Es gibt sicherlich, wenn man sich die Fragebögen anschaut, an der einen oder anderen Stelle auch eine Art Kontinuum, dass es Fragebögen gibt, die sowohl Aspekte von Patientenzufriedenheit als auch Aspekte von Lebensqualität abbilden. In der

Regel ist es so, dass, wenn uns Fragebögen vorgelegt werden, wir sie schon anschauen, um zu sehen: Was genau bilden sie denn ab? Wie gesagt: Ich würde nicht von einem Verlust von Informationen ausgehen bezüglich der Lebensqualität, wenn wir die Patientenzufriedenheit nicht berücksichtigen.

Bezüglich der umgekehrten Kausalität von Lebensqualität und Patientenzufriedenheit: Das ist durchaus möglich, klar. Das kann in beide Richtungen gehen. Wenn ich aber primär den patientenrelevanten Endpunkt messen möchte und ihn direkt messen kann, würde ich das vorziehen. Wenn ich im Grunde genommen die Patientenzufriedenheit als einen Ersatz für die Lebensqualität messen möchte, ist das ein Surrogat. Wenn das validiert ist, kann man das verwenden. Solche Daten wurden uns allerdings bisher nicht vorgelegt.

Moderator Jürgen Windeler: So, jetzt habe ich eine Wortmeldung ganz rechts.

Marianne Abele-Horn: Ich habe eine Frage: Sprechen Sie von Mortalität oder von Letalität?

Beate Wieseler: Von Mortalität.

Marianne Abele-Horn: Bezogen auf die Bevölkerung, nicht aufs einzelne Krankheitsbild?

Beate Wieseler: In diesem Sinne dann eher von Letalität.

Marianne Abele-Horn: Letalität, was auf ein Krankheitsbild bezogen ist?

Beate Wieseler: Ja.

Marianne Abele-Horn: Das ist wichtig, dass man das epidemiologisch klärt.

Dann hätte ich noch ... Darf ich zur Patientenzufriedenheit etwas sagen, obwohl ich von der DGHM komme?

Beate Wieseler: Aber sicher!

Marianne Abele-Horn: Ich finde es insofern wichtig, weil es die Adhärenz der Behandlung beim Patienten steigert. Insofern halte ich das für einen wichtigen Punkt. Ich denke, gerade wenn man Leitlinien anschaut, die nicht umgesetzt werden, weil die Leute nicht zufrieden damit sind, hat das schon einen gewissen Stellenwert.

Beate Wieseler: Vielen Dank. - Sie haben die Adhärenz ins Spiel gebracht. Das ist tatsächlich ein Punkt, der in vielen Stellungnahmen erwähnt wird. Eine gesteigerte Adhärenz hat aus unserer Sicht nur einen Vorteil für Patienten, wenn es zu einem besseren Ergebnis bezüglich der Gesundheit kommt, also zu einem verbesserten Ergebnis zum Beispiel bezüglich der Symptomatik.

Marianne Abele-Horn: Schon auch, wenn er zufrieden ist. Wenn er der Behandlung nicht folgt oder sie abbricht, weil er unzufrieden ist, wird die Behandlung keinen Erfolg haben. Ich denke, sie hat nur einen Erfolg, wenn der Patient mitgeht, wenn er mit diesem Behandlungsregime zufrieden ist und es akzeptieren oder besser akzeptieren kann.

Beate Wieseler: Die Adhärenz wird graduell gegebenenfalls beeinflusst. Aber auch da ist der direkt relevante Punkt das Ergebnis beispielsweise bezüglich der Symptomatik.

Moderator Jürgen Windeler: So, jetzt habe ich eine Wortmeldung von einem Herrn ganz links.

Andreas Reimann: Die Gleichsetzung „hohe Patientenzufriedenheit - hohe Adhärenz“ ist sehr gefährlich. Das muss man ganz klar sagen. Überhaupt wird der Adhärenz-Begriff leider aus unserer Sicht viel zu synonym mit dem Begriff „Compliance“ verwendet. Er ist letztlich ... Man sagt zwar ganz bewusst „Adhärenz“, weil man meint, damit neuere Diskussionen aufzugreifen, aber man tut es eigentlich nicht. Adhärenz tritt nur dann ein, wenn vorher individuell - das ist natürlich dann die Mikroebene zwischen Arzt und Patient - eine Einigung über das dem Patienten Mögliche und Sinnvolle stattgefunden hat. Genau das wird nicht nachgeprüft.

Bei den meisten sogenannten Adhärenzuntersuchungen und Programmen - meistens sind es letztlich Compliance-Programme nach dem Motto: Wie viel von dem verschriebenen Medikament wird tatsächlich dispensiert oder eventuell auch eingenommen? - ... Das ist aber nicht Adhärenz. Das aber nur nebenbei. Das halte ich in diesem Fall in der Tat für gar nicht so hilfreich, weil es ablenkt und weil es in der Tat methodisch nicht richtig ist.

Unsere große Sorge - auch als Allianz Chronisch Seltener Erkrankungen - ist: Wenn das IQWiG - und das IQWiG ist nicht irgendetwas, das IQWiG ist inzwischen ein weltweit anerkanntes und auch wissenschaftsmitbestimmendes Institut für Health Technology Assessment - sagt - ich sage das jetzt etwas zugespitzt -: „Uns interessiert die Patientenzufriedenheit nicht“, entsteht auch kein weiterer Forschungsbedarf. Ich sehe genau diesen Mangel bei den Skalen. Bei Patientenzufriedenheit ist er natürlich evident.

Der Missbrauch, den ich vorhin dargestellt habe, ist auch evident. Er wird aber nicht besser, wenn Sie streichen, sondern er wird schlechter. Denn dann kümmert sich keiner mehr darum. Dann ist es Blümchenforschung. Das haben wir bei der Lebensqualität gesehen. Lebensqualität ist dann interessant, wenn sie in der Tat relevant bewertet wird.

Noch einmal: Patientenzufriedenheit ersetzt nicht Lebensqualitätsmessung - überhaupt nicht -, kann aber ein Hinweis sein. Allein schon aus dem Grund, einen weiteren Forschungsanreiz zu geben, würde ich ihn in den Methoden lassen, was nicht heißt, dass Sie das berücksichtigen müssen. Denn Sie schreiben an anderer Stelle zu Recht: Skalen müssen validiert sein, Instrumente müssen validiert sein - egal, was sie nun messen.

Man sieht das bei der Verfahrensbewertung oder bei der Bewertung im Krankenhaus von Krankenhausprozessen, die gar nichts mit medikamentösen Behandlungen zu tun haben. Hier hat das Picker Institut Verdienstvolles geleistet und eine neue Form der Patientenzufriedenheitsmessung eingeführt, indem man nämlich vorher fragt: Was ist eigentlich relevant? Sind Wartezeiten für dich relevant beispielsweise? Oder ist die Sauberkeit relevant? Oder was auch immer. Das hilft Krankenhäusern, besser zu werden, wenn sie das sorgfältig durchführen, in ihren patientenorientierten Prozessen. Das wünsche ich mir natürlich auch für die pharmazeutische Industrie und auch für Medizinprodukte.

Daher noch einmal mein Petitum: Lassen Sie es drin! Das heißt noch lange nicht, dass Sie alles bewerten müssen.

Moderator Jürgen Windeler: Herr Rasch war jetzt dran.

Andrej Rasch: Ich schließe mich den bereits getätigten Schlussfolgerungen an. Wir verstehen die im SGB V genannten Kategorien auch nicht abschließend.

Des Weiteren gibt es anzumerken, dass die Berücksichtigung der Patientenzufriedenheit auch nicht zwangsläufig gleichrangig mit anderen Kategorien gesehen werden muss. Sie soll auch keine anderen Kategorien ersetzen. Natürlich wären die vorhandenen Kategorien, wenn die vom IQWiG als adäquat in die Bewertung eingeschlossen werden, natürlich sehr gut geeignet, wie das bislang gehandhabt wird.

Nur sehen wir zahlreiche Verfahren, in denen die Studien in Gänze nicht eingeschlossen werden. Wir sehen aber auch Verfahren, in denen bestimmte Endpunktkategorien oder bestimmte verwendete Instrumente aus welchen Gründen auch immer - aus methodischen Schwächen oder aus Schwierigkeiten mit der Durchführung - als nicht adäquat oder nicht verwertungswürdig erachtet werden. Dann, in diesen vorliegenden Fällen, ist es, glaube ich, auch im Sinne des Konzeptes der bestverfügbaren Evidenz schon auch sehr sinnvoll, die vorhandenen Instrumente der Patientenzufriedenheit dann zu dem Evidenzkörper hinzuzählen und dem G-BA als Empfehlung weiterzuleiten.

Moderator Jürgen Windeler: Frau Wieseler.

Beate Wieseler: Herr Rasch, diese Argumentation kann ich so überhaupt nicht nachvollziehen. Wenn wir keine adäquaten Daten bezüglich patientenrelevanter Endpunkte haben, sollen wir ersatzweise Informationen zur Patientenzufriedenheit zeigen?

(Andrej Rasch schüttelt den Kopf.)

- So habe ich Sie jetzt verstanden.

Das würde aus meiner Sicht nicht zu einer adäquaten Entscheidungsfindung führen.

Andrej Rasch: Gerade das habe ich eben verneint. Nicht „ersatzweise“, sondern in den Fällen, in denen bestimmte Endpunkte als nicht verwertungswürdig eingestuft werden seitens des IQWiG, stellt die Patientenzufriedenheit die bestverfügbare Evidenz dar. Diese sollte nicht komplett gestrichen werden, sondern als Evidenzkörper entsprechend aufbereitet werden.

Gerade bei Konstellationen wie zum Beispiel seltenen Erkrankungen gibt es durchaus Fälle, bei denen es keine gängigen Instrumente zur Messung der Lebensqualität gibt, keine gängigen Responseschwellen, die akzeptiert sind. Da ist es doch evident, dass man die bestverfügbare Evidenz im Sinne der Patientenzufriedenheit auch dazuzählen könnte - gerade bei solchen Indikationsgebieten, wo keine gängigen Standards hinsichtlich der methodischen Messung von Lebensqualität herrschen.

Moderator Jürgen Windeler: Ich habe die Zwischenbemerkung, dass man ein bisschen aufpassen muss, damit nicht eine vielleicht sehr gute Absicht mit sehr erkennbaren Instrumentalisierungen vermenget wird. Ich glaube, dass einige Punkte in eine gut nachvollziehbare Richtung gehen. Aber wenn das Ganze eine Schlagseite bekommt - ich glaube, dass der Begriff der bestverfügbaren Evidenz hier extremst ausgenutzt - um nicht sogar zu sagen: missbraucht - wird, bin ich nicht sicher, ob wir diesem Ziel hier etwas Gutes tun. - Jetzt spricht Herr Sauerbruch.

Tilmann Sauerbruch: Ich möchte eine kurze Bemerkung machen. Sie haben initial in dieser Diskussion juristisch argumentiert, im weiteren Verlauf aber auch methodisch. Die methodische Begründung erschien uns nicht ausreichend in diesem Papier.

Moderator Jürgen Windeler: Ich würde vorschlagen: Wir sammeln erst noch einmal. - Frau Schatz.

Isabelle Schatz: Wir würden uns ganz gern dem anschließen, was Herr Reimann schon gesagt hat - und zwar, das Potenzial, das wir darin sehen, den Endpunkt weiterhin zu bewerten, und dies als supportive Evidenz heranzuziehen, weil wir uns für die Evaluierung der Instrumente, die hier kritisiert wurden, erhebliches Potenzial versprechen.

Was ist das Spezifische daran? Warum setzen wir uns ein für die Beibehaltung der Patienten- und Behandlungszufriedenheit? Es sind patientenberichtete Studienendpunkte. Was haben wir noch für patientenberichtete Studienendpunkte? Die Lebensqualität. Wir wissen aber, dass es herstellerseitig mit der Erhebung von patientenberichteten Studienendpunkten momentan noch recht dürftig aussieht.

Die Frage ist, inwiefern die spezifische Perspektive des Patienten in die Nutzenbewertung mit einbezogen wird. Hier sagen wir: Wir haben herstellerseitige Bemühungen, so etwas zu erheben und vorzulegen. Wir würden uns wünschen, dass das auch weiterhin berücksichtigt sowie bewertet wird und dass hier vielleicht Verbesserungsbedarf, was die Validierung solcher Instrumente betrifft, aufgezeigt wird, dass aber nicht per se ein patientenberichteter

Studienendpunkt, der eine eigene Qualität hat, an und für sich aus dem Verfahren ausgeschlossen wird.

Moderator Jürgen Windeler: Jetzt kommt Herr Lange dran.

Stefan Lange: Es gibt ein paar Stichwörter. Ich hätte auch noch etwas zum Evidenzkörper und zur Gesamtevidenz zu sagen, aber das spare ich mir jetzt.

Das habe ich nicht ganz verstanden, Frau Schatz, mit den PROs. Wir haben an vielen Orten patientenberichtete Endpunkte - das fängt bei den Symptomen an - über die gesundheitsbezogene Lebensqualität. Ich bin nicht sicher, ob man da noch einen Gefallen tut, indem man die Patienten nach ihren Zufriedenheiten befragt. Ich glaube, dass das wieder die Gefahr birgt, diesen Begriff etwas zu entwerten.

Ich habe tatsächlich noch eine Frage, vielleicht noch tatsächlich einen Punkt. Es wurde eben angedeutet, das alles ließe sich durch Lebensqualitätsinstrumente gar nicht abbilden. Ich glaube, da muss man differenzieren zwischen generischen und erkrankungsbezogenen Lebensqualitätsinstrumenten, die erkrankungsbezogen eine ganz andere Sensitivität haben. Da kann sich so etwas vielleicht dann doch ausdrücken.

Ich hätte trotzdem eine Frage an Herrn Reimann, weil er gesagt hat - das habe ich noch nicht so richtig verstanden -, das sei gar kein Ersatz und solle gar kein Surrogat sein für so etwas wie gesundheitsbezogene Lebensqualität. Dann frage ich mich allerdings: Was ist eigentlich das Konstrukt? Was ist denn Zufriedenheit? Ob nun Patienten- oder Behandlungszufriedenheit - wir haben für Lebensqualität ein ganz gutes Konstrukt inzwischen. Ich habe so etwas - ehrlich gesagt - für Zufriedenheit an sich - nennen wir es von mir aus „Patientenzufriedenheit“ - noch nicht so wirklich gefunden.

Man kann es natürlich auch umgekehrt sehen. Man kann das auch als Anreiz verstehen, da vielleicht ein bisschen voranzugehen, das erst einmal tatsächlich in Abgrenzung von Lebensqualität zu definieren und dann zu schauen, welchen Wert es tatsächlich sozialrechtlich hat oder haben kann. Denn beim Stichwort „Zufriedenheit“ kann man sich schon - das hatten Sie selbst gesagt - verschiedene Missbrauchsszenarien vorstellen, die wir lieber nicht haben wollen.

Das wäre ein bisschen mein Petitum und ein Stück weit die Antwort auf methodische Probleme, die wir mit diesem Begriff haben, und darauf, warum wir ihn jetzt herausgenommen haben.

Moderator Jürgen Windeler: Bitte.

Andreas Reimann: Das ist genau der Punkt. Zu Recht sagen Sie: Hier muss man stärker abgrenzen. Und wenn man sagt: Das Gesundheitssystem - ich weiß, dass das jeder sagt, aber keiner meint - ist für den Patienten da - es ist klar, dass das keiner meint; ich bin nicht naiv,

aber ich behaupte mal, es gibt vielleicht Leute, die es doch ernst meinen -, dann muss man den Patienten auch fragen: Wie zufrieden bist du eigentlich mit dem, was jetzt gerade mit dir gemacht wird? Zufriedenheit ist immer ein Differenzgefühl zwischen Erwartung und Erlebtem. Wenn man mehr erlebt, als man erwartet hat, ist möglicherweise eine Zufriedenheit da. Es gibt Menschen, die hochzufrieden sind, obwohl sie schwerstkrank sind, und die auch mit einer Behandlung hochzufrieden sein können, obwohl sie vielleicht sehr belastend ist. Es ist in der Tat also weiterer Forschungsbedarf da. Nur mein Petitum war gerade: Wenn Sie das hier herausstreichen, hat es einfach eine Negativaussage: „Es ist eigentlich egal. Das ist eigentlich wurscht. Wozu sollen wir das machen? Selbst oder gerade das IQWiG streicht es raus, also ist es zum Beispiel für die Versorgungsforschung irrelevant.“ Ich wünsche mir gerade, dass es relevant wird, in der Tat sorgfältig gemacht und nicht missbraucht wird.

Moderator Jürgen Windeler: Sie haben mir natürlich jetzt die Gelegenheit gegeben, die ich auch nutze, zu sagen, dass sich das IQWiG jedenfalls sehr für die Patientenbelange interessiert und sich dafür einsetzt. - Jetzt ist Herr Dintsios dran.

Ch.-Markos Dintsios: Noch mal zurückkommend auf die Diskussion, die vorangegangen ist, und auch die Fragestellung: Das IQWiG selber - was hat man denn für Erfahrungen gemacht mit diesem Endpunkt, die einen dazu führen, ihn aus dem Methodenpapier herauszustreichen?

Ich mag jetzt 2 Szenarien spielen. Wir haben in unserer Stellungnahme Beispiele vorgetragen, in denen dieser Endpunkt bedient wurde. Entscheidungsrelevant war er nicht. Das heißt: Selbst die Diskussion um den Evidenzkörper an sich in der praktischen Handhabung hatte keine große Auswirkung. Es hat auch nicht gestört in Wahrheit.

Der zweite Punkt ist: Ich bleibe jetzt noch mal beim Evidenzkörper. Nicht nur die bestverfügbare Evidenz, vielleicht ist es manchmal auch die allein verfügbare Evidenz. Ich erinnere an das Hämophilie-Produkt eines Herstellers. Es hatte eine Zulassung, die auf einer pharmakokinetischen Studie fußte, rekombinante Produkte. Ich nenne das Produkt jetzt nicht, um kein Productplacement zu betreiben. Es hat aber auch unter anderem ein spezifische Instrument bedient HämoSats - „Sats“ steht für Satisfaction. Wieso nicht sich auch das mit anschauen? Wenn man den Endpunkt herausnimmt aus dem Methodenpapier, wird man sich nie die Mühe machen, sich dieses Instrument selbst unter den relativ strikten Kriterien des IQWiGs anzuschauen. Deshalb noch einmal mein Petitum - da schließe ich mich auch Herrn Reimann an -: Es stört keinen in Wahrheit, es sei denn, es kommt das Argument „aus pragmatischen Gründen“, weil wir nicht die Zeit oder die Ressourcen dafür haben. Dann ist es allerdings ein ökonomisches Argument und kein methodisches und vor allem kein rechtliches.

Ich würde mir vom IQWiG wünschen, nicht so viel Rechtsexegese zu betreiben. Denn Sie sind Methodiker. Ich kenne Sie ja. Ich habe selbst eine IQWiG-Vergangenheit. So viel juristischen Sachverstand, glaube ich, hat jemand anderes im IQWiG, der heute nicht präsent ist. Deswegen noch einmal: Lassen Sie es doch sein. Es stört wirklich keinen und Sie am wenigsten.

Moderator Jürgen Windeler: Herr Neubauer.

Aljoscha S. Neubauer: Vielleicht noch einmal einen Gesichtspunkt zur Einordnung, wieso so unterschiedliche Schwerpunkte auf Patientenpräferenz gelegt werden. Wir hören auf der einen Seite mehr aus der praktischen Sicht, wo es mehr um die Effectiveness geht. Daher ist natürlich Patientenpräferenz etwas, das ganz entscheidend ist.

Frau Wieseler, Ihre Argumentation kommt mehr aus der methodischen Beurteilungsebene, wo Sie in der Regel auf randomisierten kontrollierten Studien, also relativ nah an der echten Efficacy sind. Da ist natürlich die Patientenrelevanz weitgehend durch andere Maße abbildbar. Das gilt natürlich in dem Maße nicht, wenn wir mehr in Richtung einer realen Anwendung, also einer Effectiveness gehen, wo dann natürlich Patientenzufriedenheit - dann kommen die anderen Aspekte, etwa Adhärenz etc., wieder hinein - die große Rolle spielt. Insofern wäre ich sehr beim Herrn Dintsios, als man sagt: Der Aspekt spielt eine Rolle. Die Frage ist wahrscheinlich, an welcher Stelle und in welcher Gewichtung, um diesen Missbrauch, den wir, glaube ich, alle nicht wollen, zu verhindern.

Moderator Jürgen Windeler: Ich sehe jetzt im Moment keine weiteren Wortmeldungen. - Doch, ich habe mich nicht richtig umgesehen. Ich gebe erst einmal Herrn Stengel das Wort.

Dirk Stengel: Wir haben uns hier methodisch ganz schön ausgelebt. Ich versuche, hier noch einmal einen kleinen Konsensvorschlag zu unterbreiten. Tatsächlich ist es so, dass in den USA momentan das Promise-Instrument eine ganz große Rolle spielt. Das ist im Grunde genommen ein adaptives System, das die bisherigen Schwächen sehr starrer Instrumente wie zum Beispiel SF-36 oder EQ-5D versucht, aufzuheben durch computer-adapted testing. Hier wird versucht - den Beitrag von Stefan Lange fand ich sehr gut -, das Prinzip der Zufriedenheit als eine Domäne mit zu berücksichtigen. Vielleicht könnte man mindestens in das Papier mitaufnehmen, dass derartige Entwicklungen möglicherweise momentan betrachtet bzw. auch finanziert werden - leider momentan noch nicht in Europa. Es gibt eine große Item-Datenbank. So etwas ist in Entwicklung. Vielleicht erwartet uns so etwas in 2 Jahren.

Moderator Jürgen Windeler: Dann sehe ich jetzt - auch mit sorgfältigem Umsehen - tatsächlich keine weiteren Wortmeldungen mehr.

Wie weit wir jetzt die Differenzierung zwischen Efficacy und Effectiveness bemühen sollten und was das eigentlich genau ist - da lohnt sich sicherlich, mal eine eigene Veranstaltung darüber zu machen.

Ansonsten würde ich diesen Punkt abschließen. Wir werden noch einmal Ihre Argumente abwägen und uns dann entscheiden, ob wir bei unserem ursprünglichen Vorgehen bleiben, oder ob wir das noch einmal adaptieren.

Dann kommen wir jetzt zu TOP 2.

Tagesordnungspunkt 2

Einbezug von Patientenpräferenzen in (K)NB

Moderator Jürgen Windeler: Ich will an dieser Stelle mein Bedauern ausdrücken, dass leider kein Vertreter der Deutschen Gesellschaft für Gesundheitsökonomie teilnehmen konnte bzw. sich die Deutschen Gesellschaft für Gesundheitsökonomie nicht in der Lage gesehen hat, einen Vertreter hierher zu schicken, was für die Diskussion insgesamt, aber auch speziell für diesen Punkt ein bisschen bedauerlich ist. Aber nichtsdestotrotz würde ich jetzt Herrn Gerber-Grote bitten, den Punkt zu übernehmen.

Andreas Gerber-Grote: Das IQWiG beschäftigt sich nun seit 4 Jahren mit der Frage der Patientenpräferenzen intensiv. Sie alle haben wahrscheinlich wahrgenommen: Es gab 2 Pilotprojekte, die publiziert wurden: eines zum Analytic Hierarchy Process und eines zur Conjoint-Analyse. Wir hatten in der Methodenversion 4.1 dazu kurz Stellung genommen. Damals gab es schon sehr viel Interesse. Wir hatten damals vertröstet, dass, erst nach Publikation des Pilotprojekts zur Conjoint-Analyse, das in diesem Juni publiziert wurde, wir die Diskussion groß eröffnen wollen, damit man auf Basis beider Pilotprojekte diskutieren kann.

Deswegen unsere konkrete Frage - sie basiert ausgewählt auf einer Stellungnahme des vfa, aber ich denke, sie steht stellvertretend auch für andere Stellungnahme -: Ein Stellungnehmender - eben der vfa - schreibt, dass die Methodengegenüberstellung und Diskussion der Erhebung von Patientenpräferenzen nicht anhand der Kriterien der klassischen Sozialempirie wie Repräsentativität, Objektivität, Reliabilität und Validität erfolgten. Unsere Frage: Könnte der Stellungnehmende uns bitte erläutern, wie die von ihm genannten Kriterien angewendet werden sollen und in welcher Form die genannten Kriterien erfüllt werden sollten?

Moderator Jürgen Windeler: Herr Dintsios.

Ch.-Markos Dintsios: Das kommt aus meiner eigenen Feder. - Im Methodenpapier wird Bezug genommen, nachdem die Verfahren kurz angerissen werden, auf die Schwierigkeit, bei den Teilnehmern die entsprechenden Fragen zu beantworten zur Anzahl usw. usf. - sicherlich wichtige Aspekte, aber nicht unbedingt die sehr harten Kriterien, die man in der Sozialempirie anwendet. Unser Petition war ganz einfach, sich einer Diskussion zu öffnen bzw. auch im Rahmen der beiden Pilotprojekte, die durchgeführt wurden, eine Gegenüberstellung zu wagen. Ich denke, die Begriffe, die genannt wurden, „Reliabilität“, „Validität“ und „Objektivität“ sind selbst definiert. Ich muss sie hier nicht neu definieren. Das würde dann wirklich eine Veranstaltung sprengen.

Darüber hinausgehend glauben wir auch, dass man aus diesen Pilotprojekten durchaus Erkenntnisse hierzu bereits ziehen kann. Um es abzukürzen: Wir fanden die Aussagen bei der Gegenüberstellung der Methoden bzw. die begleitenden Aussagen etwas zu kurz gegriffen

und hätten eine vielleicht tiefergehende Auseinandersetzung mit den entsprechenden Kriterien gewünscht. So ist dieser Passus zu verstehen.

Moderator Jürgen Windeler: Gibt es weitere Wortmeldungen dazu? - Das ist gerade nicht der Fall. Dann würde ich eine Frage anschließen, die das, was Herr Gerber-Grote gerade beschrieben hat, ein bisschen - wie soll ich mal sagen - in der konkreten Praxis beschreibt. Uns allen - ich denke, auch fast allen, die hier sitzen - ist klar, was man unter „Repräsentativität“, „Reliabilität“, „Validität“ und solchen Kriterien versteht. Dürfen wir Äußerungen - das betrifft nicht nur den vfa - für solche Anforderungen an die Bewertung von Patientenpräferenzen bzw. an die Einbeziehung von Patientenpräferenzen in Bewertungen so verstehen, dass diese Kriterien erst zum Beispiel bei Instrumenten oder bei Erhebungen sorgfältig erfüllt sein müssen, bevor wir sie in Bewertungen entnehmen, oder wie soll ich diese Äußerungen verstehen? Denn wir haben in der Tat eine Reihe von Äußerungen gelesen, die gesagt haben: Das muss man noch mal intensiver darlegen. - Aber uns interessiert natürlich, speziell auch im Methodenpapier, nicht nur zu diskutieren, wie man das alles sehen kann, sondern die Frage: Welches sollen eigentlich die Anforderungen sein? - Herr Dintsios.

Ch.-Markos Dintsios: Wenn niemand anderes sich meldet ... Ich will nicht wieder monopolisieren wie letztes Mal, Herr Windeler.

Nein, Sie leiten schon den zweiten Schritt ab, aber ich denke, die Diskussion muss erfolgen. Dann kann man sich auch bemühen um die Aufstellung eines Kriterienkataloges, der als Voraussetzung für die Verwertung dieser Instrumente gelten kann. Aber insofern man die Diskussion gar nicht anreißt - tiefergehend bezogen auf solche Kriterien -, würde ich nicht wagen, die Erfüllung dieser Kriterien voranzustellen, sondern ich würde erst einmal diskutieren: Sind sie anwendbar? Wenn ja, im Vergleich zwischen den beiden Verfahren? Gibt es auch noch andere Verfahren, die infrage kämen? Ich denke an Best-/Worst-Scaling usw. Es gibt auch adaptive Verfahren bei der Conjoint-Analyse. Es gibt ein großes Spektrum. Diese Methoden sind nicht neu. Sie sind für die Biometrie neu. Sie waren für mich auch neu; das muss ich zugeben. Ich komme aus der Gesundheitsökonomie; da haben sie natürlich eine gewisse Diffusion erfahren, aber jetzt seien sie weiter zu verbreiten. So ist das zu verstehen. Sie fordern schon den zweiten Schritt. Ich erinnere ein bisschen an die Diskussion zu den adjustierten indirekten Vergleichen. Da gab es genau die gleiche Diskussion, inwieweit sie verwertbar sind. Am Anfang hatten wir keine scharfen Kriterien. Dann kam die Diskussion um das Kriterium der Devianz. So verstehe ich die Diskussion. Wenn die weiter gediehen ist, kommt auch der Schritt zwangsläufig, den Sie anreißen, der Aufstellung von Gütekriterien für Patientenpräferenzhebungsinstrumente. Das ist die Position des vfa oder zumindest meine.

Moderator Jürgen Windeler: Die ich schwer verstehe. Aber das kann ich gleich kommentieren. - Jetzt kommt Herr Mühlbacher dran.

Axel Mühlbacher: Ich glaube, die Aussagen und Nachfragen hinsichtlich der Darstellung der beiden Methoden AHP und DCE sind sicherlich vor dem Hintergrund gemacht worden, dass

die Aussage immer so gefasst ist, dass wir methodisch noch nicht in der Lage sind, das abzubilden, was man gern hätte, im Bewertungsprozess von klinischer Evidenz. So habe ich es zumindest persönlich gelesen.

Dann werden diese Aussagen garniert mit Feststellungen, die für Experten der Methoden sicherlich nicht haltbar sind. Ich glaube, ich habe in meiner Stellungnahme ein paar angeführt, die sich so nicht mit der Literatur decken. Wenn man den Methodenbericht von außen betrachtet, stellt sich die Frage: Warum haben wir das nicht in letzten 3 Jahren gemacht und diese Fragen zur Validität und zur Verwendbarkeit dieser Resultate gestellt? Denn die Methodenberichte sind schon seit 3 Jahren abgegeben worden. Insofern ist es vielleicht auch ein bisschen die Ungeduld, die dahinter steckt, zu sagen: Wann gehen wir daran? Wann ist eine Methode fertig? Ich kann für mich persönlich gern feststellen, dass ich natürlich verwundert war, über bestimmte Feststellungen zu sagen: „methodisch nicht reif genug“, aber dann die entsprechende Argumentation zu bringen oder sogar auf falsche Dinge zu verweisen.

Moderator Jürgen Windeler: Herr Gerber-Grote.

Andreas Gerber-Grote: Ich denke, was im Detail falsch ist, muss man immer gemeinsam diskutieren. Da habe ich schon oft erlebt, dass Leute meinten, mir oder uns nachgewiesen zu haben, etwas wäre falsch. Ich glaube, da muss man in die Detailformulierung gehen. Das können wir gern machen. Deswegen will ich darauf im Moment auch gar nicht eingehen. Das würde, glaube ich, die Mehrheit der hier Anwesenden eher langweilen.

Aber nehmen wir doch mal den Punkt der Reliabilität. „Reliabilität“ bedeutet, dass ich bei den gleichen Patientinnen und Patienten, wenn ich das gleiche Instrument - sagen wir einmal: tatsächlich eine AHP - nehme, entweder in einem gewissen Zeitraum hintereinander oder in einer anderen Form das gleiche Ergebnis erreiche. Da ist uns eigentlich relativ wenig bekannt, dass es diese Studien gäbe. Insofern geht es gar nicht um eine Diskussion, sondern tatsächlich darum, dass dies erst untersucht werden muss. Das finde ich persönlich gar nicht schlimm, weil ich ein Anhänger der Erhebung von Patientenpräferenzen bin.

Wir können auch gern noch mehr Instrumente diskutieren. Herr Dintsios hat uns eines genannt: Best-Case/Worst-Case. Es gibt noch viele weitere Instrumente. Die kann man natürlich alle hier aufnehmen. Ich glaube, derzeit ist es so sinnvoll zu sagen: Wir machen eine gewisse Forschung, wie ich es genannt habe, zur Reliabilität erst einmal an 2 Instrumenten: einmal, ob die überhaupt in sich reliabel sind, wenn ich sie wiederhole. Bezogen auf die Validität wäre die große Frage: Wenn das Konstrukt „Patientenpräferenzen“ das ist, was ich erheben will - wie auch immer wir das jetzt theoretisch bestimmen -, würde „Validität“ auch bedeuten: Da muss zum Beispiel bei 2 unterschiedlichen Instrumenten, wenn ich sie anwende, auch dasselbe herauskommen. Auch da gibt es ganz, ganz wenige Studien. Die, die sich hier darum kümmern, kennen die Literatur. Es gibt natürlich auch Studien, die zeigen, dass, wenn ich 2 Instrumente nehme, etwas Unterschiedliches herauskommt.

Das ist für mich per se im Moment noch gar kein Grund zu sagen, wir sollten das eine oder das andere nehmen, oder sie treffen sich nicht in der Validität. Aber die Argumente, die ich jetzt angeführt haben, zeigen doch genau, dass wir derzeit nicht einfach sagen können: „Wir setzen diese Methoden ein“, sondern dass wir genau die Forschung brauchen, die jetzt so langsam losgeht, wie ich von einigen befreundeten Instituten weiß: Sie nutzen tatsächlich an derselben Patientenpopulation dasselbe Instrument meinetwegen 4 bis 8 Wochen noch einmal, um zu schauen: Ist es überhaupt reliabel? Insofern, glaube ich, können wir gut vertreten, was wir hier geschrieben haben. Insofern ist es sogar gut, dass wir so dezidiert auf diese Begriffe hier gar nicht eingegangen sind.

Ich erlaube mir, abschließend etwas zur Repräsentativität zu sagen. Zum einen kann man sich natürlich streiten. In der Sozialempirie gibt es unterschiedliche Einschätzungen, ob Repräsentativität überhaupt das Ziel schlechthin ist, oder ob man nicht vielleicht eher sagen soll: Es geht einfach um die Frage: Kann ich in dem Moment, in dem ich eine Patientenpopulation habe, feststellen, wie verzerrt sie gegenüber dem ist, was ich an Durchschnittspopulation zu erwarten habe? Nehmen wir einmal Diabetes mellitus, Bluthochdruck oder irgendetwas; das ist egal.

Das heißt: Je nachdem, wie hoch wir das hängen, wird das schließlich dazu führen, dass wir beim Kriterium Repräsentativität, wenn wir alle ganz streng sein wollen, im Grunde genommen gleich schon sagen können: Das werden die Verfahren nie erreichen können. Denn wie sollen wir diese Patientinnen und Patienten so auswählen, um ein entsprechendes Maß an Repräsentativität zu erheben? Deswegen, glaube ich, ist es sogar aus meiner Sicht gut gewesen und zum Schutz der weiteren Entwicklung dieser Instrumente - wie gesagt: Ich stehe dem sehr positiv gegenüber -, dazu nicht Stellung zu nehmen. Deswegen glaube ich, dass es im Moment richtig war, sich so zu entscheiden.

Moderator Jürgen Windeler: Herr Danner.

Martin Danner: Aus unserer Sicht muss die Methodendiskussion wirklich noch intensiv fortgeführt werden. Es ist auch so, dass man sich fragen kann, ob eine allein sozial-empirische Herangehensweise nicht unterkomplex ist. Die Frage ist, ob hier nicht interdisziplinär gearbeitet werden muss und beispielsweise auch andere Wissenschaftszweige wie die Ethik oder normative Wissenschaften heranzuziehen sind.

Denn das bisherige Modell, das von den Instrumenten verfolgt wird, ist die Vorstellung, dass man Werthaltungen den Patienten zuschreibt, wie man somatische Befunde zuschreibt. Man schaut dann: Was kommt heraus?

Es ist allerdings so, wenn man sich ein bisschen näher damit befasst, dass Werthaltungen natürlich auch immer zerlegt werden können in Argumentationen, wohinter dann auch wieder empirische Befunde stehen. Insofern ist es wiederum die Frage, inwieweit diese reine sozialempirische Herangehensweise adäquat ist. Das gilt insbesondere für die Fragestellung,

inwieweit es zulässig ist, zu aggregieren, weil wir uns hier nicht darüber unterhalten, wie wir Patientenpräferenzen von Einzelpersonen erheben, sondern da wir daran systemisch gehen, geht es um die Frage: Bis zu welchem Grad ist es zulässig, Werthaltungen zu aggregieren, oder inwieweit ist irgendwann der Punkt gekommen, an dem es angezeigt ist, zu differenzieren? Das schlägt sich dann auch bei Nutzenbewertungen nieder in der Frage: Müssen Subpopulationen gebildet werden?

Insofern noch einmal der Appell, sich nicht nur auf die rein sozialempirische Betrachtungsweise zu fokussieren, sondern auch ein bisschen über den Tellerrand hinauszublicken. Da sehen wir durchaus noch weiterführenden wissenschaftlichen Diskussionsbedarf.

Moderator Jürgen Windeler: Herr Dintsios.

Ch.-Markos Dintsios: Ich greife jetzt die Wortmeldung von Andreas Gerber-Grote so auf, dass man sich sozusagen eine Zukunftsagenda beim IQWiG zumindest überlegt hat. So gesehen, wäre es nicht schlecht, diese Termini, die vorhin zitiert wurden, ins Methodenpapier einzubringen. Wenn es der Methodensicherung dient, Fragen hinsichtlich sozialempirischer Gütekriterien zu beantworten, und das ein eigenes Ziel des Instituts sein sollte, könnte das durchaus auch Eingang in die Methodenversion 4.2 finden - wie gesagt, um das Ganze hier nicht wieder in Langeweile zu gestalten. Wir fanden es nur ganz einfach mit den Ausführungen in Methodenversion 4.2 etwas dürftig, in dem Sinne, dass Sachen beschrieben wurden - das war sehr deskriptiv -, die alleine keine qualitativen Aussagen zum Verfahren selbst und per se ermöglichten. Die Diskussion soll ja weitergehen. Also schließen wir uns dem auch an. Deswegen würden wir bitten, das als Ausblick in die Zukunft durchaus mit in die Methodenversion 4.2 aufzunehmen.

Wenn es erlaubt ist, ein kurzer Kommentar: Herr Danner bemüht auch andere Disziplinen in seinem Plädoyer. Da gibt es natürlich das Problem: Wenn Sie das ausdifferenzieren, kommen Sie irgendwann in die diskursive Ethik hinein. Die können Sie nicht quantifizieren. Das Institut hier hat Biometriker zuhauf, also genügend. Die Leute leben vom Zählen und vom Quantifizieren. Das heißt: Wenn Sie Patientenpräferenzen gewichten und erheben wollen, müssen Sie sie auch quantifizieren. Die Ethik ist keine Wissenschaft, die uns Quantifizierung bis dato ermöglicht - seit Aristoteles' Ethica Nicomachia zumindest nicht.

Stefan Lange: Über die ethischen Dimensionen will ich mich jetzt ungern auslassen, eben so wenig darüber, ob das quantifizierbar ist oder nicht und wovon wir leben. Das wissen wir schon selbst, Markos.

Es ist schon so, dass es auch in anderen Bereichen mal ganz schön wäre, wenn einmal von entsprechender Seite - sprich: gern auch vom vfa oder allen anderen Beteiligten in diesem Raum - konkrete Vorschläge gemacht würden. Es sollte nicht immer gesagt werden: Das IQWiG soll jetzt auslegen, darstellen und sich in was weiß ich nicht allem erschöpfen. - Wir haben das Ganze selbst angestoßen. Das sind Generalauftragprojekte gewesen, um dies

voranzubringen. Irgendwann ist aber Schluss. Irgendwann müssen auch mal die anderen wieder ran. Dann sagt uns doch: Was ist überhaupt Validität in diesem Zusammenhang? Ich habe keine Ahnung. Wie valide soll das dann sein? Wie erhoben? Wie reliabel soll das sein? Wie erhoben? Und wie repräsentativ soll das sein? Wie erhoben? Dann kann man doch mal einen Vorschlag machen, damit es überhaupt in Kosten-Nutzen- oder von mir aus auch Nutzenbewertungen eingesetzt werden kann.

Dann kann man sich mit dem konkret auseinandersetzen. Aber dieses allgemein Bla-bla, das IQWiG solle sich bitte schön da auslassen, ist in der Regel nicht so wahnsinnig hilfreich. Wir sind nicht alleine auf der Welt, sondern es gibt zahlreiche Forschungsinstitutionen in diesem Land, die, glaube ich, auch daran interessiert sein könnten, ebenso wie die Industrie, hier weiterzukommen. Das wäre einfach meine Bitte.

Moderator Jürgen Windeler: Herr Mühlbacher.

Axel Mühlbacher: Vielen Dank für das Sprungbrett. - Beide Studien - noch einmal - sind vor 3 Jahren abgegeben worden. In welchem Kontext reden wir mit Ihnen methodisch? In welchem Kontext geben wir Tipps ab? Mein Bereich bzw. meine Mitarbeiter und ich arbeiten seit 7 Jahren an Patientenpräferenzen. Ich glaube, wir haben hinreichend publiziert. Wir haben auch in Publikationen zu den Unterschieden zum Beispiel von AHP und DCE Stellung genommen. Ich wollte vorher nicht darauf eingehen, weil das vielleicht nicht genau der Platz war.

Noch einmal konkret: Wir werfen 2 Methoden in einen Topf. Die 2 Methoden wurden ursprünglich angedacht, dasselbe zu tun, und sie tun es einfach nicht. In diesem Methodenbericht werden sie immer noch in den gleichen Topf geworfen, obwohl die eine Methode 2000 einen Nobelpreis veranlasst hat für ein Paper, das 1974 geschrieben wurde. Durchaus ein Argument für die Ökonomen, den Nobelpreis so zu vergeben, war die Reliabilität. Die andere Methode, die AHP, ist ein mathematisches Verfahren. In den letzten 10 Jahren wurde hinreichend kritisiert, dass es einen mathematischen Lapsus gibt, der zum Beispiel zu einem Rankreversal führt - einfach nur mathematisch. Das ist in der Literatur beschrieben.

Seit 3 Jahren stehen wir da und diskutieren diese Dinge. Es ist nicht so, als würden wir nichts machen. Nur: Wo finden die Möglichkeit, genau diese Dinge zu diskutieren, sodass sie eingehen und dann im Methodenpapier sicherlich mit einer kritischen Linse - das würde ich Ihnen auch raten - aber stetig weiterdiskutiert werden? Denn die Ungeduld, wie ich vorher gesagt habe, kommt vielleicht daher, dass man sagt, in den letzten 3 Jahren hätten wir doch den einen oder anderen Schritt machen können.

Moderator Jürgen Windeler: Herr Lange.

Stefan Lange: Das ist jetzt mal ein Statement. Sie sagen: Das eine Verfahren - das habe ich so verstanden -, DCE ist zu favorisieren, weil das andere Verfahren ... Ich bin kein Experte

auf dem Gebiet, möglicherweise gibt es noch ein drittes oder viertes Verfahren; das kann man erstmal so stehenlassen.

Trotzdem bleibt die Frage: Wie reliabel muss das sein? Nennen Sie uns einfach mal eine Zahl. Wie gemessen? Cronbach's Alpha? Irgendwie? Wert 0,734? Wie valide soll es sein? Gemessen - keine Ahnung wie. Wie repräsentativ? Wen sollen wir denn fragen, Herr Mühlbacher? Wie wählen wir die Leute aus? - Nicht wir, sondern im Grunde genommen geht es ein Stück weit ... Das Ganze könnte in Dossiers für eine Kosten-Nutzen-Bewertung für die Aggregation von Endpunkten dienen. Auch die Industrie muss ein großes Interesse daran haben, Kriterien zu kennen, die eingehalten werden müssen. Ich glaube, Herr Gerber-Grote hat das schon sehr klar und eindrucksvoll gesagt: Gerade der Punkt der Repräsentativität wird ein Riesenproblem werden. - Da wünschte ich mir schon konkretere Vorstellungen.

Moderator Jürgen Windeler: Ich möchte das nur in einem Punkt ergänzen wollen. Ich nehme eine sehr deutliche Lücke zwischen Äußerungen wahr, die lauten - das betrifft jetzt den vfa, aber auch andere -: An diese Instrumente sind alle möglichen - sicherlich übrigens berechtigten - Forderungen zu stellen, von denen wir weitestgehend wissen, dass sie bisher nicht erfüllt sind - für viele Instrumente jedenfalls nicht, aber vielleicht für eins, das Sie besser kennen als ich, vielleicht doch - unter der gleichzeitigen Forderung, Patientenpräferenzen müssten aber unbedingt in Nutzenbewertungen Eingang finden und berücksichtigt werden.

Dieser Unterschied, dass sie berücksichtigt werden sollen, dass wir aber keine vernünftigen Daten haben, dass wir natürlich auch in Dossiers zum Beispiel keine Daten über Patientenpräferenzhebungen, die diese Kriterien wie „repräsentativ“ usw. erfüllen, finden, macht mich etwas stutzig. Das ist der Grund, warum das hier drauf steht. Wie - außer dem Hinweis, wie wir es geschrieben haben, sollte es nicht sein - soll es denn sein?

Axel Mühlbacher: Der erste Punkt war, zu sagen: Die Information ist in den Dossiers nicht enthalten. Es wäre eine Polemik zu sagen, dass, wenn sie enthalten sind, nicht gelesen wurden. Das ist aber festzustellen in einem der Protokolle.

Das Zweite ist: Ich möchte die Frage noch einmal wiederholen: In welchem Kontext unterhalten wir uns über die Methoden? Denn es gibt zum Beispiel die Unterscheidung zwischen der Frage: „Misst dieses Instrument verlässlich?“ und „Müssen diese Informationen bei der Bewertung berücksichtigt werden?“ Das sind zwei unterschiedliche Paar Schuhe. Das heißt: Wir können 1 Stunde darüber reden, wie gemessen wird. Dann können wir das Problem der Multikriterien- oder der multiplen Entscheidungsproblematik diskutieren. Daher kommt die Forderung: Bitte benutzt diese Information, um zu einer transparenten Entscheidungsfindung zu kommen in irgendeinem Prozess.

Kommen wir zum Messen. Natürlich: Reliabilität, Validität und die Repräsentativität - gerade die - werden uns auf Dauer beschäftigen, so, wie Sie es auf Dauer beschäftigen wird, wie Sie

das bei klinischen Studien machen, und so, wie wir die Probleme bei der Lebensqualitätsmessung auch haben. Wir werden hoffentlich, wenn die Studien nachgefragt werden, immer besser werden und diese Dinge tun können.

Der zweite Punkt ist - jetzt komme ich ganz bewusst auf den Punkt des Bewertens -: Werden diese Informationen nachgefragt? Das ist, wenn man den Methodenbericht 4.2 durchliest, an ein paar Punkten unsystematisch. Woher kommt denn die Evidenz darüber, was Patienten wollen? Woher kommt der Backup, zu behaupten, dass ein bestimmtes Kriterium, Zielkriterium oder ein bestimmter Endpunkt wichtiger als das oder der andere sei? Das kommt daher, dass im Bewertungsverfahren beim G-BA die Patienten am Tisch sitzen. Das hatten Sie vorher als Argument für die Patientenzufriedenheit angeführt. Auch da - wenn man da nicht sauber argumentiert - geht es um die Bewertung.

Ich sage Ihnen eines: Alles, was vorher nicht gemessen wurde, kann beim G-BA nicht bewertet werden. Insofern ist das Argument, dass die Zufriedenheit dann berücksichtigt wird innerhalb des Bewertungsprozesses beim G-BA, schlicht nicht zielführend, sondern wir müssen vorher schon hingehen und diese Dinge messen. Denn Sie sagen: In der Bewertung sollten sie berücksichtigt werden. Aber wenn wir durch Kapitel 3 und Kapitel 4 gehen, haben wir sogar Probleme zuzuordnen, wann, in welchem Kontext und wie diese Informationen benutzt werden. Allein eine Klarstellung in diesen beiden Kapiteln - zu sagen, dass diese Information im Rahmen des Dossiers dazu führt, dass die Gewichtung von Endpunkten überdacht wird - würde schon dazu führen, dass mehr Studien gemacht werden.

Moderator Jürgen Windeler: Herr Dintsios.

Ch.-Markos Dintsios: Ich greife einige Argumente auf, die von Ihnen, Herr Windeler, und von Stefan Lange gefallen sind, und weise darauf hin: Zum Beispiel im Zusammenhang der Erhebung von Patientenpräferenz - Instrument A springt an, Instrument B liefert andere Ergebnisse. Diese Probleme kennen wir auch von der gesundheitsbezogenen Lebensqualität. Das ist nichts Neues. Da haben wir auch den Fall gehabt - auch in der frühen Nutzenbewertung: Das eine Instrument war sensitiver, ist angesprungen. Das andere hat gar nichts gezeigt. Was macht man jetzt? Die Problematik, die sich so stellt, Stefan, methodisch, gilt doch auch für die gesundheitsbezogene Lebensqualität. Das ist doch nichts Neues für das IQWiG.

Zum zweiten Punkt, den Herr Windeler angeregt hat, Fertiglösungen in den entsprechenden Stellungnahmen zu liefern. Wir haben angeregt, dass diese Punkte hier aufgegriffen und in einem nächsten Schritt abdiskutiert werden, weil wir glauben: Vielleicht ist es jetzt noch zu früh. Das heißt doch eher: die Forschungsagenda fürs IQWiG; so war das angedacht. Wir können doch auf andere Literaturquellen hinweisen. Natürlich: Zur Discrete-Choice-Erhebung gibt es einiges im Bereich ISPOR. Das gibt es eine Taskforce; die hat sich damit befasst, eine Arbeitsgruppe. Sie hat auch Gütekriterien für diese Methoden aufgestellt. Aber ich dachte: Wir sind noch einen Schritt vorher.

Abschließend bin ich auch bei Herrn Mühlbacher. Natürlich, wenn man die Anreize nicht setzt, wird sich auch keiner damit im Gesundheitssystem direkt befassen wollen. Ich nenne ein Beispiel aus einer Bewertung eines Produktes. Das Produkt will ich nicht nennen. Da gab es die Unterteilung in zwei Subpopulationen. Für die eine war die Mortalität wichtig, und für die andere war das Ziel, Leben zu verlängern, nicht mehr wichtig. Beide bezogen sich bei der IQWiG-Bewertung - daraus hat das IQWiG auch die Empfehlungen abgeleitet, auf den Kardinalendpunkt Mortalität. Aber ich frage mich: Ist das nicht eine *Contradictio in adiecto*, wenn ich sage: Die zweite Subpopulation hatte nicht einmal als primäre Zielsetzung die Verlängerung des Lebens? Dann kapriziere ich mich wieder auf diesen Endpunkt Mortalität. Das sind doch die Fragestellungen, die die Patientenpräferenzhebung versucht, mit zu erörtern und mit zu lösen. Also müssten Sie vom IQWiG ein genuines Interesse daran haben.

Das verwundert uns wiederum umgekehrt, Herr Windeler. Ich kann das Argument genauso umdrehen und sagen: Sie begnügen sich mit einem semiquantitativen Vorgehen. Das macht Ihnen das Leben leichter. 9 Werturteile, glaube ich, hat Daniel Strech bei der IQWiG-Matrix festgestellt. Wenn es Sie glücklich macht - wieso nicht? In der Müller-Werbung stand: Wenn es schön macht - wieso nicht?

Moderator Jürgen Windeler: Nach mir kommen Andreas Gerber-Grote und Herr Danner dran.

Sie haben gesagt, dass die Bewertung und das Messen 2 unterschiedliche Dinge seien. Damit bin ich erst einmal, naheliegender Weise, sehr einverstanden. Im Moment nehme ich aber einen Widerspruch wahr. Voraussetzung für das Werten ist also das Messen.

(Axel Mühlbacher: Ja!)

Jetzt nehme ich einen Unterschied wahr, dass Sie sagen, es gibt Instrumente.

(Axel Mühlbacher: Ja!)

Und Herr Dintsios sagt, es gibt sie nicht. Er sprach gerade von Forschungsagenda und davon, das müsse erst noch erforscht werden.

(Ch.-Markos Dintsios: Richtig!)

Ist das ein Widerspruch, den ich nur wahrnehme, oder habe ich da etwas missverstanden, oder wie soll ich das verstehen? Denn noch einmal: Die Voraussetzung für die vielfältig geäußerten Forderungen, Patientenpräferenzen in die Bewertungen - mit „Bewertung“ meine ich jetzt Nutzenbewertung im Allgemeinen und nicht nur Werturteile - einzubeziehen, ist, dass man sie messen kann. Jetzt habe ich Sie so verstanden ... Ich habe kein Problem mehr, wenn mir jemand sagt: Es gibt auch Instrumente, die das können. - Super. Die haben wir noch nicht in unserem Methodenpapier stehen. Deswegen muss das geändert werden. Auch super.

Aber eine Äußerung, die heißt: „Patientenpräferenzen müssen dringend in die Bewertung einbezogen werden, aber es gibt die Instrumente nicht“, macht für mich wenig Sinn.

Andreas Gerber-Grote: Ich wollte nur kurz einfügen, weil es auch immer wieder darum ging: Haben wir überhaupt Gütekriterien genannt? Es gibt auch die ISPOR-Taskforce. Wenn man sich unsere Gütekriterien, die wir genannt haben - also eben nicht die, die hier laut Sozialempirie mit bestimmten Begriffen verfasst sind -, anschaut, muss man doch feststellen, dass viele diese Kriterien tatsächlich aus dem Dokument der ISPOR-Taskforce kommen. Das wollte ich hier nur noch ergänzen, falls es im Abgleich in der deutschen Sprachverfassung mit der englischen Sprachverfassung nicht deutlich geworden ist. Wir haben uns schon am internationalen Standard orientiert, der bei der ISPOR liegt, wenn die 10 Kriterien bei der Conjoint-Analyse gefordert werden, um eine Art Güte der Durchführung einer Conjoint-Analyse festzustellen.

Moderator Jürgen Windeler: Ich würde gern Herrn Mühlbacher hören, um zu versuchen, meine Unklarheiten zu sortieren.

Axel Mühlbacher: Die Antwort auf Ihre Fragen ist: Ja, ja, ja, nein, soweit ich die letzte Frage für mich beantworten kann.

(Heiterkeit)

Moderator Jürgen Windeler: Dann müssten Sie sie bitte zuordnen.

Axel Mühlbacher: Die letzte Frage war die Divergenz von Aussagen von Herrn Dintsios und mir. Von meiner Seite gibt es da eigentlich keine Divergenz.

Ich möchte aber gern auf den Punkt eingehen, dass die Voraussetzung der Berücksichtigung von Patientenpräferenzen deren Messbarkeit an sich ist. Ich wäre unheimlich froh, wenn wir die methodische Diskussion, die längere, auch wirklich führen. Aber - ich will das Pferd von hinten aufzäumen - das bei der Bewertung. Sie leben in der Annahme, dass, angenommen die Präferenzen wären nicht messbar oder mit einem zweifelhaften Instrument, man diese in der Bewertung nicht nutzen sollte. Das ist naheliegend, denn die Information könnte falsch sein.

Es gibt einen Punkt: In der Bewertung werden Sie auf jeden Fall Zielkriterien gewichten. Das heißt: Wir stehen vor der Wahl: Werden die Zielkriterien gewichtet durch ein Gremium von 20 Leuten, die implizit und intransparent bestimmte Annahmen tätigen und damit diese Annahmen in die Entscheidungsfindung eingehen lassen, oder steuern wir langsam und ordentlich in eine Situation hinein, in der wir genau diese impliziten und intransparenten Annahmen mit weiteren Informationen füttern? Diese weiteren Informationen - das ist die Forderung nach der Berücksichtigung von Patientenpräferenzen - wären dann Studien. Die Gefahr ist natürlich groß, die eine Methode zu verteidigen und die andere zu verdammen. Darum geht es nicht, sondern es geht darum, sich auf dem Weg weiterzuentwickeln und diese Studien als Information zu nutzen. Denn noch einmal: Bewertet wird, gewichtet wird schon

jetzt implizit und intransparent. Die Forderung nach der Berücksichtigung innerhalb dieses Bewertungsverfahrens stellt allein darauf ab, dass man sagt: Wir können uns Schritt für Schritt weiterbewegen. Lassen Sie uns diese implizite Gewichtung ein bisschen lüften, die Intransparenz beseitigen, und wir tun das unsere, die Methoden zu verbessern.

Moderator Jürgen Windeler: Jetzt hab ich Herrn Danner und dann Frau Wieseler und Herrn Dintsios.

Martin Danner: Ich will auch noch einmal an dieses Bonmot anknüpfen: Was nicht vorher gemessen wurde, darf nicht bewertet werden. - Letztlich ist es doch so, dass, wenn man sich die Nutzenbewertungen anschaut, verschiedene Endpunkte angeboten werden als diejenigen, die irgendeine Relevanz werden. Dann ist es doch eine Frage der Argumentationslast, ob derjenige, der diese Endpunkte anbietet, sagen kann: „Die treffen die Patientenpräferenzen und sind relevant“, oder ob derjenige, der das Bewertungsverfahren durchführt, darlegen muss, welches aus seiner Sicht die patientenrelevanten Endpunkte sind, die maßgeblich sein sollen. Das heißt mit anderen Worten: Man kommt an der Frage nicht vorbei zu sagen, dass man eine Bewertung braucht. Daher kann der Satz: „Es darf nur bewertet werden, was vorher gemessen wurde“ natürlich mit einem Fragezeichen versehen werden.

Wenn die Argumentationslast nämlich so ist, dass derjenige, der einreicht, diese Daten liefern muss, hätten Sie nie einen Nutznachweis erbringen können, weil Sie nie die Daten hätten liefern können, dass gerade die Endpunkte, die in den Zulassungsstudien eingeschlossen wurden, wirklich diejenigen sind, die die Patientenpräferenzen treffen. Insofern ist es eine normative Frage: Wer trägt die Argumentationslast?

Aber das führt aus meiner Sicht wiederum an dem Thema vorbei. Es ist gerade schon gesagt worden: Man muss eine Bewertung durchführen. - Ich würde auch zustimmen, dass es darauf ankommt, auf welcher Basis diese Bewertung erfolgt. Wenn es keine andere Erkenntnis gibt, muss man eben die Patienten, die beteiligt werden sollen, fragen. So findet es aktuell statt. Wir sind die Letzten, die sagen: Es könnten nicht noch weitere Instrumente hinzugezogen werden, wenn sie die Gütekriterien erfüllen, um Licht ins Dunkel zu bringen. - Insofern ist das sicherlich ein Prozess der Anreicherung. Aber das ist auf der anderen Seite sicherlich kein Prozess, bei dem man sagen kann, man kann jetzt noch nichts Vernünftiges tun, nur weil die Instrumente noch nicht hinreichend vorliegen.

Nochmals: Es ist auch nichts völlig Außergewöhnliches. Aus unserer Sicht ist die Nutzenbewertung, aber auch die Kosten-Nutzen-Bewertung von Arzneimitteln ein Unterfall von Technikfolgenabschätzung. Da gibt es umfassende Methodendiskussionen, die interdisziplinär laufen. Da kann ich nur noch mal dafür werben, sich da auch anschlussfähig zu halten.

Moderator Jürgen Windeler: Jetzt kommen Frau Wieseler, Herr Dintsios und Herr Mühlbacher dran.

Beate Wieseler: Ich möchte noch einmal auf den Punkt von Herrn Mühlbacher zurückkommen, dass diese Gewichtung, die da zweifelsohne vorgenommen wird, zu implizit und intransparent sei.

Ich denke, im Verfahren herrscht eine hohe Transparenz. Die Gewichtung wird nicht mit den hier beschriebenen Methoden gemacht. Aber ich denke, sie wird, soweit möglich, in diesen Verfahren transparent. Schon im IQWiG-Bericht ist es nicht so, dass wir mit einem Gesamturteil aus der Bewertung herauskommen, das nicht in seinen Einzelkomponenten und in seiner Gewichtung nachvollziehbar wäre.

Also: Wir stellen die vorhandene Evidenz pro Endpunkt dar. Wir beschreiben, warum wir zu einem Gesamturteil kommen und zu welchem wir kommen. Das ist eine Notwendigkeit für unsere Bewertung aus dem SGB V. Aber wir stellen das, wie gesagt, zunächst einmal nicht aggregiert, sondern einzeln dar und begründen unsere Abschätzung. Diese Darstellung einschließlich der Abschätzung geht in ein Stellungnahmeverfahren. Viele der hier Anwesenden sind an diesen Stellungnahmeverfahren regelhaft beteiligt. Da können auch alternative Gewichtungen und Abschätzungen vorgebracht und diskutiert werden. Dann ist es so, dass im Prozess beim G-BA diese disaggregierten Daten auch noch einmal angeschaut und gewichtet werden. Auch der G-BA begründet seine abschließende Einschätzung in den Dokumenten zu seinem Beschluss. Ich halte das Verfahren nicht für so intransparent, wie Sie es vielleicht dargestellt haben.

Moderator Jürgen Windeler: Herr Dintsios.

Ch.-Markos Dintsios: Vielleicht eine Richtigstellung: Herr Windeler, ich hatte nicht gesagt, es gebe kein Instrument. Ich hatte nur gesagt: Man muss sie weiterentwickeln. Da haben Sie mich zu konservativ interpretiert.

Zu den Punkten, die auch Frau Wieseler gerade angemerkt hat: Natürlich ist das Verfahren in seinen Folgen durchaus transparent. Nur: Gewichtungen wird es immer geben, wo Entscheidungen getroffen werden müssen. Das ist genauso mit der Rationierung. Das kann man entweder explizit oder implizit machen. Selbst die Vorgehensweise des IQWiG, anhand der entsprechenden oberen Konfidenzintervalle eine Aussage für die Klassifizierung zu treffen, hat implizit Werturteile. Das ist, glaube ich, hinlänglich bekannt. Das ist auch der Grund, wieso der Gemeinsame Bundesausschuss in seiner ersten Beschlussfassung zu Ticagrelor damals nicht der Empfehlung des IQWiG bzw. der Begründung dieser Empfehlung gefolgt ist. Man kann den Text lesen; er ist öffentlich zugänglich. Das ist die Fragestellung, mit der wir uns hier befassen müssen.

Mein Petitum - noch einmal - ist, nicht zu sagen: Wir haben das in der Forschungsagenda gehabt, wir haben uns damit befasst, und wir lassen es sein. Denn in den Stellungnahmen - zumindest in unserer Stellungnahme - wurde dezidiert genannt, dass Zulassungs- oberbehörden, die FDA und die EMA, sich mit diesen Fragen befassen, unter anderem mit

multikriteriellen entscheidungsanalytischen Verfahren, weil sie sich auch immer in ihrem Quasidunstkreis, was die Zulassung angeht, bewegen und das Verfahren transparenter machen wollen, indem sie partizipative Elemente mit einbringen wollen. Das ist letztlich auf der Agenda - nicht auf der Forschungsagenda allein, sondern auch auf der Agenda der Zulassungsoberbehörden sowohl in Übersee als auch in Europa.

Noch einmal abschließend mein Petition: Man möge sich damit weiter befassen und man möge vielleicht im Dialog versuchen, Gütekriterien mit zu definieren, insofern sie noch nicht existent sind. Sie sind in vielen Fällen auch nicht existent. Aber diese Problematik hatten wir genauso bei der Entwicklung des Konstrukts „Lebensqualität“ oder „Instrumente, die sie abgreifen sollen“. Das ist hier nichts Neues, um Himmels Willen.

Moderator Jürgen Windeler: Herr Mühlbacher.

Axel Mühlbacher: Vielen Dank. - Ich glaube, ich muss wirklich eines vorausschicken: Ich respektiere Ihre Arbeit, die Sie machen müssen, mit dem Dossier. Ich respektiere die Arbeit des G-BA. Ganz insbesondere bewundere ich die Mitarbeit von Patienten, die wirklich noch mehr als wir betroffen sind, die wir hier sitzen, und sich bei den Prozessen beteiligen.

Ich möchte nur eines ganz kurz betonen: Wir reden über Gütekriterien. Und Sie befragen Personen und Personengruppen. Welche Gütekriterien stellen Sie denn an die anekdotische Evidenz, die derzeit vorherrscht? Und: Inwieweit ist die von Intransparenz entfernt oder nicht? Das können wir irgendwann, ein anderes Mal diskutieren.

Es geht einfach darum, diese Evidenz, die vorherrscht und die nicht darum geht, klinische Effekte zu messen ... Sondern vielmehr sind diese klinischen Effekte zu überführen in ein Nutzenmaß, eine Nutzenbewertung oder ein Maß des Gesamtnutzens. Dazu braucht es die Gewichtung. Die Frage, die immer wieder kommen wird, ist die Frage: Woher kommt die Information für diese Gewichtungen? Die ist in den Dossiers derzeit nicht in der Form ausgeführt, dass sie die Gütekriterien, über die wir vorher geredet haben, befriedigen würde. Mehr will ich nicht sagen.

Wir haben Methoden, die genau diese Information beschaffen können - bis zu einem gewissen Grad -, wo wir hoffen, dass sie sich in Form der Gütekriterien weiterentwickeln. Es geht jetzt eigentlich nur um die Frage: Im Methodenbericht 4.2 wird diese Form der Information gewünscht. Kann man sie so im Methodenbericht verankern, dass sie auf Dauer Sinn macht?

Moderator Jürgen Windeler: Okay, danke. - Gibt es weitere Wortmeldungen? Ich schaue ganz nach rechts und ganz nach links, ob ich nichts übersehe. - Weitere Wortmeldungen sehe ich jetzt nicht. Dann würde ich mit dieser Form von Schlusswort diesen Tagesordnungspunkt abschließen.

Dann kommen wir zu Tagesordnungspunkt 3, zum direkten Nutzen diagnostischer Verfahren. Das ist so etwas wie ein Dauerbrenner. - Herr Sauerland

Tagesordnungspunkt 3

Direkter Nutzen diagnostischer Verfahren

Stefan Sauerland: Für das Methodenpapier wurden in der jetzigen Version die Kapitel zur diagnostischen Intervention sowie zu den diagnostisch-prognostischen Interventionen grundlegend umgestaltet. Ein zentraler Satz am Anfang von Kapitel 3.5 erklärt, dass die alleinige Gewinnung diagnostischer Informationen ohne therapeutische oder präventive Konsequenzen regelhaft keinen sozialrechtlich relevanten Nutzen habe. Hieran haben sich einige Stellungnehmende gestoßen, zum Beispiel die AWMF, die diagnostische Interventionen auch dann für sozialrechtlich erforderlich hält, wenn damit eine Diagnosesicherung und -klassifizierung insbesondere gemäß ICD erfolgen kann. Auch die DGHO sieht es für die Diagnosesicherung und -kodierung gemäß ICD als erforderlich an, dass diagnostische Interventionen durchgeführt werden, die keine therapeutischen Konsequenzen haben.

Uns interessiert, inwieweit es tatsächlich in der klinischen Praxis vorkommt, dass diagnostische Tests durchgeführt werden - allein mit dem Ziel, zum Beispiel eine ICD-Kodierung präzisieren zu können. Aus unserer Sicht kann nur die sozialrechtlich festgelegte Orientierung am individuellen Patientennutzen maßgeblich sein. Spätestens, wenn es um invasive diagnostische Maßnahmen geht oder eine Strahlenbelastung involviert ist, sprächen aus unserer Sicht auch ethische und strahlenschutzrechtliche Aspekte gegen eine Diagnostik ohne therapeutische Folgen. Daher die Frage: Vielleicht können uns die klinischen Kollegen diesbezüglich beruhigen und erklären, worin denn der Nutzen einer Diagnosesicherung oder einer präzisierten ICD-Klassifizierung besteht, die keine therapeutische Konsequenz hat.

Moderator Jürgen Windeler: Herr Wörmann.

Bernhard Wörmann: Das wurde auch von uns angemerkt. Ein typisches Beispiel ist ein Patient mit einer akuten myeloischen Leukämie, bei der nach WHO-Klassifikation heute eine molekulare oder genetische Differenzierung erforderlich ist - zum Beispiel mittelfristig oder auch jetzt schon für die Krebsregisterdokumentation.

Wenn wir das Beispiel akute promyelozytäre Leukämie nehmen, sage ich Ihnen: Es gibt heute 4 Verfahren, um diese Diagnose molekulargenetisch zu sichern. Man kann eine normale Zytogenetik, eine FISH-Analyse, eine Whole-Genome-Analyse oder eine PCR machen. Man kann sich ein Verfahren vorstellen; es wäre ganz sinnvoll zu schauen, welches dieser Verfahren das effektivste ist, um die richtige Diagnose zu stellen. Dann wird erst die klinische Therapie eingeleitet werden. Aber das Festlegen dieser Diagnose ist zunächst relevant. Da kann man sich Kosten-Nutzen-Verfahren vorstellen, in denen genau diese Frage gestellt wird. Deswegen haben wir das angemerkt.

Moderator Jürgen Windeler: Herr Lange.

Stefan Lange: Ich habe ein bisschen Zweifel, dass Sie auf den Punkt von Herrn Sauerland eingegangen sind. Noch einmal: Die Risikostratifizierung bei einer akuten myeloischen Leukämie oder überhaupt bei akuten Leukämien hat das Ziel, eine bessere Therapie für die Patienten zu finden. Also geht es natürlich um eine therapeutische oder überhaupt um eine Konsequenz - nicht unmittelbar, aber das ist das Ziel der Diagnostik. Sonst würden Sie sie ja nicht machen. Man macht keine Risikostratifizierung, wenn das keine Konsequenz hat. Wenn Sie den Patienten sowieso nichts anbieten können, brauchen Sie ihnen nicht zu sagen, dass sie den nächsten Tag sterben. Das macht nur traurig. - Aber das ist doch unbenommen so.

Wenn ich Sie jetzt richtig verstanden haben, sagen Sie: Manchmal haben wir noch gar nicht die entsprechenden Instrumente, auch therapeutischer Art, um diese Prognose gegebenenfalls zu verbessern. Aber es wäre hilfreich, diese Informationen jetzt schon in Krebsregistern oder sonst wo zur Verfügung zu haben, um überhaupt forschen zu können. Habe ich Sie da richtig verstanden?

Bernhard Wörmann: Nein. - Wir können einen Patienten mit akuter promyelozytärer Leukämie in 98 % der Fälle heilen. Das heißt: Wir haben die Instrumente schon. Aber die richtige Diagnose zu stellen, ist trotzdem wichtig. Der Weg zur Diagnose kann Gegenstand eines Verfahrens sein.

Wenn ich 4 verschiedene Verfahren habe, um zur richtigen Diagnose zu kommen, kann man die Methoden vergleichen. Dann kann ich ein Verfahren zur Kosten-Nutzen-Bewertung machen, um dann bei dieser Diagnose anzukommen. Egal, welches ich nehme - das hat keinen Einfluss auf die Überlebenszeit, weil ich alle Patienten dann gleichbehandle. Aber die Diagnose muss trotzdem erst gesichert sein.

Das heißt: Ich kann kein Verfahren haben, in dem ich Zytogenetik oder FISH vergleiche und schaue, wer am Ende überlebt. Aber die Diagnosesicherung muss schon am Anfang stehe.

Moderator Jürgen Windeler: Ich versuche eine Rückfrage: Das bedeutet doch übersetzt eigentlich, dass alle 4 Verfahren den gleichen Nutzen haben, weil sie dazu dienen, Patienten zu charakterisieren, damit eine bestimmte erfolgreiche Therapie gemacht werden kann, und dass es jetzt eigentlich nur noch darauf ankommt, von diesen 4 Verfahren mit gleichem Nutzen das - pointiert ausgedrückt - billigste oder das schnellste Verfahren herauszusuchen oder ein anderes Kriterium zu benutzen. Aber der Nutzen der Verfahren steht eigentlich nicht infrage.

Bernhard Wörmann: Es steht nicht infrage, wenn die Diagnose gestellt ist, wie der Patient dann behandelt wird. Aber wir können uns Verfahren vorstellen, in denen die erste Frage, die Sie gerade artikuliert haben, schon gestellt wird, oder habe ich das völlig falsch verstanden?

Es gibt schon Kosten-Nutzen-Erwägungen. Es können schon Verfahren da sein. Das heißt: Deswegen kann man diese Wertigkeit der reinen Diagnosestellung nicht einfach nicht aufnehmen. Deswegen denke ich, dass das mit hineingehört.

Unser Gefühl war, dass Sie die dominierende Fragestellung ... Das ist natürlich etwas, das wir gut nachvollziehen können. Dass die größte Menge von Verfahren dazu da ist, irgendwelche Parameter zu erheben, die potenziell nur prognostisch, aber nicht prädiktiv sind, ist ein großes Feld, bei dem wir extrem empfindlich sind. Da sind wir völlig auf Ihrer Seite. Trotzdem glauben wir, dass es eine vorgeschaltete, vielleicht auch zunehmend methodisch bedingte Situation gibt, in der die Diagnose gesichert werden muss, auch wenn es keinen unmittelbaren Einfluss auf die Therapie hat. Deswegen glaube ich, dass man bei der Diagnosesicherung trotzdem anfangen muss in dem Papier. Wir würden nicht direkt mit dem Nutzen angefangen haben.

Wenn Sie sagen: „Die Verfahren kann es nie geben bei Ihnen“, sehe ich das anders. Ich glaube, sie wird es geben.

Moderator Jürgen Windeler: Herr Lange und dann Herr Sauerbruch.

Stefan Lange: Also gut, ich glaube, dann gab es da ein beiderseitiges Missverständnis. Um die Rückfrage von Herrn Windeler zu präzisieren: Wir haben eine Situation, in der es im Grunde genommen ein klares Krankheitskonstrukt mit klaren Anweisungen gibt, wenn die Diagnose gestellt ist, wie zu handeln ist. Jetzt gibt es verschiedene Möglichkeiten, zu dieser Diagnose zu kommen. Die können Vor- und Nachteile haben.

Aber es ist im Grunde genommen im Methodenpapier als Möglichkeit vorgesehen, dass das eine Situation sein kann. Wir halten sie für vergleichsweise selten. Typischerweise ist das in der Vergangenheit anders gesehen worden, dass das eigentlich das alleinige Ziel sei, aber eigentlich die klare Situation, dass mit einem neuen Verfahren im Grunde genommen schon alles hinterher klar ist, ist wahrscheinlich tatsächlich - noch einmal - vergleichsweise selten.

Aber ich glaube: Mit dem, was Sie gerade gesagt haben, sind wir sehr d'accord. In unserem Methodenpapier ist eine solche Evaluationsmöglichkeit vorgesehen.

Moderator Jürgen Windeler: Herr Sauerbruch.

Tilmann Sauerbruch: Wir haben auch darauf hingewiesen. Sie schreiben auf Seite 64: „Die alleinige Gewinnung diagnostischer Informationen (ohne therapeutische oder präventive Konsequenzen) hat regelhaft keinen sozialrechtlich relevanten Nutzen.“ Dazu muss ich Ihnen sagen, dass Diagnostik durchaus Prognose machen kann, ohne dass das erfüllt ist - weder Therapie noch Prävention -, nämlich Einschätzung des Patienten. Insofern ist dieser Satz nicht korrekt. Diese Einschätzung des Patienten brauchen Sie durchaus. Dafür brauchen Sie auch Diagnostik. Es kann also gut sein, dass daraus hervorgeht, dass Sie keine Therapie und keine Prävention machen.

(Zuruf: Beispiel!)

Moderator Jürgen Windeler: Herr Sauerland.

Stefan Sauerland: Ich denke, diesen Punkt hat Herr Wörmann gerade schon sehr schön ausdifferenziert, indem er gesagt hat, dass uns interessieren muss, was an der prognostischen Information tatsächlich prädiktive Information ist, was also die Therapie leiten kann.

Das reine Gewinnen von Informationen zur Information des Patienten, der dann vielleicht sein Leben besser planen kann oder erwartbar Entscheidungen beispielsweise zu Familie, Hauskauf oder ähnlichen Dingen besser organisieren kann, ist sozialrechtlich nicht als „medizinische Konsequenzen“ zu fassen. Es ist klar, dass es da durchaus einen Schnittbereich gibt, wo diagnostische oder prognostische Informationen für den Patienten in seinem Privatleben Bedeutung haben, aber das ist nicht das, was im Sozialgesetzbuch miterfasst werden soll.

Tilmann Sauerbruch: Darf ich darauf antworten?

Moderator Jürgen Windeler: Klar.

Tilmann Sauerbruch: Es ist von ganz entscheidender Relevanz in der Behandlung des Patienten. Wenn Sie sagen, es gebe keine Therapie und keine Prävention, müssen Sie doch ein diagnostisches Kriterium in der Hand haben, um das sagen zu können. Warum wollen Sie das hier verbannen? Das ist medizinisch nicht einzusehen.

Moderator Jürgen Windeler: Herr Lange.

Stefan Lange: Zur Präzisierung: Es geht nicht darum, dass es zwangsläufig eine Therapie geben muss, sondern es geht darum, dass man in Kenntnis dieser Information bessere Entscheidungen treffen kann. Wo wir nur differenzieren müssen, ist: bessere medizinische Entscheidungen. Das ist sozialrechtlich relevant.

Was Herr Sauerland gerade versucht hat - das ist nicht ganz unheikel -, pointiert auszudrücken: Bessere individuelle Entscheidungen des Patienten im Hinblick zum Beispiel auf - das zu sagen, ist ganz extrem - den Abschluss einer Lebensversicherung können eigentlich nicht sozialrechtlich relevant sein. Wir haben schon natürlich den Fokus auf medizinische Entscheidungen, die übrigens auch lauten können: Es ist besser für den Patienten zum Beispiel im Hinblick auf Lebensqualität, auf eine bestimmte Therapie zu verzichten. - Aber es geht immer um eine Entscheidung. Wenn keine Entscheidung ansteht - ich glaube, das ist trivial -, kann Diagnose keinen Nutzen für den Patienten haben.

Moderator Jürgen Windeler: Herr Kunz, Herr Mühlbacher.

Hans-Detlev Kunz: Herr Lange, Sie sprechen von medizinischen Entscheidungen. Was tatsächlich passiert, entscheidet der Patient. Der trifft die Entscheidung. Nicht der Arzt oder das Diagnostikum. Die Entscheidung, nichts zu tun, ist auch eine therapeutische Konsequenz. Die trifft dann auch der Patient. Dafür braucht er aber eine ziemlich gesicherte Diagnostik. Nichts zu tun, ist eine therapeutische Entscheidung, die sich aus dem Ergebnis der Diagnostik

ableitet. Daraus entscheidet dann der Patient, wenn er noch einwilligungsfähig ist, was zu tun ist oder was nicht zu tun ist.

Stefan Lange: Herr Kunz, da sind wir völlig einer Meinung.

(Andreas Reimann: Ich habe es nicht verstanden: Dann sind Sie völlig anderer oder einer Meinung?)

Moderator Jürgen Windeler: Einer Meinung. - Herr Mühlbacher, bitte.

Axel Mühlbacher: Ich habe eine Frage an Herrn Sauerland: Angenommen, es ließe sich wissenschaftlich valide nachweisen, dass die Information, ohne dass sie einen Therapieerfolg oder überhaupt eine Therapie nach sich zöge, den Patientennutzen steigert, würden Sie dann bei Ihrer Aussage hinsichtlich des Sozialgesetzbuches bleiben?

Stefan Sauerland: Wir haben uns schon einmal mit diesem Thema relativ detailliert beschäftigt und geschaut, ob es Studien gibt, in denen tatsächlich das Durchführen einer diagnostischen Intervention, ohne dass sich irgendeine Therapie anschloss, dazu geführt hat, dass sich Patienten besser fühlen. In der Regel wird es dann nicht um die Endpunkte Mortalität und Morbidität gehen, sondern es wird um den Endpunkt gesundheitsbezogene Lebensqualität gehen. Da gibt es einzelne Studien, wenn ich mich richtig erinnere, die dabei ein kleines bisschen sehen, aber die Mehrzahl der Studien sieht dort nichts. Letztlich ist aber entscheidend, dass hier die 3 Endpunkte Mortalität, Morbidität oder Lebensqualität beeinflusst würden. Wenn es tatsächlich so wäre, dass man mit einer diagnostischen Intervention Lebensqualität positiv beeinflussen könnte, könnte ich mir auch vorstellen, dass man dann sagen müsste: Dann ist das sozialrechtlich ein Punkt, der gewertet werden müsste.

Moderator Jürgen Windeler: Herr Lange.

Stefan Lange: Selbstverständlich. Der Punkt - deshalb ist es vielleicht auch im Methodenpapier so pointiert gemacht worden -, ist, dass wir davon ausgehen - dafür gibt es in der Tat zahlreiche Evidenz -, dass das Vermitteln von prognostischer oder diagnostischer Information einen Effekt hat, typischerweise einen negativen Effekt. Es ist also erst einmal ein Schaden.

Umgekehrt ist es natürlich auch vorstellbar, dass es einen positiven Effekt hat. Nur dieser Effekt - darauf wollen wir eben hinweisen - kann nicht nur allein durch Plausibilität, Annahmen oder sonst etwas, sondern muss durch geeignete Studien nachgewiesen werden.

In der Tat ist es so, wie Herr Sauerland sagt: Es gibt eine Übersicht zu der Frage, ob Patienten, die beunruhigt sind und eher zur Hypochondrie neigen, durch eine ständige Diagnostik irgendwie beruhigt werden können, indem ihnen gesagt wird: Nein, da ist nichts. - Das ist das, was gemeinhin angenommen wird: Wenn jemand beunruhigt ist, ich eine Diagnostik mache und eine Erkrankung ausschließe, sollte diese Person beruhigt sein. - Nein,

das ist nicht der Fall. Dieser eigentlich als plausibel angenommene positive Effekt existiert nicht.

Umgekehrt gibt es - noch einmal - zahlreiche Situationen nachgewiesener Art - auch durch höchste Evidenz im Evidenzkörper -, dass diagnostische oder prognostische Informationen erst einmal schadet, weil sie Leuten erst einmal schlechte Laune macht.

Moderator Jürgen Windeler: Jetzt Herr Dintsios.

Ch.-Markos Dintsios: Wir haben auch dazu Stellung genommen. Ich mache es mal ein bisschen plakativ, weil ich sagen muss: Wir haben uns ein bisschen vor den Kopf gestoßen gefühlt.

Erstens ist das IQWiG ein Auftragsinstitut. Primär erledigt es 90 % seiner Arbeit für den Gemeinsamen Bundesausschuss. Das heißt: Wenn so ein Auftrag käme, würde der Gemeinsame Bundesausschuss als Walter über das Sozialrecht und als sozusagen unfehlbarer Anwender des SGB V - selten wird er vors Bundessozialgericht zitiert und verliert dort - schon wissen, was er tut. Deswegen sehe ich nicht die Notwendigkeit dieser Pointierung im Methodenpapier Version 4.2.

Das Zweite ist - wenn ich Stefan Lange richtig folge; ich möchte gern korrigiert werden, falls ich es falsch verstanden habe -: Wir hatten eine Zeit, in der man HIV diagnostizieren konnte. Therapien gab es nicht. Ich erinnere an das Jahr 1987, glaube ich; man möge mich korrigieren. Ich glaube, 1985 ist der Virus in Frankreich und den USA isoliert worden. 1987 hat man die PCR-Tests schon entwickelt. Die hochaktive antiretrovirale Therapie kam 1992, glaube ich. Ich bin ein bisschen älteres Semester; man möge mich gern korrigieren. Hieße das dann, dass wir, um zum Beispiel die Krankheitslast für Gesundheitssysteme zu ermitteln, keine HIV-Testungen als diagnostische Maßnahme machen sollten, weil wir zum damaligen Zeitpunkt die Alternative hatten, den Leuten zu sagen: Ihr habt HIV und werdet an AIDS versterben? Interpretiere ich das richtig?

Das Dritte ist: Ich lebe, obwohl ich vom Balkan komme, in Zeiten der Aufklärung und erinnere gern an Jean-Jacques Rousseau: Information hat einen Eigenwert an sich. Den würde ich mir nicht vom IQWiG, auch nicht vom Sozialrecht nehmen lassen wollen. Es gibt eine Implikation in § 12. Es ist nicht die Lebensversicherung. Das ist eher ein amerikanisches Problem, ob die Menschen das in voller Kenntnis des Risikos, wenn sie eine Lebensversicherung abschließen, verheimlichen. Das haben wir hier in Zentraleuropa weniger ausgeprägt. Es gibt ein anderes Problem, nämlich die Lebensplanung. Wenn ich anhand eines Gentests weiß, dass ich eine Generkrankung entwickeln werde, mache ich meine Familienplanung anders. Vielleicht bekomme ich keine Kinder. Also gibt es dann, falls ich später Interventionen auf meine Kinder geben könnte, weniger Kosten für das System. Ich kann es so oder so herleiten. Ich muss sagen: Ich bin verblüfft von dieser Argumentation - bis heute noch.

Moderator Jürgen Windeler: Herr Schinzel als Nächster.

Stefan Schinzel: Ich habe eine Frage, die ich nicht so ganz verstehe, und zwar: Diagnostik sieht häufig so aus, dass man keine hundertprozentige Entscheidung oder Antwort bekommt. Tests haben meistens eine gewisse Sensitivität und eine gewisse Spezifität. Ich kann unter Umständen durch Hinzunahme eines weiteren diagnostischen Tests Sensitivität und/oder Spezifität steigern. Wo würden Sie denn da die Grenze ziehen, wo Sie sagen: Das macht keinen Sinn mehr, in der Frage eine sichere Entscheidung zu treffen für eine weitere Therapie?

Moderator Jürgen Windeler: Ich sammle noch einmal Wortmeldungen. Das sind in der Reihenfolge Herr Sauerbruch, Herr Reimann, Herr Köbberling und Herr Danner. Dann kommt Stefan Sauerland dran. - Herr Sauerbruch, bitte.

Tilmann Sauerbruch: Ich komme noch einmal auf den von mir eingangs zitierten Satz zurück: „Die alleinige Gewinnung diagnostischer Informationen (ohne therapeutische oder präventive Konsequenzen) hat regelhaft keinen sozialrechtlich relevanten Nutzen.“ Woher wissen Sie das denn eigentlich vorher? Das frage ich rein logisch.

Moderator Jürgen Windeler: Herr Reimann.

Andreas Reimann: Ich hätte auch diese Frage. Gerade bei den seltenen Erkrankungen haben wir natürlich die Situation, dass es eine Symptomatik gibt, dass die Patienten bzw. ihre Angehörigen, wenn die Patienten nicht einwilligungsfähig sind, darunter leiden dass man nicht weiß, um was es geht, und dass man zumindest eine Antwort sucht, damit dann nicht nur die Lebensplanung, sondern möglicherweise gesundheitsbezogene Konsequenzen gezogen werden. Beispielsweise kann ein bestimmtes Verhalten vermieden oder durchaus gesucht werden.

Das heißt: Ist es eigentlich vorstellbar, dass es eine Situation im Falle einer unklaren Diagnose gibt, aus der gar keine Entscheidung folgt? Sie haben vorhin, Herr Windeler - deswegen habe ich extra noch einmal nachgefragt - gesagt: Sie sind sich einig, wenn eine Entscheidung daraus folgt. Diese Entscheidung kann Nichtbehandlung bzw. Nichtintervention oder Prävention sind, wobei HIV ein sehr eindrückliches Beispiel war. Das kann aber auch andere Zusammenhänge haben, etwa besondere Risikofaktoren, die man vermeidet. Das kann aber eben auch Intervention bedeuten. Ist es denkbar, dass es eine Diagnose gibt, die überhaupt gar keine Konsequenz nach sich zieht?

Deswegen glaube ich: Wenn man die Frage verneint, wird man zu dem kommen, was Sie sagen, dass, wenn eine informierte Entscheidung des Patienten möglich ist, eine Diagnose durchgeführt werden sollte.

Moderator Jürgen Windeler: Die nächste Wortmeldung von Herrn Köbberling.

Johannes Köbberling: Es gibt das Moment, dass Patienten sich häufig viel besser fühlen, wenn nur irgendeine Diagnostik gemacht wurde. Das ist in der Tat in der klinischen Erfahrung gar nicht so selten, aber das bringt trotzdem große Probleme mit sich. Klassisches Beispiel, das immer spaßeshalber zitiert wird: Der Patient sagt: Herr Doktor, seitdem das EKG gemacht wurde, geht es mir schon viel besser.

Nun könnte jemand auf die Idee kommen: Dann machen wir bei solchen Patienten ein EKG aus suggestiven oder Placebogründen - wie man das auch immer nennen will -, dann haben wir die ganze Diskussion um die Sinnhaftigkeit des therapeutischen Placebos. - Das wollen wir hier nicht machen.

Aber wir haben noch ein zusätzliches Problem: Wir bekommen dann - wenn wir uns ein EKG machen, müssen wir uns das anschauen - eine Information für eine nicht vorher dagewesene Fragestellung. Wir hatten gar keine Fragestellung an das EKG, wir machen das aus einem ganz anderen Grund. Das ist, wie wir alle wissen, wieder ein großes Problem: Was machen wir mit Befunden, an die wir keine A-priori-Fragestellung hatten? Wir kommen in ganz viele Probleme hinein, wenn wir uns darauf einlassen, Diagnostik aus anderen Gründen zu machen, als eine therapeutisch relevante Information zu bekommen.

Moderator Jürgen Windeler: Herr Danner.

Martin Danner: Zwei Punkte. Das eine ist: Aktuell sind diese Tests während der Schwangerschaft zur frühzeitigen Feststellung der Trisomie 21 in der Diskussion und auch in der Bewertung beim Gemeinsamen Bundesausschuss. Da ist für mich die Frage: Würde das IQWiG sagen: So etwas lehnen wir ab, denn ein Schwangerschaftsabbruch kann nicht die therapeutische Option sein? Wie geht man damit um?

Das zweite ist eher eine Anmerkung. Aus unserer Erfahrung ist es weniger die Diskussion zu fragen: Gibt es eine therapeutische Option, wenn man eine Diagnostik durchführt? Vielmehr ist die Frage: Welcher Nachweis ist dann an diese therapeutische Konsequenz zu stellen? Wir haben bei diagnostischen Verfahren häufig die Situation, dass man natürlich viel danach machen kann und eine Vielfalt von Patienten hat. Ich erinnere an die Diskussion zur Messung des Sauerstoffs im Blut bei Neugeborenen, um Herzfehler festzustellen. Da kann man sich danach viele verschiedene Operationen vorstellen usw. Dabei die Evidenz nachzuweisen, dass, wenn der Test in der Welt wäre, der Outcome schließlich bei den Kindern und den späteren Erwachsenen besser wäre, ist sozusagen eine sehr hohe Anforderung. Deswegen noch einmal die Nachfrage, was dann unter dem Thema ohne therapeutische Konsequenz zu verstehen ist.

Moderator Jürgen Windeler: Herr Sauerland hat jetzt 6 oder 7 Fragen zu beantworten. Danach kommt Herr Stengel dran.

Stefan Sauerland: Ich fange bei Ihnen an, Herr Danner, weil das gerade am frischesten im Kopf ist. Die Pulsoxymetrie bei Neugeborenen bewerten wir gerade. Da ist durchaus

erkennbar, dass es Studien gibt, die sich tatsächlich auch mit dieser therapeutischen Konsequenz beschäftigen. Denn wenn man eine Pulsoxymetrie dann nicht durchführt, könnten tatsächlich Herzfehler übersehen werden. Daraus würden sich natürlich therapeutische Konsequenzen in dem Sinne ergeben, dass keine Therapie erfolgt, wo eigentlich eine Therapie notwendig wäre und dann die Kinder Schaden nähmen.

Beim anderen Punkt, den Sie ansprechen, die Trisomietests und der Endpunkt Schwangerschaftsabbruch, glaube ich: Diesen Punkt zu diskutieren, würde ein paar Stunden brauchen. Aber es ist durchaus ein Endpunkt vorstellbar, die Belastung bzw. die Lebensqualität der Eltern durch Geburt eines Kindes mit einer Fehl- oder Missbildung. - Das nur dazu.

Zum Argument, das Herr Dintsios genannt hat - HIV und was haben wir eigentlich zu tun mit der Krankheitslast -: Das IQWiG hat nicht die Aufgabe, die Krankheitslast zu definieren. Ich kann mir aber sehr gut vorstellen, dass es im Bereich HIV auch in der Zeit, in der es keine verfügbare Therapie gab, durchaus auch Alternativdiagnosen und Differentialdiagnosen gab, und dass es durchaus Sinn machte, eine HIV-Testung durchzuführen, weil dann eben ein anderer Teil der Patienten, nämlich der, der glücklicherweise kein HIV hatte, dann einer Therapie zugeführt werden konnte.

Jetzt zu Herrn Sauerbruch: Woher wissen wir denn vorher, ob der Patient eine therapeutische Konsequenz hat? - Nein, das wissen wir natürlich nicht. Das ist immer nur im Gruppenbezug sinnvoll. Wenn wir eine Studie haben, wird natürlich nicht erwartet, dass sich bei allen Patienten eine therapeutische Konsequenz ergibt, aber zumindest doch bei einem gewissen Teil der Patienten. Dieser Anteil sollte ein gewisses Ausmaß haben.

Das führt mir zur Frage von Herrn Schinzel, wann denn hier eine noch weitere Steigerung von Sensitivität oder Spezifität auch noch als Zusatzinformation messbar und statistisch nachweisbar sei. - Selbstverständlich, wenn ich eine Sensitivität von 98 % auf 99 % anhebe, ist das vielleicht nicht mehr ganz so relevant. Aber wenn ich es nachweisen kann und wenn es Studien dazu gibt, dann ist vermutlich - das hängt vom Endpunkt und den Konsequenzen ab, die sich in der Therapie ergeben - auch problemlos vorstellbar, dass sich eine solch kleine Änderung in der Testgüte in einem Nutzen niederschlägt.

Dann noch zur Frage von Herrn Reimann: Ich hatte das so verstanden, dass Sie die Verhaltensänderung oder das Vermeiden eines gewissen Verhaltens nicht als therapeutische Konsequenz ansehen. Ich sähe da durchaus weniger Probleme.

(Andreas Reimann: Entschuldigung!)

Moderator Jürgen Windeler: Vielleicht stellen Sie das eben richtig.

Andreas Reimann: Dann habe ich mich unklar ausgedrückt, ganz im Gegenteil: Eine Verhaltensänderung kann eine therapeutische Konsequenz sein. Es kann auch eine Kon-

sequenz sein, nicht zu intervenieren oder eben zu intervenieren. Das sind die Konsequenzen, die passieren können.

Ich habe gefragt: Ist es vorstellbar, dass es eine Diagnostik bei einer Situation unklarer Symptomatik gibt? Um die geht es jetzt, nicht etwa um EKG, damit der Patient, der Hypochonder beruhigt ist; da sollen wir schon die Fälle auseinanderhalten; das ist nun etwas völlig anderes. Ein Patient, der eine klare somatisch-nachvollziehbare Symptomatik hat, und nicht einer Diagnose sicher zugeordnet werden kann - das ist ein typischer Fall bei seltenen Erkrankungen. Ist eine Situation vorstellbar, in der eine Diagnose in solchen Fällen keine therapeutische Konsequenz hat in dem von mir beschriebenen Sinne: Intervention, Nichtintervention, Veränderung von Verhalten?

Stefan Sauerland: Ich kann mir durchaus vorstellen, dass es Erkrankungen gibt - gerade auch im Bereich der seltenen Erkrankungen -, für die es derzeit keine kausale Therapie gibt und auch keine, die im Sinne einer Verhaltensänderungen und Verhaltensvermeidung als Therapie gewertet werden könnte.

Aber die Situation, die Sie beschreiben, dass ein Patient mit einem eher unklaren Krankheitsbild kommt, würde dafür sprechen, dass man eine ganze Reihe von Differentialdiagnosen in Betracht ziehen muss. In einem solchen Fall kann ich mir nicht vorstellen, dass alle der infrage kommenden Differentialdiagnosen keine therapeutischen Konsequenzen haben.

Wenn dann doch zumindest eine dieser möglichen Diagnosen eine Therapie nach sich zöge, wäre unser Kriterium schon längst erfüllt. Dann würden sich therapeutische Konsequenzen ergeben. Dann würde man natürlich sinnvollerweise einen solchen Patienten mit einer unklaren somatischen Symptomatik einer Diagnostik unterziehen.

Bevor wir die nächste Runde aufmachen, zu Herrn Köbberling: Diese Idee, dass die Diagnostik einen Placeboeffekt hat, ist durchaus sehr wichtig, im Kopf zu haben. Ich denke auch, dass wir uns, ähnlich wie es in ähnlichen Bereichen bereits passiert, auch im Bereich der Evaluation diagnostischer Intervention durchaus irgendwann damit beschäftigen müssen, dass wir Placebodiagnostik als Komparator in Betracht ziehen. Das wird sicherlich sehr selten sein, aber ich kann mir so etwas vorstellen. Gerade im Bereich der EKG-Durchführung gibt es schon Studien, die zwar schon 30 Jahre zurückliegen, aber an denen man gesehen hat, dass ein EKG durchaus gewisse, aber doch sehr überschaubare Wirkungen hat. Das hängt eben sehr vom Patienten ab.

(Johannes Köbberling: Da sind wir bei der Diskussion um die therapeutische Intervention, nicht um die diagnostische!)

- Ja!)

Moderator Jürgen Windeler: Herr Reimann, ich würde gern auch noch einmal eine Antwort auf Ihre Frage versuchen, wie ich sie verstehe. Wir glauben: Es ist in der Tat eben nicht vorstellbar, dass einer diagnostischen Intervention keine neuen Entscheidungen folgen. Deswegen überrascht uns die Diskussion, die wir jetzt auch gerade wieder führen. Denn der Anspruch ist häufig der, dass Anbieter, Hersteller, Protagonisten von diagnostischen Tests meinen, die Qualität des Tests erübrige sich in Sensitivität und Spezifität. Nein, das ist es eben nicht. Es kommt auf die Konsequenzen an.

Ich glaube auch - ich habe das nicht systematisch untersucht -: Es gibt keinen einzigen Anbieter und keinen einzigen Protagonisten eines diagnostischen Tests, der sich damit begnügen würde zu sagen: Da kann man eine Krankheit genau feststellen. - Vielmehr hat jeder die explizite und die Patienten sogar wahrscheinlich auch die Erwartung - implizit ist das, glaube ich, immer dabei -, dass damit irgendetwas anders gemacht werden kann, am besten natürlich besser. Genau für die Frage, ob mit der diagnostischen Information etwas besser gemacht werden kann, interessieren wir uns. Ob das Therapie, Intervention, Prävention, Verhaltensänderungen oder Nichtstun ist - uns kommt es darauf an, dass - das ist der einzige Punkt, der vielleicht ein bisschen spezifischer ist, dass das medizinische - sagen wir etwas allgemeiner: gesundheitliche - Konsequenzen haben sollte.

Andreas Reimann: Dazu kann auch die Vermeidung weiteren Diagnoseaufwands gehören. Gerade, weil Sie sie zu Recht angesprochen haben: Es gibt Erkrankungen mit einer Reihe von Differentialdiagnosen, die abgeklärt werden müssen. Der Patient sollte nicht von Arzt zu Arzt und von Diagnostik zu Diagnostik laufen - nicht nur, weil das natürlich sicherlich nicht seiner Lebensqualität zuträglich ist, sondern auch, weil das ressourcenverbrauchend ist - möglicherweise abhängig von der Art der Diagnostik sogar schädlich. Daher: Mit dieser Äußerung, Herr Windeler, kann ich leben, weil Sie das noch einmal klargestellt haben.

Moderator Jürgen Windeler: Herr Lange und dann Herr Stengel.

(Dirk Stengel: Hat sich erledigt!)

- Okay, dann nur Herr Lange.

Stefan Lange: Dann würde ich das gern in der Tat noch einmal als Antwort auf die Frage von Herrn Danner präzisieren - bezogen auf Schwangerschaftstests. Natürlich ist das eine medizinische Konsequenz. Noch einmal: Das Wortspiel um „therapeutisch“ oder „medizinisch“ könnten wir zum Anlass nehmen, um das in der Tat so weit zu präzisieren. Es geht aber tatsächlich um medizinische Konsequenzen. Das ist in erster Linie, wo wir glauben, dass man den patientenrelevanten Nutzen im Sinne des Sozialgesetzbuches auch bei der Diagnostik festmachen muss.

Ich glaube, beim zweiten Beispiel, das Sie genannt haben, Herr Danner, geht es um eine Screeningsituation. Screening ist das Beispiel par excellence, wo, glaube ich, praktisch keine Zweifel mehr bestehen, dass man irgendeine Form von Evidenz haben muss, damit es

hinterher vernünftige Konsequenzen gibt, die in der Tat mit einem Benefit für die Betroffenen verknüpft sind. Insofern sind wir da sehr auf einer Linie. Wir haben uns erlaubt, diesen Gedanken des Screenings ein Stück weit auch auf die „mehr oder weniger Normaldiagnostik“ zu erweitern.

Last but not least - das ergibt sich, glaube ich, von selbst -: Lieber Markos, vielleicht warst du noch zu jung, um die Diskussion um die HIV-Diagnostik mitzuerleben. Über die Frage, ob das überhaupt erlaubt ist, gab es ernsthafte Fragen: Darf man bei Menschen einen HIV-Test machen, weil es gerade keine Konsequenzen gibt? Die gibt es immer noch. Das ist im Grunde genommen genau das Beispiel, das es sehr klar macht, wo die Problematik beim Vermitteln einer Information ohne Konsequenz ist.

Moderator Jürgen Windeler: Jetzt gibt es noch zwei Wortmeldungen: einmal hier und dann Herr Kunz und Herr Sauerbruch.

Sven Klebs: Wir haben jetzt schon relativ lang über mögliche Fälle diskutiert, bei denen dieser Satz, den Sie eingefügt haben, nicht greifen würde, bei denen am Ende keine therapeutische Konsequenz, medizinische Entscheidung oder irgendetwas anderes dahintersteht.

Wir sind nicht wirklich auf irgendein Beispiel gekommen - zumindest habe ich keines wahrgenommen -, bei dem dieser Punkt wirklich greift. Bei allen Beispielen, die wir diskutiert haben, haben wir am Ende explizit oder implizit hergeleitet, dass es doch dann nicht unter diese Kategorie fällt, dass am Ende keine therapeutische Konsequenz, kein präventiver Ansatz oder Ähnliches damit verbunden wäre. Welche konkreten Beispiele hatten Sie denn im Kopf, bei denen Sie sagen: Die sind aus unserer Sicht diejenigen, die wir sozialrechtlich nicht sehen? Das würde ich gern ein bisschen besser verstehen. Denn bisher waren alle Beispiele immer gegensätzlich.

Moderator Jürgen Windeler: Herr Lange, direkte Antwort?

Stefan Lange: Die Frage, wenn ich Sie richtig verstanden habe ... In der Tat, genau. Wir haben über Beispiele diskutiert, bei denen es genau darum geht: Gibt es eine medizinische Konsequenz, und gibt es Evidenz dafür, dass diese Entscheidungsfindung positiv durch die Information beeinflusst werden kann? Aber es geht um medizinische Konsequenzen; darüber sollten wir uns einig sein.

Jetzt fragen Sie nach Beispielen, bei denen das nicht gegeben ist. Da gibt es Hunderttausende. Herr Wörmann war sich mit uns einig, dass es gegenwärtig und schon seit Jahren eine Flut von prognostischen Biomarkern in der Wissenschaft gibt, die auch teilweise in der Tat in der Praxis angewendet werden, um prognostische Stratifizierungen vorzunehmen. Da gibt es zahlreiche Tests, wo eben die Ergebnisse, die dann als Konsequenz resultieren, zum Beispiel therapeutische Konsequenz, bei denen, die angeblich ein hohes oder niedriges Risiko haben,

exakt die gleichen sind. Also brauche ich diesen Test nicht. Solche Dinge gibt es en masse. Darüber kann ich Ihnen gern Literatur zur Verfügung stellen.

Noch einmal die Frage und die Antwort an Herrn Schinzel: Ich bin nicht sicher, ob wir die Frage richtig beantwortet haben. Wenn ein Test behauptet, dass er eine höhere Sensitivität habe, also mehr Fälle entdeckt als irgendetwas, das schon da ist, muss einen das immer hellhörig machen. Denn dann muss man sich fragen, ob das denn wirklich die Fälle sind, für die wir wissen, dass wir sie mit einem bestimmten Arzneimittel, einer bestimmten Maßnahme oder einer bestimmten präventiven Maßnahme wirklich gut versorgen können. Typischerweise sind das zusätzliche Fälle, bei denen wir das gar nicht wissen und bei denen wir erst einmal wissen müssen: Gibt es denn überhaupt für diese Fälle eine vernünftige Konsequenz?

Das Gleiche gilt übrigens auch für den Ausschluss von Patienten. Wenn wir Patienten sagen: Wir entdecken im PET irgendwelche leuchtenden Punkte, deswegen werden Sie von einer vorher etwas prognostisch günstigeren in eine ganz ungünstige Situation verbannt, dann meint man plötzlich, man könne auf jegliche Therapie verzichten. Wir wissen gar nicht, ob das wirklich gut für die Patienten ist. Das ist unsere Forderung, dass man im Grunde genommen Evidenz dafür benötigt, dass man wirklich zu besseren Entscheidungen kommt. Die diagnostische Information und die diagnostische Güte allein im Sinne von höherer Sensitivität und Spezifität sind eben oft nicht handlungsleitend in diesem Sinne.

Moderator Jürgen Windeler: Stefan Lange hat gerade den Begriff „zusätzliche Patienten“ benutzt. Ich will nur pointieren: Es geht nicht um mehr, sondern um andere Patienten dabei. Das kann das Problem sein. - Jetzt kommt Herr Kunz dran.

Hans-Detlev Kunz: Hier wird der Begriff „therapeutische Konsequenz“ abstrakt benutzt. Muss dann diese therapeutische Konsequenz einen evident nachgewiesenen Nutzen haben? Denn wir haben leider einige Diagnostiken, und wir haben sehr viele therapeutischen Konsequenzen. Aber diese Konsequenzen sind evident nicht belegt.

(Stefan Lange: Das ist ein Problem!)

Dann sprechen wir aber von einem sogenannten zusammengesetzten Nutzen, der aber evident nicht belegt ist. Wie will man eigentlich aus dieser Situation heraus? Dann dürfte man ja überhaupt keine Diagnostik mehr machen, weil die therapeutischen Konsequenzen, die zur Verfügung stehen, evident nicht als nutzbringend belegt sind.

Moderator Jürgen Windeler: Ich sammle erst einmal wieder. - Herr Sauerbruch.

Tilmann Sauerbruch: Wir hatten eine Frage bezüglich der Methodik auf Seite 67. Wenn Sie einen neuen diagnostischen Test prüfen, wer bestimmt dann, was der Goldstandard ist, gegen den Sie das werten?

Moderator Jürgen Windeler: Dann würde ich noch die Wortmeldung von Herrn Mühlbacher, wenn er bereit ist, mit dazu nehmen.

Axel Mühlbacher: Herr Lange, da sind wir an dem Punkt und sagen: Informationen können schädlich sein. Ich stimme Ihnen voll zu.

Wo ich Ihnen nicht zustimme, das ist Ihr Nutzenbegriff. Der Nutzenbegriff ist der, den Sie im Glossar selbst „enger Nutzenbegriff“ nennen. Sie trennen ihn ab, was Sie von Ihrem Glossar im Internet „weiter Nutzenbegriff“ nennen. Ich will das ausführen. Nehmen wir das Beispiel Alzheimer. Ein Therapieerfolg ist nicht garantiert. Wir gehen davon aus: Es gibt keinen Therapieerfolg. Angenommen, wir gehen davon aus, es gäbe eine Möglichkeit, die Plaques zu identifizieren. Das heißt: Wir haben eine Information darüber vorliegen, ob jemand Alzheimer hat. Wir wissen aber: Diese Information führt zu keinem kurativen Erfolg und zu keiner Therapie. Jetzt wäre es interessant zu wissen, ob es eine bestimmte Population gibt, die diesem Wissen einen Nutzen oder sogar einen Negativnutzen zuspricht.

Angenommen wir könnten identifizieren, dass es eine bestimmte Population gibt, die aus dieser Information einen Patientennutzen zieht, wäre das der weite Nutzenbegriff. Das ist der Nutzenbegriff, den jede Person an diesem Tisch eigentlich normalerweise in ihrem Leben hat. Nach Ihrem engen Nutzenbegriff würden wir über 100 % der Population keinen Nutzen sehen. Wenn wir aber den weiten Nutzenbegriff anwenden, würden wir einen Nutzen bei den Personen schaffen, die diese Informationen haben wollen und auch bereit wären, dafür etwas zu tun. Es gibt auch eine Population, für die das nicht gilt. Wir haben die Studie - ich kann sie Ihnen gern geben -, wo man genau das differenzieren kann. Denn das wäre ein Beispiel für eine Präferenzstudie, um darzustellen, dass es Leute gibt, die es nicht wissen wollen. Bei denen können wir davon ausgehen: Da entsteht ein Schaden. Aber es gibt auch eine vielleicht sogar geringere Population, die sagt: Ich will es wissen.

Stefan Lange: Ich finde das ein wunderbares Beispiel, aber es macht deutlich, dass wir vergleichende Studien benötigen - auch für diese Frage. Wir müssen diese Effekte - „hat irgendetwas Positives“ oder „hat irgendetwas Negatives“ - quantifizieren können. Das geht nur durch vergleichende Studien. Das geht nicht durch Sensitivität und Spezifität, sondern das geht durch vergleichende Studien. Das ist der eigentliche Punkt.

Das Problem ist - Sie haben völlig Recht -: Es gibt welche, bei denen es einen Nutzen geben könnte, und welche, bei denen es einen Schaden, einen Nachteil oder einen Vorteil geben könnte. Das Problem ..., das haben wir eigentlich immer bei Interventionen. Auch bei Arzneimitteln gibt es Patienten, die von der Gabe gar nicht profitieren, die nur Nebenwirkungen und Nachteile haben. Es gibt auch andere, die nur Positives haben. Wir müssen dies immer irgendwie zueinander in Beziehung setzen. Das Einzige, was wir versuchen, ein bisschen auf den Punkt zu bringen, ist: Auch diagnostische Information bedarf einer vergleichenden Evaluation.

Moderator Jürgen Windeler: Stefan Sauerland.

Stefan Sauerland: Dann beantworte ich jetzt die beiden Fragen von Herrn Kunz und von Herrn Sauerbruch.

Herr Kunz, wenn die therapeutische Konsequenz nicht belegt, also unsicher ist - es gibt zwar eine therapeutische Konsequenz in der Praxis, aber wir wissen gar nicht, ob das dem Patienten etwas nutzt und guttut -, ist natürlich in unserer gemeinsamen Betrachtung von Diagnose und Therapie das gesamte Bündel in seinem Nutzen nicht belegt. Dann würden wir tatsächlich auch ein Problem sehen, selbstverständlich. Aber das ist eben auch ein Vorteil für diese gemeinsame Betrachtung. Diesen Vorteil gibt es dann, wenn man eine Studie hat, wo beides gleichzeitig getestet wird. Wenn ich eine Studie habe, in der Patienten diagnostisch untersucht werden und direkt anschließend therapiert werden, kann ich zwei Fliegen mit einer Klappe schlagen, indem ich hier einen Nutzen finde. Dann ist klar, dass sowohl die Diagnostik erforderlich ist, als auch diese Therapie nutzbringend ist. Das ist der Idealfall.

Aber wir haben eben eine Kopplung. Diese Kopplung spricht auch Herr Sauerbruch an, wenn er fragt, was wir denn als Goldstandard verwenden. Zum einen: Wir verwenden meistens den Begriff „Referenzstandard“, denn einen echten Goldstandard gibt es in vielen Bereichen nicht. Da müsste man die Histologen bemühen. Dafür müsste man den Patienten aufschneiden oder ähnliche Dinge machen. Wir haben meistens Referenzstandards. Diese sind in diesen Fällen oft inkomplett oder zusammengesetzt. Das heißt: Ein Teil der Patienten wird vielleicht nur nachuntersucht. Bei einem Teil der Patienten kann man vielleicht eine Biopsie machen.

Aber auch hier: Diese diagnostische Güte ist nur ein Hilfskriterium. Wir brauchen eigentlich anschließend an das Stellen einer Diagnose Evidenz, die zeigt, dass diese Patienten mit dieser Diagnose von einer Therapie profitieren. Daraus können Sie ableiten, was wir denn für Ansprüche an einen Referenzstandard stellen. Er muss eigentlich zu dem passen, was wir in den Therapiestudien finden, oder was in der Medizin allgemein als diagnostisches Kriterium angesetzt wird, um hierüber eine Therapie begründen zu können. Die Kopplung ist das Entscheidende.

Moderator Jürgen Windeler: Herr Sauerbruch noch einmal.

Tilmann Sauerbruch: Wenn Sie einen neuen diagnostischen Test prüfen, müssen Sie auch wissen: Gegen was prüfen Sie ihn? Ist er besser oder schlechter? Das wollen Sie ja wissen. Das muss doch jemand bestimmen.

Stefan Sauerland: Wir haben 2 verschiedene Situationen. Entweder gibt es noch überhaupt keinen diagnostischen Test,

(Tilmann Sauerbruch: Ja!)

dann wird der diagnostische Test gegen Nichts getestet oder, wenn ich den Ideen von Herrn Köbberling folge, vielleicht sogar gegen eine Placebointervention. Die zweite Situation ist deutlich häufiger: Ein diagnostischer Test wird gegen einen anderen diagnostischen Test verglichen, etwa PET/CT versus Kernspintomografie. In diesen Fällen würden wir uns einen Referenzstandard wünschen, der tatsächlich den Erkrankungsstatus möglichst adäquat abbildet. Wir haben in den Bewertungen - gerade jetzt in der Onkologie - häufig die Situation, dass man häufig Biopsien machen oder nachbeobachten kann. Das alles ist relativ unproblematisch. Ich verstehe nicht ganz Ihren Punkt.

Tilmann Sauerbruch: Noch einmal: Wer bestimmt, welches Ihr Referenztest ist? Irgendjemand muss sich doch dafür entscheiden.

(Bernhard Wörmann: In Studien wird das doch gemacht!)

Stefan Sauerland: Im Endeffekt müssen wir das machen.

(Tilmann Sauerbruch: Bei Ihnen im IQWiG!)

- Ja, selbstverständlich. Wenn Sie der Meinung sind, dass wir uns bei diesen Entscheidungen zu einem Referenzstandard oder gegen einen Referenzstandard falsch entschieden hätten, müssten Sie tatsächlich beim Kommentieren eines Berichtsplans sagen: Liebes IQWiG, da habt ihr euch aber vertan.

Moderator Jürgen Windeler: Ich sehe keine weiteren Wortmeldungen. - Dann schließe ich jetzt diesen Tagesordnungspunkt. Ich schlage vor, dass wir uns um 13:40 Uhr, also in einer halben Stunde, wieder treffen.

(Pause: 13:10 Uhr bis 13:40 Uhr)

Moderator Jürgen Windeler: Ich begrüße Sie wieder. Ich hoffe, Sie haben sich stärken und austauschen können. Ich hoffe, Sie konnten die Diskussion noch etwas fortführen und sich auf den zweiten Teil vorbereiten.

Wir kommen jetzt zum Tagesordnungspunkt 4.

Tagesordnungspunkt 4

Darstellung der Evidenz in Potenzialbewertungen

Moderator Jürgen Windeler: Ich will vorausschicken, dass wir weniger zum neuen Abschnitt „Potenzialbewertungen“ - das auch -, aber generell zu Potenzialbewertungen ungefähr so viele Eingaben, Kommentare und teilweise Fragen wie zu Tagesordnungspunkt 1 bekommen haben. Das waren die beiden gravierendsten einzelnen Punkte bei den Stellungnahmethemen.

Ich will allerdings auch gleich dazusagen, dass die meisten Punkte, die uns in Sachen „Potenzialbewertungen“ erreicht haben und die in den Stellungnahmen zu lesen waren, weder das IQWiG noch Entscheidungen betreffen, die das IQWiG treffen kann. Das betraf etwa die Frage der Fristen. Das betraf die Frage der Beteiligungsmöglichkeiten. Das betraf die Frage der Vertraulichkeit. Das alles sind Fragen, zu denen ich im Moment nur sagen kann, dass sich die Stellungnehmenden weniger an den G-BA als vielmehr an den Gesetzgeber wenden müssen, weil wesentliche Punkte im Gesetz festgelegt und von uns nicht zu beeinflussen sind. Das bedeutet auch, dass Beteiligungsmöglichkeiten - das betrifft sowohl externe Sachverständige als auch irgendeine Form von Stellungnahmeverfahren in diesen Potenzialbewertungen -im IQWiG nicht möglich sind.

Ich kann auch dazusagen, dass das IQWiG bestimmte Aspekte, etwa der Grad der Vertraulichkeit bei Unterlagen, auch nicht begrüßt, zumal das in gewissem Gegensatz zu allen sonstigen Verfahren und allen sonstigen Dingen, die das IQWiG vertritt, steht. Aber auch das ist von uns nicht zu beeinflussen. Also Appell an diejenigen, die meinen, das irgendwie ändern zu wollen: Da gibt es andere Adressaten.

Wir haben einen Punkt bei den Potenzialbewertungen dann doch als so unklar angesehen, dass wir ihn hier ansprechen und thematisieren wollen. Dazu gebe ich das Wort an Herrn Sauerland.

Stefan Sauerland: Danke. - Mit dem neuen Kapitel 3.8 zur Potenzialbewertung sind gerade im Vergleich zur Nutzenbewertung ganz neue Maßstäbe erforderlich geworden. Das bedeutet, dass hierbei nun erstmals auch andere Grade der Ergebnissicherheit auftauchen. So werden nichtrandomisierte Vergleichsstudien mit methodischen Schwächen hier als „Studien sehr geringer Ergebnissicherheit“ einsortiert. Eine Stufe darunter finden sich „Studien minimaler Ergebnissicherheit“, wobei hierunter primär nichtvergleichende Studien, also einfache Fallserien, zu verstehen sind. Mehrere Stellungnehmende, also DGOU, Sanofi-Aventis, BVMed, BPI und Weitere kritisieren die vom IQWiG gewählten Adjektive „sehr gering“ und „minimal“ als sprachliche Herabwürdigung der bestverfügbaren Evidenz im Bereich nichtmedikamentöser Verfahren.

Uns kann es hier und heute weniger darum gehen, schönere Adjektive zu suchen, uns interessiert vielmehr, warum die Stellungnehmenden offenbar der Auffassung sind, dass

Fallserien regelhaft mehr als minimale Ergebnissicherheit bieten. Anders herum könnte man auch fragen: Wenn nicht Fallserien minimale Ergebnissicherheit bieten, gibt es denn dann unter dieser Evidenzstufe weitere Quellen, die man für die Ableitung eines Potenzials heranziehen sollte oder müsste?

Dirk Stengel: Es geht tatsächlich gar nicht darum, dass man sagen müsste: Fallserien bieten regelhaft in irgendeiner Weise eine bessere Evidenz als zum Beispiel eine vergleichende Studie, um Gottes Willen. Das einzige, was uns ein bisschen aufgefallen ist und uns vielleicht auch ein bisschen stört, ist, dass derzeit immer noch Termini gebraucht werden, von denen ich mindestens glaube, dass sie derzeit in der methodischen Gemeinschaft ein bisschen anders gehandhabt werden.

Was zum Beispiel gar nicht Verwendung findet - das sage ich nur als Beispiel -, ist das GRADE-Instrument für die Bewertung von Verzerrungen, vom Bias, vom Outcome von Effektstärken usw. Das heißt: Hierbei sehen wir immer noch eine gewisse Verhaftung in der üblichen Evidenzhierarchie.

Ich glaube: Wir sind uns alle einig, dass das Wesentliche an einer wissenschaftlichen Untersuchung die methodische Qualität ist. Die ist erst einmal möglicherweise auch unabhängig vom eigentlichen Studiendesign zu sehen.

Moderator Jürgen Windeler: Ja, bitte.

Stefan Sauerland: Direkt dazu: Das Verharren an einer Evidenzhierarchie, bei der bestimmte Studientypen bestimmte Aussagesicherheiten zugeordnet sind - gerade in diesem nicht-randomisierten Bereich -, läge uns, ehrlich gesagt, ziemlich fern.

Im GRADE-Ansatz sind einige Ideen dazu vorhanden, wie man Studien auf- oder abwerten kann. Auch da gibt es solche Begriffe wie „critical validity“, also wirklich schwere Probleme bei der Validität oder Aussagekraft von Studien.

Das andere, was ich vielleicht erwähnen möchte, auch weil es relativ aktuell ist, ist, dass sich die Cochrane Collaboration derzeit recht intensiv mit einer Bewertung nichtrandomisierter Studien beschäftigt. Wir sind da auch in Kooperation mit den Kollegen gegangen und haben ein Tool pilotieren dürfen, das tatsächlich genau diese Ansprüche erfüllt, eine evidenzstufenübergreifende Bewertung von Einzelstudien vorzunehmen, um dann einen Gesamtevidenzkörper, um dieses schöne Wort noch einmal zu bringen, zu gewährleisten.

Dirk Stengel: Nur kurz als Kommentar: Ich weiß das, denn ich komme gerade vom Cochrane Colloquium zurück. Ich kenne das Instrument. Wir haben auch lange darüber diskutiert.

Moderator Jürgen Windeler: Herr Winkler.

Olaf Winkler: Erst einmal herzlichen Dank für die Einladung zur Anhörung. Bevor ich tiefer auf die methodische Fragestellung eingehe, müssen wir uns noch einmal klarwerden, dass wir hier vom Gesetzgeber durch das Versorgungsstrukturgesetz einen neuen Rechtsbegriff eingeführt bekommen haben, was auch in Ihrer methodischen Einleitung entsprechend gewürdigt wird.

Hier geht es um ein niedrigschwelligeres Niveau der Darstellung der Validität der Aussagen, die entsprechend für methodische Bewertungen herangezogen werden. Das ist vom Gesetzgeber klar in der Gesetzesbegründung dargelegt worden. Jetzt besteht die Kunst darin, dies mit Leben zu füllen.

Als wir Ihren Vorschlag erstmalig gesehen haben, haben sich sehr viele Fragezeichen für uns dargestellt. Denn letztlich muss man ein Methodenpapier deswegen erstellen, damit diejenigen, die davon Gebrauch machen - in diesem Falle sind es eventuell Hersteller, die Medizinprodukte in Verkehr bringen und eventuell den Antrag auf ein Potenzial stellen -, genau referenzieren und prüfen können: Passt mein Verfahren bzw. meine Methode in das Verfahren? Wie muss ich das Dossier erstellen? Welche Fakten muss ich liefern?

Sie haben sich die Mühe gemacht, eine Klassifikation der Evidenzgüte darzustellen. Uns sind leider nicht die Referenzierung - warum sie auf diese Stufe gekommen ist - und die Ausgestaltung erkenntlich. Das heißt also: Die Klarheit der Bestimmtheit der Aussagekraft wäre vielleicht in einem weiteren Prozess zu optimieren.

Ich muss ganz klar sagen: Auch uns fällt es schwer, derzeit methodische Instrumente zu finden, um dies zu qualifizieren. Das muss sicherlich aus dem lernenden System kommen. Unser Wunsch wäre es, mehr Klarheit hineinzubringen, und, was das Wording der Adjektive angeht: Wenn man unvoreingenommen ist und das außerhalb dieser Methode liest, gibt es eine gewisse subjektive Empfindsamkeit. Vielleicht kann man diesen Potenzialbegriff auch nutzen, das mehr als Möglichkeit zu sehen, tatsächlich im Rahmen der Potenzialbewertung später eine Nutzenbewertung folgen zu lassen. - So weit mein erster Kommentar.

Stefan Sauerland: Die Fragezeichen, die Sie hatten, als Sie den Entwurf zum Versorgungsstrukturgesetz sahen, hatten wir sicherlich auch. Es wäre vermessen, wenn wir sagen würden: All diese neuen Aspekte, die sich durch die Potenzialbewertungen aufgetan haben, hätten wir schon im Hinterkopf gehabt.

Insofern ist das eine Methodik, die erst einmal ein Vorschlag ist, die sich in nächsten Jahren erweitern und präzisieren kann. Aber mir wäre bezugnehmend auf die Ausführungen von Herrn Lange heute Vormittag wichtig zu fragen, ob es konkrete Punkte gibt, bei denen Sie sagen: Das muss man ganz präzise anders regeln. Da wissen wir genau, dass hier irgendein Studientyp falsch einklassifiziert bzw. höher zu werten ist oder dass irgendein Effekt geringer zu werten ist. Wir sind immer dankbar für Ideen, die unsere Methodik verfeinern können. Die allgemeinen Bitten: „Schreibt das doch genauer hin“ oder „Erklärt das doch noch einmal“,

helfen relativ wenig. Es gibt auch durchaus Stellungnehmende, die geradezu davor gewarnt haben, dass man vonseiten des IQWiG mit Festlegungen, die vielleicht noch relativ unbegründet sind, gleich Pflöcke in die Erde schlägt, die man am Ende nicht wieder entfernen kann.

Moderator Jürgen Windeler: Guido Skipka hatte sich gemeldet, dann Frau Soskuty.

Guido Skipka: Vielleicht eine kleine Randbemerkung zu den Adjektiven, die vielleicht besser oder weniger gut gewählt sind: Das dürfte den meisten klar sein - manchmal geht es in der Diskussion ein bisschen durcheinander: Diese Adjektive beziehen sich nicht auf die Studie, nicht auf die Qualität und sogar nicht auf die Durchführung einer Studie. Das kann alles wunderbar gelaufen sein, etwa eine Fallkontrollstudie nach allen Regeln der Kunst.

Die Adjektive beziehen sich bei uns allein auf die Ergebnissicherheit für unsere Fragestellung. Das ist die Nutzenbewertung. Das heißt, eine Studie kann bei uns eine sehr geringe Ergebnissicherheit haben, obwohl es eine wunderbare Studie ist, die sehr gut durchgeführt ist und für andere Fragestellungen auch extrem sichere Ergebnisse liefert.

Meinen zweiten Punkt habe ich vergessen; vielleicht komme ich gleich noch dazu.

Moderator Jürgen Windeler: Einmal kurze Denkpause. - Frau Soskuty.

Gabriela Soskuty: Ganz kurz zu diesen Begriffen: Es geht nicht darum, dass wir sagen, es müssten schöne oder passendere Worte gefunden werden, sondern uns ist nicht ganz klar: Wie messen Sie denn „sehr gering“ oder „mäßig“? Was heißt das denn ganz genau? Das ist uns ein bisschen zu unbestimmt.

Moderator Jürgen Windeler: Herr Lange.

Stefan Lange: Wie würden Sie es denn machen, Frau Soskuty?

Gabriela Soskuty: Dazu kann ich wenig sagen, da ich selbst weder Methodikerin noch Gesundheitsökonomin bin. Ich bin der Meinung, man müsste Wege finden, die Studien so einzuteilen ... Gegen die Einteilung der Studien ist gar nichts zu sagen, wie Sie sie gruppiert haben. Mich stört einfach nur die Frage: Wie messen Sie denn „gering“ oder „sehr gering“? Es ist nicht, dass ich es anders machen will. Ich will einfach nur eine klare Aussage: Was heißt denn „gering“?

Moderator Jürgen Windeler: Herr Skipka versucht eine Antwort.

Guido Skipka: Ich würde mich nicht so sehr an den Begriffen „gering“, „sehr gering“ oder „minimal“ festhalten. Entscheidend ist, dass wir hier auf der Ebene der nichtrandomisierten Studien 3 gestufte Kategorien gebildet haben, um uns bei der Potenzialbewertung die Ergebnissicherheit zu quantifizieren.

Was genau „gering“ heißt, haben wir versucht, abzubilden. Wir halten Ergebnisse für sicher in geringem Maßstab, wenn es nichtrandomisierte Studien sind, die aber zumindest prospektiv geplant sind und beispielsweise eine Confounder-Kontrolle haben. Dann würden wir erst einmal vom Studientyp her sagen: Diese Ergebnisse haben eine geringe Ergebnissicherheit.

Wenn Sie genauer lesen, steht etwas weiter, dass diese Einschätzung dann noch von weiteren Faktoren abhängen kann. Da verweisen wir auf unseren allgemeinen Abschnitt zum Verzerrungspotenzial. Da orientieren wir uns wie schon seit Jahren natürlich an den Systemen, die man beispielsweise bei Cochrane kennt, etwa Risk-of-Bias-Tool. Wir schauen uns zum Beispiel die ITT-Problematik in den Studien oder weitere Dinge an. Das kann natürlich dazu führen, dass man in der Ergebnissicherheitseinstufung herauf- oder auch herabgeht.

Moderator Jürgen Windeler: Herr Schinzel hat sich gemeldet.

Stefan Schinzel: Was uns beim Lesen ein bisschen irritiert hat, ist der Begriff „minimal“. Wenn man das mit der Evidenzklassifizierung vergleicht, wie sie in der aktuellen Verfahrensordnung des Gemeinsamen Bundesausschusses steht, ist das Kategorie IV. Es gibt aber noch darunter Kategorie V, also Assoziationsbeobachtungen oder Einzelfallberichte. Dass das natürlich eine sehr niedrige Evidenz ist, ist klar. Aber „minimal“ war aus unserer Sicht die unterste Stufe, die es überhaupt gibt. Daher haben wir uns darüber schon gewundert.

Moderator Jürgen Windeler: Jetzt hat noch einmal Herr Stengel das Wort, dann Herr Sauerland.

Dirk Stengel: Ich möchte empfehlen ... Ich glaube, wir sind uns sehr nahe, denn ich glaube: Die einzelnen Begriffe scheinen sowohl im Kontext als auch im Inhalt vorzukommen. Ich würde nur darauf hinweisen, dass man vielleicht das GRADE-Instrument noch ein bisschen spezifischer herausstellt, weil es gerade in den letzten 2 Jahren einen enormen Aufschwung gibt und es zumindest international die wichtigste Bewertungsbasis ist - sowohl für die Cochrane Collaboration aber auch für eine Institution wie McMaster usw. Ich würde es nur dringend empfehlen. Es gibt da auch ganz wunderbare Tabellen, an denen man sich orientieren kann und die man möglicherweise auch parallel hier einfach einmal abbilden kann, um zu zeigen: Es gibt einen entsprechenden Konsens - sowohl inhaltlich als auch bei der Definition.

Moderator Jürgen Windeler: Stefan Sauerland.

Stefan Sauerland: Ob GRADE ein Konsens ist, der weltweit so akzeptiert ist, will ich einmal dahingestellt sein lassen, weil GRADE durchaus auch einen relativ hohen Aufwand erfordert. Frau Nothacker kann vielleicht etwas dazu sagen, ob sich das in der AWMF so verbreitet. Das müssen wir ein bisschen abwarten. Insofern sehe ich da durchaus auch Alternativen, zum Beispiel dieses neue Bewertungsinstrument, das von der Cochrane Collaboration vorgelegt wurde.

Auch klar ist, dass die Bewertung von nichtrandomisierten Studien deutlich aufwendiger und deutlich komplizierter ist. Das bedeutet, dass dieses klare Ableiten von bestimmten Ergebnissicherheiten aus einer Studie oder aus mehreren Studien natürlich eine subjektive Komponente beinhaltet. Wir werden uns bemühen, die subjektive Komponente so gering wie möglich zu halten. Aber wie das immer so ist bei Reliabilitätsmessungen: Es wird sich sicherlich nicht völlig vermeiden lassen, dass darin subjektive Elemente verbleiben.

Zu den Ausführungen von Herrn Schinzel: Vielleicht kann man das Problem um den Begriff „minimal“ so lösen, dass man sagt: Er bezieht sich minimal auf die Verfügbarkeit klinischer Evidenz. Denn was sich in der G-BA-Verfahrensordnung darunter befindet, sind in dem Sinne keine Studien mehr, sondern zum Teil Erfahrungsberichte, Expertenmeinungen, tier-experimentelle Dinge, pathophysiologische Rationale und Ähnliches.

Da würde ich allein schon wegen der Abgrenzung zwischen klinischer Evidenz und nichtklinischer Evidenz irgendwo eine Grenze sehen. Dann könnte man sagen: Dann ist in der klinischen Evidenz die Fallserie oder vielleicht einmal ein Fallbericht tatsächlich die minimale Ergebnissicherheitsstufe.

Moderator Jürgen Windeler: Herr Lange.

Stefan Lange: Man kann es trotzdem noch einmal auf den Punkt bringen. Es ist im Grunde genommen Common Sense, dass man in der Regel - das ist ganz entscheidend - für einen Nutznachweis beispielsweise Vorstellungen hat, wie solche Studien auszusehen haben. In der Regel werden das eben randomisierte Studien mit patientenrelevanten Endpunkten sein.

Die spannende Frage ist: Was stellt man sich als Regelfall als Mindestvoraussetzung für ein Potenzial vor? Es ist wirklich nicht hilfreich zu sagen: „Wir müssen uns den gesamten Evidenzkörper anschauen“, sondern da ist bitte schön irgendwo eine Linie einzuziehen. Wir glauben zum Beispiel, dass reine Überlegungen, gedankliche Konstrukte über den Einsatz eines Verfahrens nicht ausreichend sind, ein Potenzial abzuleiten. Darüber kann man natürlich diskutieren. Aber das ist eigentlich der Sinn und Zweck heute, ob man nicht ein Stück weit solche - ich glaube, der Begriff passt schon ganz gut - Minimalanforderungen für den Regelfall formulieren kann. Wir halten das für sinnvoll - übrigens auch um der Transparenz und Nachvollziehbarkeit willen.

Moderator Jürgen Windeler: Herr Sauerbruch, bitte.

Tilmann Sauerbruch: Wir hatten in diesem Kontext, aber auch in anderen Kontexten noch die Frage: Inwiefern werden Fachgesellschaften in die Potenzialbewertung einbezogen?

Moderator Jürgen Windeler: Den Punkt haben in der Tat eine ganze Reihe von Stellungnehmenden - nicht nur Fachgesellschaften, sondern auch Firmen und Verbände formuliert. Noch einmal: Das Verfahren zeichnet sich durch 2 Spezifika aus - ich rede jetzt ausdrücklich darüber, was das IQWiG angeht -: einen sehr hohen Grad an Vertraulichkeit,

was bedeutet, dass wir nicht nur wie bei den Dossierbewertungen die Unterlagen zum Beispiel Fachexperten nicht zur Verfügung stellen können. Das bedeutet auch, dass wir über Art, Umfang und Zahl der Aufträge nicht reden können.

Da kommt also eine Beteiligung von Fachgesellschaften schlicht nicht infrage, um es ganz klar und eindeutig zu sagen. Das ist nicht unsere Entscheidung, um das auch noch einmal klar zu sagen, sondern die Entscheidung des Gesetzgebers oder, soweit das irgendwie noch modifizierbar war, des G-BA.

Der zweite Punkt sind die sehr engen Fristen. Wir haben 6 Wochen für eine Potenzialbewertung Zeit. Diese 6 Wochen - das können Sie sich, glaube ich, vorstellen - reichen gerade eben für die üblichen und notwendigen Abläufe im Institut. Da wäre also Zeit für Stellungnahmen oder irgendwelche Einbeziehungen schlichtweg nicht verfügbar. Die Frist ist ein geringer gewichtiges Argument - das will ich gern zugeben ... Das erste Argument ist der hohe Grad der Vertraulichkeit, der eine Einbeziehung Dritter, Externer, schlichtweg nicht erlaubt.

Stefan Sauerland: Vielleicht darf ich das noch ein bisschen ergänzen. Herr Windeler hat von „nicht erlaubt“ gesprochen. Wir haben es versucht - um es klar zu sagen -, durch einen Fragebogen, den wir durch klinische Experten ausfüllen lassen, mehr über das Erkrankungsbild zu erfahren. Das ist aber in der Regel äußerst schwierig, wenn wir die Intervention, um die es geht, überhaupt nicht benennen dürfen.

Ich nehme ein fiktives Beispiel: Da kommt jemand und hat eine neue diagnostische oder therapeutische Intervention im Bereich „Asthma“ oder „koronare Herzerkrankungen“. Jetzt können wir diesem externen Sachverständigen überhaupt nicht sagen, worum es geht, sondern wir können ihn nur bitten: Erklär uns doch einmal, wie die Asthmadiagnostik und -behandlung funktioniert. - Oder: Wie funktioniert das bei der koronaren Herzerkrankung? - Dass das natürlich wenig hilfreich ist, ist klar.

Daher: Wir haben diesen Weg versucht, aber solange man nicht klar benennen kann, was eigentlich die Fragestellung ist und um welchen Punkt im diagnostisch-therapeutischen Algorithmus es geht, ist eine solche Beteiligung leider im gegenwärtigen Kontext nicht sinnvoll.

Moderator Jürgen Windeler: Ja, bitte, Frau Soskuty.

Gabriela Soskuty: Ich möchte etwas zum Vertraulichkeitsthema sagen. Was Sie gerade gesagt haben, Herr Sauerland, verstehe ich einerseits. Andererseits ist das ein Problem, das uns in der Industrie bzw. in der Entwicklung permanent begegnet, dass wir vertrauliche Dinge haben, wofür wir Experten brauchen, denen wir eigentlich nicht erzählen wollen, was wir machen. Das kann man ganz einfach lösen, indem man Vertraulichkeitsvereinbarungen unterzeichnen lässt. Ich glaube, das wäre im Prinzip nicht das Problem - vor allen Dingen nicht in dem Zusammenhang, dass hinterher, nach der Potenzialbewertung darin steht, dass

hinterher das IQWiG sowieso alles veröffentlicht, was dann im Rahmen der Potenzialermittlung zusammengetragen wurde. Hinterher ist es eh öffentlich.

Es geht nicht darum, dass 20 Fachgesellschaften angefragt werden, die alle eine Stellungnahme abgeben sollen, sondern vielleicht um die Einbeziehung einer Fachgesellschaft, die tatsächlich genau zu diesem Thema vielleicht noch eine fundierte Bemerkung machen kann.

Moderator Jürgen Windeler: Vielleicht noch einmal: Das ist ein Thema, das wir hier nicht ausweiten wollten, aber wir können das gern tun. Noch einmal: Wir können natürlich selbstverständlich, Frau Soskuty, keine Vertraulichkeitsvereinbarungen mit Fachgesellschaften schließen, sondern wir müssten Vertraulichkeitsvereinbarungen mit einzelnen Personen schließen.

Wenn Sie sich - der G-BA kann dazu, wenn er sich dazu in der Lage sieht, Auskunft geben; ich werde das nicht tun - die Vorrichtungen, die Prozesse anschauen, ist hier für Geheimhaltungsvereinbarungen mit Externen kein Raum.

Der zweite Punkt, den Sie gerade angesprochen haben, betraf die Veröffentlichungen. Wir veröffentlichen überhaupt nichts.

(Stefan Lange: Das dürfen wir ja nicht!)

Wir dürfen gar nichts veröffentlichen.

Vielleicht noch zu dem einen Punkt, der vorhin angesprochen worden ist - ich glaube, Sie, Herr Winkler, waren das -, zur Frage, dass man das doch ein bisschen besser darlegen könne, damit Hersteller daraus lernen können: Das Problem ist, dass Bewertungen, bei denen wir und schließlich auch der G-BA in seiner Beschlussfassung kein Potenzial anerkennen, überhaupt nie das Licht der Welt erblicken - weder die Bewertung noch das Thema. Dafür gibt es Gründe. Das bedeutet natürlich auch, dass Hersteller und zukünftige Antragsteller aus solchen negativ beschiedenen Anträgen nichts lernen können, weil sie das nie erfahren.

Noch einmal: Das IQWiG veröffentlicht überhaupt nichts. Das IQWiG veröffentlicht schon überhaupt gar keine vertraulichen Dokumente. Das IQWiG veröffentlicht seine Potenzialbewertung bzw. die Potenzialbewertung des IQWiG wird veröffentlicht mit dem Richtlinienbeschluss des G-BA - aber auch dort nur die Dinge, die die Öffentlichkeit erreichen können.

Monika Nothacker: Das heißt also: Das IQWiG wäre sowohl dafür, dass externe Sachverständige eingebunden würden, wenn das möglich wäre, als auch dafür, dass diese abgelehnten Potenzialbewertungen veröffentlicht würden. - Ist das richtig?

Moderator Jürgen Windeler: Das IQWiG würde es begrüßen, wenn bei allen seinen Produkten die gleichen Beteiligungsmöglichkeiten für - jetzt will ich wieder ausdrücklich sagen: nicht Fachgesellschaften, mich interessieren die Institutionen nicht so sehr -

(Monika Nothacker: Das habe ich nicht gesagt! Ich habe gesagt ...)

externe Sachverständige, die üblicherweise aus Fachgesellschaften kommen, bestünden und sie mit eingebunden werden. Bei normalen Nutzenbewertungen passiert das so. Bei Dossierbewertungen passiert das wegen der Fristen in ähnlicher Weise, wie es Stefan Sauerland gerade beschrieben hat. Dort kann es um die Klärung konkreter Fragen gehen, weil dort auch das Produkt genannt werden kann. Bei den Potenzialbewertungen geht das nicht. Wir würden uns vorstellen, dass da eben auch - das hängt am Verfahren - eine Beteiligung möglich ist.

Ich habe - das sage ich auch ganz offen - erhebliche Bedenken und glaube, dass das auch faktisch nicht realisierbar sein wird, dass in den Verfahren - das gilt übrigens für die Dossierbewertungen genauso -, in denen das IQWiG für seine Bewertungen Einblickmöglichkeiten in vertrauliche, als „Betriebs- und Geschäftsgeheimnisse“ gekennzeichnete Dokumente hat, eine relevante Beteiligung an der Bewertung - auch das findet bei Dossierbewertungen übrigens nicht statt - möglich ist. War das klar genug?

Monika Nothacker: Ja.

Stefan Lange: Es würden sich viele Probleme lösen lassen, wenn die Industrie nicht darauf bestehen würde, dass in Studienberichten Dinge als „Betriebs- und Geschäftsgeheimnisse“ gekennzeichnet werden können. Das gilt nicht nur für die Medizinprodukteindustrie, sondern auch für die Arzneimittelindustrie.

Moderator Jürgen Windeler: Herr Dintsios, Ihr Stichwort.

Ch.-Markos Dintsios: Ich habe schon darauf gewartet, ich wusste, dass das kommt.

Entnehme ich Ihren Ausführungen, Herr Windeler, dann auch, dass bei nicht als vertraulich gekennzeichneten Dokumenten ... Im Rahmen der Dossierbewertung - bleiben wir bei dem Beispiel - beziehen Sie medizinische Fachexperten und Patientenvertreter ein. Wir erfahren dann immer bei den IQWiG-Bewertungen deren Namen. Wir erfahren sehr vieles über ihre potenziellen Interessenskonflikte. Was wir nicht erfahren, obwohl wir wissen, dass es einen Fragekatalog gibt, ist, was Sie darauf geantwortet haben. Kann es sein, dass Sie selbst Ihre eigenen Aussagen vertraulich schalten? Ich glaube, die meisten kommen nicht von der Industrie.

Beate Wieseler: Herr Windeler hat eben geschildert, in welchem Rahmen die Beteiligung der externen Sachverständigen in der Dossierbewertung aufgrund des Verfahrens möglich ist. Die Sachverständigen sind aufgrund der Struktur des Verfahrens nicht an der Bewertung beteiligt.

Die Bewertung, so wie sie veröffentlicht wird, wird vom IQWiG vorgenommen. Deshalb ist nicht klar, was in diesem Zusammenhang ... Wie gesagt, der Fragebogen ist einsehbar. Die Notwendigkeit, die Antworten der Experten auf diese Fragen zu veröffentlichen, sehe ich nicht.

Moderator Jürgen Windeler: Wir haben jetzt etwas den Faden verloren.

(Heiterkeit)

Es scheint ja, um noch einmal zum Faden zurückzukommen - Stefan Lange hat das vorhin formuliert -: Der Begriff „minimal“, der für die Ergebnissicherheit an dieser Stelle steht, heißt, dass wir wenigstens - bei weniger würden wir uns extremst schwertun, ein Potenzial anzuerkennen, um es vorsichtig zu sagen - klinische Daten im Sinne von Fallserien erwarten würden. Das ist der Minimalstandard.

Bei allen anderen Dingen, die in der Kategorie V des G-BA abgebildet werden könnten, also Assoziationsstudien oder pathophysiologische Überlegungen oder Meinungen von Experten oder auch Patienten, würden wir uns - ich beschreibe das noch einmal so - extremst schwertun. Das heißt immer auch: Es mag einen Fall geben, in dem die biologische und die pathophysiologische Plausibilität so durchschlagend ist, dass man gar nicht anders kann, und wo jeder sagen würde: Wir brauchen gar keine empirischen Daten. - Das mag sein, aber das würde ich als eine Rarität bezeichnen wollen.

Wenn ich das noch einmal zusammenfassen darf: Mir lässt die Erläuterung des Begriffs „minimal“ und das Verständnis dieses Begriffes ...

(Wortmeldung von Marianne Abele-Horn)

- Sofort. Jetzt habe ich genug geredet, jetzt kommt die Diskussion wieder ein bisschen in Schwung. Bitte schön.

Marianne Abele-Horn: Nur eine Frage - Entschuldigung -, weil ich da thematisch nicht so zu Hause bin. Wer bewertet es denn, was für Berufsgruppen ... Wenn Sie sagen, Sie bewerten es an Ihrem Institut: Wer bewertet diese klinischen Fragestellungen? Sind das Ärzte, die da sitzen, oder wer macht das? Wie muss ich mir das vorstellen?

(Monika Nothacker: Das ist ganz interdisziplinär, nicht wahr, Herr Windeler?)

Moderator Jürgen Windeler: Das IQWiG beschäftigt auch Ärzte, aber an den Bewertungen sind sowohl Ärzte als auch Nichtärzte beteiligt.

(Monika Nothacker: Also vorwiegend Biologen und Ernährungswissenschaftler, oder?)

- Wir haben Soziologen, Pharmakologen, Pädagogen, Pharmazeuten, alles, was das Herz begehrt.

(Zuruf: Und Psychologen!)

Marianne Abele-Horn: Und wie machen Sie das, wenn Sie eine klinische Frage bewerten müssen?

Moderator Jürgen Windeler: Wir müssen keine klinische Frage bewerten, sondern ...

Marianne Abele-Horn: Eine Studie halt, die mit ...

Moderator Jürgen Windeler: Tja, wie machen wir das?

(Heiterkeit)

Ich versuche, das in 2 bis 3 Sätzen zusammenzufassen, ohne das weiter zu vertiefen. Wenn wir einen Auftrag vom G-BA bekommen, machen wir uns mit dem Thema vertraut, lesen uns ins Thema ein und machen eine sehr systematische akribische Suche nach den vorhandenen Studien und nach den vorhandenen Daten. Vorher veröffentlichen wir einen Berichtsplan, der zur Diskussion gestellt wird, bei dem Fachgesellschaften - das darf ich an dieser Stelle auch erwähnen - jede Menge Möglichkeiten haben, sich ausführlich zu beteiligen, was sie oft auch tun, und uns Rückmeldungen zu geben, ob wir uns in unseren insbesondere klinischen Betrachtungen möglicherweise geirrt haben. Auf dieser Basis wird ein Bericht gemacht, der die Evidenz zusammenfasst und versucht, die Frage, die der G-BA hat, zu beantworten. Dies geschieht - das will ich noch erwähnen - bei allen Bewertungen, die nicht unter die Dossier- und nicht unter die Potenzialbewertungen fallen, unter Beteiligung externer Sachverständiger.
- Ralf, bitte.

Ralf Bender: Das geht dann genauso, wie es im Methodenpapier beschrieben ist - sowohl der Prozess als auch die Bewertungskriterien.

Moderator Jürgen Windeler: Gut, nachdem wir das jetzt auch ...

(Wortmeldung von Andrej Rasch)

- Gut, Herr Rasch, bitte.

Andrej Rasch: Sie haben meine Anmerkung zuvor leider vergessen. - Um ganz kurz zum Thema „externe Sachverständige“ zurückzukommen - ich möchte gern einen Spielverlauf machen -: Sie haben zwischendurch erwähnt, dass die Einbindung explizit nicht möglich sei. Ich bin deswegen verwirrt, weil im Methodenpapier durchaus die Option vorgesehen ist. Deswegen nur zur Richtigstellung für das Protokoll, dass das zumindest nicht explizit unmöglich ist.

(Zustimmung von Stefan Sauerland)

Moderator Jürgen Windeler: Herr Sauerland nickt, und ich sage gern Ja, wenn das für das Protokoll bedeutend ist. - Herr Danner.

Martin Danner: Ich würde gern auf diesen Punkt zurückkommen, wonach Herr Lange vorhin gesagt hat, es wäre doch schön, wenn man eine Linie als Regelfall hätte, wann man Potenzial annehmen kann oder nicht. Ich wollte dazu eine Nachfrage stellen, wie das IQWiG seine Rolle in diesem Prozess sieht. Wenn man sich nämlich zunächst einmal anschaut, um welche Entscheidung es geht, ist es letztlich die Frage: Sollen Erprobungsstudien gemacht werden? Diese Entscheidung trifft der Bundesausschuss. Als jemand, der da mit am Tisch sitzt, ist natürlich interessant: Welche Wirkungen werden erwartet? Welche Risiken werden erwartet? Gibt es bereits etliche Therapieoptionen, oder ist es etwas, das wirklich etwas sozusagen ganz Neues ermöglichen könnte? Das sind natürlich Aspekte, die wichtig sind.

Ich muss sagen: Für uns ist es immer eine Frage, ob es sich um eine hochriskante Geschichte oder um eine handelt, die potenziell wenig Risiken in sich birgt. Danach würden wir natürlich auch ein bisschen abwägen: Können wir es verantworten, die Patienten, die in den Erprobungsstudien damit konfrontiert werden, einzubeziehen? Beziehungsweise: Wollen wir der Versichertengemeinschaft diese kleine Tür öffnen, so etwas vielleicht doch zu haben? Da kann es Abwägungsspielräume geben, weshalb man sich im Einzelfall anschauen muss, ob man ein Potenzial bejaht oder nicht.

Zu meiner Frage: Welche Rolle nimmt das IQWiG in diesem Prozess ein? Zu sagen: „Unser Begriff von Potenzial ist einer, der sich auf die Güte der Evidenzlage bezieht, und die Bewertung, ob etwas in die Erprobung kommt, ist noch einmal ein eigener Schritt, der vom G-BA zu gehen ist“, oder ist es so, dass das wie eine Sachverhaltsfeststellung für den G-BA getroffen wird: Potenzial liegt vor - Tatsache: Potenzial gegeben -, und der G-BA entscheidet nur noch einmal. - Mit anderen Worten: Gibt es da sozusagen einen Potenzialbegriff des IQWiG, der sich vom Potenzial im Rahmen der G-BA-Entscheidung unterscheidet, weil eben diese wertenden Elemente da eine Rolle spielen? Ich finde, das ist wichtig, um überhaupt sagen zu können, ob das IQWiG für sich eine Linie macht oder welche Bedeutung die Linie für die ganzen Entscheidungen hat.

Stefan Lange: Ich bin nicht ganz sicher, ob ich Ihre Frage richtig verstanden habe oder was die Implikationen waren, die Sie mit Ihrer Frage im Kopf hatten.

Klar ist, dass wir eine Empfehlung an den Gemeinsamen Bundesausschuss geben, und zwar über die Frage, ob eine Methode, die in einem Antrag beschrieben ist, das Potenzial - so steht es im Gesetz - einer erforderlichen Behandlungsalternative besitzt. An dieser Formulierung orientieren wir uns, weil das verschiedene Implikationen hat. Denn „Behandlungsalternative“ heißt: Es muss etwas anderes geben, denn sonst gibt es keine Alternative. Es muss also

irgendeinen sinnvollen Vergleich zu dem geben, was schon da ist. Der Begriff des Erfordernisses hat auch Implikationen.

Weil Sie fragten - das wäre der zweite oder der eigentliche Teil Ihrer Frage -, an welchem Potenzialbegriff wir uns orientieren, sage ich Ihnen: Das ist in erster Linie das, woran wir uns orientieren. Natürlich bewerten wir auch - das muss zwar nicht, sollte aber in den Anträgen spezifiziert werden -, ob durch die vorgelegten Unterlagen auch Studien konzipierbar sind, um hinterher gegebenenfalls den Nutzen zu beweisen. Aber das ist sozusagen erst der zweite Schritt.

Ein erster Schritt ist die Empfehlung, ob wir ein Potenzial in diesem Sinne sehen oder nicht. Davon abgesehen kann natürlich der Gemeinsame Bundesausschuss wie bei allen unseren Empfehlungen durch seine sachgerechte Entscheidung von unseren Empfehlungen abweichen. Er ist natürlich immer auch selbst als Herr oder Herrin des Verfahrens in der Lage, die Unterlagen, die dabei eingereicht werden, selbst zu bewerten.

Moderator Jürgen Windeler: Herr Danner.

Martin Danner: Eine Nachfrage: Um es an einem Beispiel zu machen: Es sind Methoden denkbar, wo das Schlimmste, was passieren kann, eine Hautreizung ist. Es sind Methoden denkbar, bei denen der Patient zwar nicht vital bedroht ist, aber das Schlimmste, was passieren kann, ist dann der Tod durch diese Intervention. Spielen solche Aspekte eine Rolle - je nachdem, wo man die Linie zieht, sodass man sagt: Beim letzten Fall muss die Evidenzlage sicherer sein als im ersten Fall? Das würde mich methodisch interessieren.

Stefan Sauerland: Wir sind, Herr Danner, in unseren Bewertungen erst einmal an diese Sechswochenfrist gebunden. Wir sind auch daran gebunden, das zu nehmen, was im Antrag steht. Das ist oft nicht viel.

Das Zweite, was wichtig ist: Wir kümmern uns um Daten. Wenn die Methode vielleicht ein Risiko bietet, es aber gar keine Daten dazu gibt, dass diese Komplikation jemals aufgetreten ist, wird es schwierig für uns, irgendetwas daraus zu machen. Wir versuchen, da tatsächlich mit gleichem Maßstab den Nutzen und den möglichen Schaden zu messen. Wenn es nur pathophysiologische Überlegungen zum Nutzen gibt, die uns nicht interessieren, interessieren uns genauso wenig die pathophysiologischen Überlegungen zu einem möglichen Schaden.

Nichtsdestotrotz beinhalten aber diese Antragsformulare, die der G-BA auf seiner Website öffentlich hält, relativ umfangreiche Angaben gerade im Bereich Medizinprodukte über Vorkommnisse mit Medizinprodukten. Das sind Elemente, die natürlich dann im G-BA eine Rolle spielen sollten, weil dort am Ende die Entscheidung getroffen wird. Da muss natürlich ein Abwägen stattfinden. Das Abwägen findet bei uns auch in gewisser Weise statt. Aber die letztliche Entscheidung - das ist auch klar - wird der G-BA treffen. Da sind auch die Patientenvertreter mit am Tisch und werden auch gehört. Das kann ich durchaus bestätigen,

dass man dann eine sehr wohl ausgewogene Abwägung zwischen Nutzen und Schaden vornimmt, selbst wenn es sich nur um potenziellen Nutzen oder potenziellen Schaden handelt.

Moderator Jürgen Windeler: Herr Mühlbacher.

Axel Mühlbacher: Zuerst einen kleinen Hinweis: Herr Sauerland, wenn bei der Entscheidung oder Bewertung beim G-BA Patienten anwesend sind, so tut das bei der Bewertung nichts zur Sache, weil vorher die Nutzen-Schaden-Abwägung auf Basis von Evidenz getroffen wurde. Nur die alleinige Anwesenheit von Patienten ...

(Störgeräusche in der Mikrofonanlage)

- Ich entschuldige mich - für das Protokoll.

Wenn die Patienten bei der Bewertung anwesend sind, heißt das nicht, dass man bei der Messung auf die Evidenz der Abwägung von Nutzen und Schaden verzichten kann. Das ist nicht das Gleiche.

Jetzt zur Frage: Gehe ich recht in der Annahme, dass das Potenzial auf patientenrelevanten Endpunkten basiert, die nur einen unterschiedlichen Evidenzgrad haben könnten?

(Widerspruch von Stefan Lange)

- Ich halte deswegen inne, weil Sie mit dem Kopf schütteln, Herr Lange.

Stefan Lange: Ich glaube, das ist im Methodenpapier auch beschrieben. Nein, ein Potenzial kann sich auch aus Surrogaten ergeben.

Axel Mühlbacher: Entschuldigung, da war ich zu ungenau. - Ich meinte: Endpunkte und Surrogate. Gehe ich recht in der Annahme: Es gibt eine Abwägung von Nutzen und Schaden, um dann zum Potenzialbegriff zu kommen? Das heißt: Der Potenzialbegriff ist dann ähnlich wie ein Nutzenmaß ein Aggregat der gemessenen Effekte. Jetzt wird es spannend, wenn Sie - immer gegenüber der Alternative - vergleichen wollen: Reichen Ihnen die patientenrelevanten Endpunkte, weil Sie mit gleichen Kriterien messen? Wenn ich es richtig verstanden habe, trifft der G-BA einmal pro Jahr die Entscheidung über die Studien an sich und vergleicht dort nicht indikationsspezifisch die Alternative mit der neuen Studie, sondern vergleicht über alle eingereichten Anträge für Studien. Das heißt: Sie haben keine indikationsspezifische Bewertung, sondern eine integrationsübergreifende?

Stefan Lange: Das ist ein Missverständnis, Herr Mühlbacher. Damit hat die IQWiG-Bewertung überhaupt nichts zu tun. In der Tat ist es so, dass der G-BA, wenn es zahlreiche Anträge mit positivem Potenzial geben sollte, natürlich aufgrund der begrenzten Ressourcen eine Auswahl treffen muss, in welchen Bereichen eine Erprobungsstudie stattfinden soll und in welchen gegebenenfalls nicht. Aber das hat erst einmal mit den Einzelbewertungen nichts zu tun. Da kommt es dann zum üblichen Problem mit der Prioritätensetzung. Ich muss

gestehen: Die Methodik des G-BA dafür ist mir gegenwärtig auch nicht bekannt, aber das hat, wie gesagt, mit der einzelnen Bewertung nichts zu tun. Das ist aber das übliche Problem - auch in der Forschung -: Wenn man begrenzte Forschungsmittel hat, muss man die natürlich auch allozieren, etwa über Indikationen, über Fragestellungen hinweg. Aber noch einmal: Das hat mir der einzelnen Bewertung nichts zu tun.

Vielleicht noch ein Stichwort, Herr Danner: Ich finde es sehr spannend, dass ich den Eindruck habe: Sie problematisieren ein Stück weit dahin gehend, dass wir unter Umständen mit unseren Potenzialbewertungen zu viel falsch Positive haben könnten. Wir sind in der Vergangenheit eher mit dem Gegenteil konfrontiert worden, dass wir zu viel falsch negativ machen. Ich spreche nicht über Potenzialbewertungen - das ist hoch vertraulich -, also: Bitte nicht falsch verstehen.

Man könnte das Problem aber sehr gut lösen. Noch einmal - das habe ich eben schon gesagt -: Diese Besorgnis kann ich durchaus nachvollziehen, insbesondere wenn wir uns in den Bereich von nichtrandomisierten Studien begeben. Denn bekannt ist, dass in diesem Bereich die Gefahr eines Publikationsbias noch viel größer als im Bereich der randomisierten Studien ist. Da wäre es natürlich extrem hilfreich, wenn wir das Ganze öffentlich machen könnten, um eben eine möglichst große Schar von Menschen dafür zu gewinnen, ihre Expertise dort einzubringen, zum Beispiel zu eventuellen Risiken, die möglicherweise noch gar nicht richtig öffentlich gemacht worden sind.

Moderator Jürgen Windeler: Ich habe jetzt Frau Soskuty und Frau Nothacker bei mir auf der Liste. Habe ich andere Meldungen übersehen? Ich schaue noch einmal ganz nach rechts und ganz nach links. - Anscheinend nicht. Frau Soskuty.

Gabriela Soskuty: Eine kurze Nachfrage: Mich irritiert immer noch der zweifache Potenzialbegriff ein bisschen. Ist es denn vorstellbar, dass das erste Potenzial, also „erforderliche Behandlungsalternative“ positiv bewertet wird? Sie sagen: „Es hat das Potenzial, eine erforderliche Behandlungsalternative zu werden“, aber das zweite Potenzial wird negativ bewertet. Aber man kann keine Studie machen. - Ist das denn überhaupt vorstellbar? Denn wenn es nicht vorstellbar ist, könnte man sich diesen zweiten Potenzialbegriff sparen.

Moderator Jürgen Windeler: Das Problem ist, dass sie beide im Gesetz stehen, und zwar in dieser unterschiedlichen Differenzierung. Sparen kann man sich das nicht, sondern man muss eine Aussage dazu machen. Vorstellbar wäre es zum Beispiel in Situationen, in denen, um Herrn Danner zu adressieren, vielleicht die vorbereitenden Informationen für eine Studie, auch was meinetwegen den Ausschluss existenzieller Risiken angeht, noch nicht ausreichen.

Eine andere Situation, die auftreten könnte, wäre, dass man zum Beispiel weiß, dass Studien zu einer Thematik, die man dort adressiert, schon laufen. Dann würde man auch sagen: Da ist vielleicht nicht das Potenzial adressiert, sondern eher die Frage: Sollte man das machen?

Ich kann mir Situationen vorstellen, die ein Potenzial bejahen, die aber sagen: Ein Potenzial für eine Erprobung - einen Grund habe ich versucht zu nennen; aber das ist jetzt alles ein bisschen theoretisch - besteht nicht oder ist jedenfalls nicht ausreichend. - Frau Nothacker.

Monika Nothacker: Das geht von der inhaltlichen Sache weg. - Zum einen ist es so: Die AWMF würde auch begrüßen, wenn diese Verfahren veröffentlicht würden - aus den von Ihnen genannten Gründen, Herr Windeler. Zum anderen möchte ich noch einmal auf die externen Sachverständigen zurückkommen. Ist es denn so, dass Sie im IQWiG die Beiträge dieser externen Sachverständigen anhand dieser Fragebögen evaluieren? Sind die ausreichend gut? Sind sie ausreichend in ihren Beiträgen? - Gibt es da eine interne Qualitätssicherung?

Beate Wieseler: Wir haben sie nicht formal evaluiert. Wir haben aber in dem Verfahren die Möglichkeit, dass wir, wenn wir den Fragebogen an die externen Sachverständigen schicken und zurückbekommen und dann Fragen für uns offenbleiben, innerhalb dieser Bewertungsfrist erneut auf die externen Sachverständigen zuzugehen, um die Dinge, die für uns offengeblieben sind, zu klären. Das tun wir durchaus wiederholt. Das tun wir dann, wenn Fragen aus dem Fragebogen offenbleiben oder wenn im Laufe des Bewertungsverfahrens die Notwendigkeit besteht, noch weitere klinische Expertise einzuschließen und einzuholen. Dann gehen wir auch nach der Beantwortung des Fragebogens durch den externen Sachverständigen erneut auf die Sachverständigen zu.

Monika Nothacker: Sie hätten aber nicht die Möglichkeit zu sagen: Hier reicht uns das nicht aus, wir möchten noch andere einbeziehen? Es kann dann immer nur dieser eine sein.

Beate Wieseler: Ich will überhaupt nicht ausschließen, dass, wenn wir eine für unsere Bewertung relevante Frage mit einem Sachverständigen nicht klären können, wir rein theoretisch versuchen können, in der Bewertung einen zweiten Sachverständigen zu gewinnen. Das ist bisher nicht vorgekommen,

(Monika Nothacker: Okay!)

dass wir mit den externen Sachverständigen, die wir einbezogen haben, die Fragen nicht klären konnten.

Stefan Sauerland: Noch eine Ergänzung: Im Bereich der Potenzialbewertung ist das, wie gesagt, äußerst schwierig. Da sind auch diese Nachfragen, die Frau Wieseler jetzt beschrieben hat, eigentlich nicht möglich, weil ich auch in der Nachfrage nicht klar ansprechen kann, worum es eigentlich geht und was wir bewerten.

In anderen Bereichen, in der konventionellen Nutzenbewertung, kann ich es mir sehr gut vorstellen, dass man tatsächlich auch mal eine Situation hat, in der 2 Schulen zueinander in Konkurrenz stehen. Dann soll man eine Bewertung machen, hat aber das Problem: Wenn man nur einen Sachverständigen oder eine Sachverständige gewönne, würde sie aus dem rechten oder dem linken Lager kommen. Dann kann es durchaus sinnvoll sein, aus beiden Lagern

jeweils einen zu holen, um auf diese Weise ein bisschen die Bandbreite der Meinungen und das Spektrum der Einflüsse unter Kontrolle zu halten.

Moderator Jürgen Windeler: Nachdem wir unter diesem Tagesordnungspunkt noch ein paar prozessuale und rechtliche Klärungen sowie den Tagesordnungspunkt „Einbeziehung von Sachverständigen“ noch mitbearbeitet haben, würde ich jetzt, da ich keine weiteren Wortmeldungen mehr sehe - auch nach kurzem Innehalten nicht -, diesen Tagesordnungspunkt jetzt abschließen wollen. Wir kommen nun zu TOP 5.

Tagesordnungspunkt 5: Darstellung von Unsicherheit in KNB

Andreas Gerber-Grote: Es geht um die Darstellung von Unsicherheit in Kosten-Nutzen-Bewertungen. Wir haben damals in den Methoden 2009 verschiedene Möglichkeiten genannt, wie Unsicherheit dargestellt werden kann, unter anderem Konturplots, Konfidenzbänder, Scatterplots. Dann haben wir, wie Sie alle wissen, eine große Kosten-Nutzen-Bewertung zu den Antidepressiva gemacht. Im Rahmen dieser Kosten-Nutzen-Bewertung und im Austausch mit Expertinnen und Experten zur Frage der Darstellung von Unsicherheit haben wir uns dann insbesondere auf das Verfahren des Net-Health-Benefit und auf das Verfahren der sogenannten Preisakzeptanzkurven in Anlehnung an die Kostenakzeptanzkurven festgelegt.

Hierzu gab es nicht so viele Stellungnehmende. Auch der vfa hat hierzu eine Frage formuliert. Deswegen zitieren wir sie und stellen eine Frage dazu. Ein Stellungnehmender schreibt, dass zur Kommunikation von Unsicherheit folgende Darstellungen im Rahmen der Kosten-Nutzen-Bewertungen sinnvoll seien: Scatterplots, Vertrauensintervalle, Konturplots, Kosten-, Effektivitäts-, Akzeptanzkurven und Kovarianzanalysen, sofern Kosten und Effekte aus derselben Quelle stammen. Könnte der Stellungnehmende uns bitte kurz erläutern, worin seiner Ansicht nach die Vorteile der unterschiedlichen Darstellungen bestehen und welche Form der Darstellung er für sinnvoll hält? Denn es musste irgendwann einmal eine Auswahl getroffen werden. Sowohl aus Ressourcengründen als auch aus Darstellungsgründen können nie alle Möglichkeiten gewählt werden.

Moderator Jürgen Windeler: Herr Dintsios.

(Störgeräusche in der Mikrofonanlage)

- Dahinten ist noch ein Mikrofon an. Ich glaube, bei Frau Nothacker. Können Sie das bitte ausschalten? - Danke!

Ch.-Markos Dintsios: Die Position war auch Ausfluss des Net-Health-Benefit-Ansatzes. Das IQWiG schlägt in seinem Methodenentwurf 4.2 vor, den Net-Health-Benefit-Ansatz sowohl für die Base-Case- als auch für die probabilistischen Sensitivitätsanalysen und auch für die deterministischen anzuwenden. Wir haben darauf hingewiesen, dass aufgrund der Kosten-

verteilung im Rahmen von gesundheitsökonomischen Evaluationen, die meist extrem schief verteilt sind - nicht immer, aber sehr oft -, ein Net-Health-Benefit-Ansatz, der sich auf den Interquartilsabstand bezieht, eventuell zu glättend wirken könne. Das blendet sehr große Teile der Verteilung aus.

Die Methoden, die wir hier schildern, haben natürlich den Vorteil, dass sie sozusagen die gesamte Varianz mit abbilden können. Wir haben sie auch so genannt, weil sie auch früher in den IQWiG-Dokumenten so mitgeführt wurden. Auf die Frage hin, welche aufgrund von Ressourcenengpässen die geeignetere sei, kann man wiederum verschiedene Kriterien ansetzen. Was will man den Entscheidungsträgern mitteilen? Beim Health-Benefit-Ansatz im Rahmen einer probabilistischen Sensitivitätsanalyse holt man sich nur Teile des Geschehens mit auf die Empfehlung und sieht nicht das gesamte Spektrum.

Das andere: Wenn ich mich als Zielgruppe für den Entscheidungsträger entscheide, würde ich eher intuitiv das vorschlagen, was er am besten nachvollziehen kann. Ob das Konturplots oder Scatterplots sind, bleibt dahingestellt - visualisierend oder nicht. Letztlich sind aber die Credibility Intervals sicherlich einer der internationalen gesundheitsökonomischen Standards. Die spezielle Konstellation einer Piggy-Back-Studie mit einer Kovarianzanalyse, glaube ich, erklärt sich von allein.

So ist auch dieser Kommentar zu verstehen. Er ist weniger die Aufforderung, diese Methoden parallel anzuwenden. Man kann sie auch alternativ anwenden. Er ist eher spezifisch für die deterministische und probabilistische Sensitivitätsanalyse gedacht, und nicht, den etwas glättenden Net-Health-Benefit-Ansatz ausschließlich zu verwenden. Das ist die primäre Zielsetzung, den für den Base-Case übrigzulassen, weil er, wie gesagt, eher mit dem Interquartilsabstand arbeitet.

Andreas Gerber-Grote: Vielleicht müssen wir dann die Diskussion sauber aufarbeiten. Wir können erst einmal, dass was von dir, Markos, „Interquartilsabstand“ genannt wurde, trennen. Wir haben erst einmal verschiedene Darstellungsmethoden. Der Interquartilsabstand ist dann nur eine wertmäßige Festlegung. Ich könnte auch ein Konfidenzintervall auch von 50 % wählen; das ist gar nicht die Debatte. Wir müssen zunächst einmal die verschiedenen Ansätze unterscheiden.

Da vielleicht nur als Beispiel: Wenn man sich dann doch genauer anschaut, was man wie kommunizieren kann ... Ich verweise da auch auf eine Publikation aus unserem Hause, unter anderem mit Herrn Stollenwerk und internationalen Expertinnen und Experten, die im Frühjahr erschienen ist. Mit dem Scatterplot zum Beispiel kann man den Entscheidungsfindern gar nichts darstellen, weil Sie in einem Scatterplot einfach nur Punktwolken haben und gar nicht wissen, welche dieser Punkte zu welcher Effizienzgrenze gehören. Damit haben Sie eigentlich überhaupt keine sinnvolle Darstellung für einen Entscheidungsfinder - genauso, wenn Sie ein Konfidenzband wählen. Das Konfidenzband ist

nie über Ihre Effizienzgrenzen definiert. Da sollte man doch noch einmal genauer hinschauen, welches Maß man wählt.

Der Net Health Benefit wirkt in dem Moment, in dem ich ihn über die 100 % seiner Bandbreite nehme, nicht glättend. Noch einmal: Der Net Health Benefit oder die Kosteneffektivitätsakzeptanzkurve sind primär erst einmal gar nicht glättend. Deswegen möchte ich das klar von dem anderen Punkt trennen: Ist der Interquartilsabstand das Richtige? Das wäre dann auch verknüpft - da gebe ich dir Recht, Markos - mit der Frage, was ist eigentlich das Ziel? Wozu soll die Kosten-Nutzen-Bewertung dienen? Da könnte man das durchaus diskutieren.

Moderator Jürgen Windeler: Herr Dintsios noch einmal.

Ch.-Markos Dintsios: Ganz kurz nur. Ich glaube, Herr Gerber hat in seinem letzten Satz auch das wiedergegeben, was ursprünglich von uns intendiert war, dass man genau diesen Interquartilsabstand im Zusammenhang mit einer möglichen Handlungsempfehlung für Entscheidungsträger noch einmal diskutiert und eine gewisse Reflexion betreibt, ob es eine andere Möglichkeit gibt. Mit den anderen Sachen sind wir völlig d'accord in Wahrheit.

Moderator Jürgen Windeler: Dann scheint in Abwesenheit weiterer Wortmeldungen dieser Punkt, jedenfalls die Interpretation der Stellungnahme, geklärt zu sein. - Ich schließe dann TOP 5 ab.

Tagesordnungspunkt 6

Verschiedenes

Moderator Jürgen Windeler: Gibt es zu TOP 6 Wortmeldungen oder Bedarf, Themen anzusprechen, die nicht auf der Tagesordnung standen? - Herr Rasch.

Andrej Rasch: Ganz kurz: Das steht auch in unserer Stellungnahme. Zum Nutzenbegriff gab es zuvor eine etwas weiter fassende Definition sowohl im allgemeinen Methodenpapier als auch im alten KNB-Papier 1.0, das immer noch aktuell ist. Diese weitere Beschreibung ist einfach abhandengekommen.

Des Weiteren sehen wir auch - das mündet wiederum in einen älteren Tagesordnungspunkt – die Verortung oder die Erhebung der Patientenpräferenzen oder der Utilitywerte, auch wenn man das zu einem Maß zusammenfügen möchte, allein im Bereich von § 35 b SGB V als etwas wenig nachvollziehbar, warum diese methodischen Verfahrensweisen nicht im Rahmen der Nutzenbewertung denkbar sind.

Moderator Jürgen Windeler: Herr Gerber-Grote.

Andreas Gerber-Grote: Man kann sicherlich darüber debattieren, wo die richtige Verortung ist. Ich glaube - wir haben auch Verweise zwischen Kapitel 3 und 4 gesetzt -: Es ist sinnvoll, sie im Kapitel 4 „Kosten-Nutzen-Bewertung“ zu belassen, weil sie sehr eng, was in § 35 b SGB V „Maß des Gesamtnutzens“ genannt wird, anknüpft. Ich denke, es muss nicht zu einer Doppeldarstellung kommen. Deswegen halte ich es für sinnvoll, sie in einer Kosten-Nutzen-Bewertung darzustellen, ohne damit ein Statement darüber abzugeben, dass sie nicht auch in anderen Bereichen genutzt werden kann.

Die Darstellung zum weiten und engen Nutzenbegriff war damals einer Diskussion mit der Deutschen Gesellschaft für Gesundheitsökonomie geschuldet, die wir im Moment nicht noch einmal neu aufrollen müssen. Deswegen haben wir erst einmal darauf verzichtet, das explizit noch einmal darzustellen.

Stefan Lange: Vermutlich meinten Sie noch einen anderen Punkt. Das ist im Abschnitt 3.1.1 „Nutzendefinition“, wo wir zunächst einmal definieren, was wir unter „patientenrelevantem Nutzen“ verstehen: Mortalität, Morbidität, Lebensqualität. In der Tat: Zuvor im Methodenpapier war noch ein Passus, in dem die Kategorien aus § 35 b SGB V, bzw. der Arzneimittelnutzenbewertungsverordnung aufgeführt sind: Verlängerung des Überlebens, Verringerung der Krankheitsdauer, Vermeidung von Nebenwirkungen usw. usf. Wir haben gedacht: Durch den Verweis auf § 35 SGB V ist das klar. Weil das alle Aspekte von Mortalität, Morbidität und Lebensqualität sind, sei es redundant, das alles noch einmal aufzuführen. Wir können auch gern die Redundanz wieder aufnehmen, wenn das zu irgendwelchen Irritationen geführt hat. Das war nicht intendiert und nicht beabsichtigt. Wir werden auch weiterhin die Kriterien der Arzneimittelnutzenbewertungsverordnung anwenden.

Moderator Jürgen Windeler: Frau Abele-Horn.

Marianne Abele-Horn: Ich habe noch eine Frage, die unsere Fachgesellschaft gestellt hat, ob es möglich ist, bei der Aktualisierung für die Mikrobiologie das Intervall etwas auszudehnen, weil 3 Jahre für die Mikrobiologie und die Methoden einfach ein bisschen kurz sind.

Moderator Jürgen Windeler: Herr Waltering.

Andreas Waltering: Ich habe die Stellungnahme so verstanden, dass Sie die 3 Jahre für ein wenig illusorisch halten. Uns geht es nur darum, dass wir unsere Gesundheitsinformationen spätestens alle 3 Jahre auf noch bestehende Richtigkeit oder möglichen Aktualisierungsbedarf überprüfen.

Insofern ist mir nicht ganz klar, inwieweit dieses Intervall Ihr Intervall für die Aktualisierung Ihrer eigenen Informationen oder Ihrer wissenschaftlichen Ausarbeitung beeinflusst. Wir schauen spätestens alle 3 Jahre nach: Ist das, was wir da stehen haben, noch aktuell, oder gibt es mittlerweile neue Erkenntnisse, die in unsere Gesundheitsinformationen einfließen? Davon unbenommen sind Ihre eigenen Aktualisierungsintervalle.

Oder habe ich das jetzt vielleicht nicht richtig verstanden?

Moderator Jürgen Windeler: Herr Mühlbacher.

Axel Mühlbacher: Ich würde gern noch auf die vorhergehende Frage eingehen. Die Frage nach der Berücksichtigung von Patientenpräferenzen in den Kapiteln 3 und 4 ist schon berechtigt, denn in Kapitel 3 wurde gestrichen, ohne zu ersetzen. In Kapitel 4 wurde ausgeführt - stark verkürzt. Es gibt zwar die Verweise von Kapitel 4 auf Kapitel 3 und von Kapitel 3 auf Kapitel 4. Sie führen aber jeweils nicht zu dem Punkt, dass man Endpunkte, wenn man mehr als einen benutzt, gewichten muss, entsprechend einer Experten- oder Patientenmeinung - jedes Mal nicht.

Das heißt: Im Methodenpapier 4.2 ist genau in den Punkten nicht ausgeführt, wie die Gewichtung durchgeführt werden soll. Die Frage ist natürlich extrem wichtig. Wird in Kapitel 3 die Nutzenbewertung unter der Prämisse gemacht, dass man davon ausgeht, dass man Endpunkte gewichten muss, dass man ein multikriterielles Entscheidungsproblem hat, oder liegt das nur bei der Kosten-Nutzen-Bewertung vor? Diese Frage ist im gesamten Methodenpapier unklar. Noch einmal zur Wiederholung: Die Verweise helfen einem da nicht weiter.

Moderator Jürgen Windeler: Ich glaube, das haben wir auch schon mal angesprochen in den ersten beiden Punkten. Wir nehmen das mit. Wir schauen einmal, ob wir an den Stellen etwas verändern und verbessern können. Es hat schon Hintergründe, dass es so da drin steht, wie es da drin steht, was nicht heißt, dass es für alle Zeiten so bleiben muss, weil die - aber ich will den Punkt 2 nicht wieder komplett aufmachen - Frage der Patientenpräferenzen eine, wie Sie gerade sehr richtig gesagt haben, sehr sinnvolle und wünschenswerte ist. Sie ist aber gerade für uns so gut wie gar nicht operationalisierbar; möglicherweise ist sie nur teilweise im IQWiG angesiedelt. Insofern kommt es ein bisschen zu diesem unscharfen und für Sie unbefriedigenden Bezügen und Äußerungen.

Im Grundsatz, glaube ich, können wir alle fröhlich bestätigen, dass es eine sinnvolle Vorgehensweise ist, Patientenendpunkte in irgendeiner Weise zu gewichten; „gewichten“ ist schon zu viel, weil das qualitativ ist. Ich will lieber von „sortieren“ sprechen.

Ein belastbares Verfahren und belastbare Daten stehen uns nicht ohne Weiteres zur Verfügung, um es zurückhaltend zu sagen.

Axel Mühlbacher: Darf ich kurz einhaken? - Ich verstehe Ihren Punkt. Den brauchen wir wirklich nicht zu vertiefen. Aber sowohl bei der frühen Nutzenbewertung als auch bei der Kosten-Nutzen-Bewertung wird gewichtet:

(Monika Nothacker: Ja!)

implizit oder explizit. Das heißt: Die Frage ist nur: Welche Informationen gehen da ein? Eine mögliche Variante wären Patientenpräferenzstudien. Eine weitere Variante wäre ein abgezeichnetes Expertenurteil in einem Gremium wie hier: Alle halten die Hand hoch. - Das ist die Frage der Transparenz. Das Methodenpapier könnte helfen zu verstehen, dass a) diese Gewichtung bei beiden Entscheidungen notwendig ist und b) Hinweise darauf zu geben, wie man diese Information im Entscheidungsprozess beschaffen könnte.

Moderator Jürgen Windeler: Ich kann beide Punkte nachvollziehen. Wir werden noch einmal schauen, ob wir - wenn nicht in der Version 4.2, dann jedenfalls in einer Folgeversion - uns dieses Punktes speziell annehmen.

Stefan Lange: Um unsere Schwierigkeiten damit ein Stück weit transparent zu machen: Länger als das IQWiG gibt es Zulassungsbehörden, die auch solche Abwägungen bei ihren Zulassungsentscheidungen treffen. Sie hatten vorher das Projekt der EMA angesprochen, das vor einigen Jahren in mehreren Paketen gestartet worden ist, sich diesem Thema zu nähern, wie man das stärker explizit machen kann. Es gibt ein paar Veröffentlichungen dazu, aber noch keine abschließende Empfehlung.

Meines Wissens - mein Wissensstand mag nicht korrekt oder veraltet sein - hat die EMA für sich keine entsprechende Empfehlung treffen können, ihren Assessoren irgendwelche konkreten Instrumente in die Hand zu geben, um eine solche Abwägung zu quantifizieren.

Lange Rede, kurzer Sinn: Wenn eine solche Institution, die schon so viele Jahrzehnte existiert, noch immer große Schwierigkeiten damit hat, würde ich um etwas Verständnis darum bitten, dass wir versuchen, uns damit Zeit zu lassen und dass wir - noch einmal - auf andere angewiesen sind, die auf diesem Gebiet gute Forschung durchführen und gute Hinweise geben, wie das gut genutzt werden kann.

Moderator Jürgen Windeler: Frau Wieseler.

Beate Wieseler: Ich möchte das ergänzen. Wir sind durchaus zu dieser Fragestellung mit der EMA in der Diskussion. Ich war am Samstag in einer Veranstaltung mit einem EMA-Vertreter zur Nutzenbewertung von Onkologika auf dem ESMO-Kongress, der dort auch noch einmal klargemacht hat, dass die EMA die komplexen Verfahren, die in diesem Methodenprojekt erarbeitet wurden, nicht in ihren Verfahren für einsetzbar hält und jetzt eine alternative Möglichkeit der Einbindung von Patienten in die Entscheidung in Pilotprojekten durchführen möchte.

Dazu gibt es eine Presseerklärung auf der Seite der EMA, die diesen Piloten und das Vorgehen der Patienteneinbindung in diese Entscheidungsfindung beschreibt. Wenn man sich das anschaut, ist das etwas, das sehr dem ähnelt, was im G-BA geschieht. Das ist eine Einbindung einzelner Organisationen und einzelner Patientenvertreter in Diskussionen auch bei der EMA ohne eine Stimme in der Entscheidung. Aber es ist, wenn Sie sich das

anschauen, ein ganz paralleles Verfahren, das die EMA jetzt zunächst einmal in Pilotprojekten ausprobieren möchte.

Moderator Jürgen Windeler: Gut. Wir haben uns darauf verständigt und geeinigt, dass das ein wichtiges Thema ist. Das lässt sich immer leicht sagen. Dass es ein schwieriges Thema ist, lässt sich auch immer leicht sagen. Nach allem, was Frau Wieseler und Herr Lange gerade beigetragen haben, wäre meine Prognose, dass wir in der grundsätzlichen Haltung: „Patientenpräferenzen sind wichtig, und es ist wichtig, sie einzubinden“ sofort alle Ja sagen und wir zu einem strukturierten, transparenten Prozess im optimalen Fall noch eine Weile brauchen. Wir - das sage ich für meine Person und damit dann auch für das Institut - werden uns darum bemühen, an dieser Stelle nicht lockerzulassen und unsere Anstrengungen und nicht - das sage ich sehr klar - Forschungsanstrengungen - das ist nicht unser Job - darauf zu richten, dass da etwas passieren muss und dass wir schauen werden, dass wir unsere Verfahren so nachvollziehbar und transparent wie möglich zu machen. Dafür wollen wir uns einsetzen. - Das war jetzt das Wort zum Ende des Punktes „Verschiedenes“.

Eine Meldung von Ihnen, und dann hat sich noch Herr Dintsios gemeldet.

Stefan Schulz: Wenn wir das eine Thema abgeschlossen haben, dann nur ein grundsätzlicher Punkt, der uns als unbefriedigender Bezug bisher im Methodenpapier 4.2 aufgefallen ist. Das sind die rechtlichen Grundlagen für die indikationsspezifische Effizienzgrenzenbestimmung. Das sucht das IQWiG auf § 12 und § 35 b zu stützen. Es gibt unserer Meinung nach - vielleicht kann das noch einmal beim weiteren Nachdenken über das Methodenpapier nachgeschärft werden - keine zwingende Stütze für die indikationsspezifische Wirtschaftlichkeitsbestimmung. Das hatten wir auch schon in der rechtswissenschaftlichen Diskussion, aber das ist noch nicht geklärt in diesem Papier. Vielleicht kann man darüber noch einmal nachdenken. Auch in § 92 SGB V - wir haben heute schon viel über das SGB V geredet - sehen wir darin keine Stütze.

Moderator Jürgen Windeler: Gut, wir nehmen das auf. - Da es zu diesem Punkt jahrelange Diskussionen mit nur dezent geänderter Rechtslage gegeben hat, würde ich jetzt vorschlagen, dass wir das auf- und Ihren Hinweis ernstnehmen, das jetzt aber nicht weiter vertiefen. - Herr Dintsios.

Ch.-Markos Dintsios: Zwei Punkte, einer rein technisch: Die Stellungnahmefristen für Kosten-Nutzen-Bewertungen, insofern sie umgesetzt werden, sind, glaube ich, mit 3 Wochen angesetzt. Wir hatten bei G09-01 damals sogar eine Verlängerung, weil es aufgrund der indirekten Vergleiche und der ganze Masse sehr umfassend war. Ich bitte zu bedenken, dass vielleicht auch bei weiteren Kosten-Nutzen-Bewertungen, sofern sie stattfinden, 3 Wochen wegen wiederum der extremen Masse zu kurz sind und man vielleicht 2 Monate spezifisch für diese Stellungnahmen ansetzen sollte.

Die Inputberichte - das waren IQWiG-eigene Produkte - zu den Antidepressiva waren bereits vom Umfang her enorm. So gesehen ist es innerhalb von 3 Wochen fast kaum möglich, pragmatisch betrachtet, eine Stellungnahme einzureichen.

Der zweite Punkt ist ein bisschen supportiv, zu dem, was vorhin der MSD-Vertreter gesagt hat. Die Diskussion ist vielleicht auch eine rechtliche. Da gebe ich Herrn Windeler völlig Recht. Aber das IQWiG selbst verwendet im Rahmen der frühen Nutzenbewertungsmatrix, wenn ich sie so nennen darf, auch etwas, das indikationsübergreifend verwendet wird - mit den entsprechenden Konfidenzintervallen, mit den entsprechenden impliziten Gewichten. Also ist diese Fragestellung - indikationsspezifisch oder indikationsübergreifend - gar nicht mal so fremd für das IQWiG. In Fragen der frühen Nutzenbewertung hat sich das IQWiG für eine indikationsübergreifende Lösung entschieden.

Andreas Gerber-Grote: Ich würde erst einmal zu dem einen Punkt der 3 Wochen antworten: Es ist so, dass laut Gesetz, wenn nach § 35 b SGB V im AMNOG-Verfahren eine Kosten-Nutzen-Bewertung angestrengt wird, die Frist zwischen Vor- und Abschlussbericht 3 Monate beträgt. Wenn wir Ihnen 2 Monate lang erlauben würden, eine Stellungnahme zu erarbeiten, hätten wir nur noch einen Monat, um das zu bearbeiten. Insofern ist das leider rechtlich bedingt. Auch hierbei müssten wir Sie im Grunde genommen darauf verweisen, auf den Gesetzgeber entsprechend einzuwirken. Wir haben uns das gut überlegt. Mehr als 3 Wochen sind da nicht drin. Denn wir müssten nach der Stellungnahme irgendwann eine Dokumentation und Würdigung schreiben. Deswegen ist das sehr knapp. Auch für uns ist das sehr knapp und sehr sportlich. Aber angesichts der in § 35 b gesetzten Fristen ist hieran leider im Moment nichts zu ändern.

Moderator Jürgen Windeler: Frau Wieseler.

Beate Wieseler: Zu „indikationsübergreifend“ bzw. „indikationsspezifisch“: Ich denke, sowohl in der Kosten-Nutzen-Bewertung als auch in der Dossierbewertung greifen wir indikationsspezifische Charakteristika der Erkrankung auf, die in die Bewertung eingehen - Endpunkte, die in dieser Indikation relevant sind. Dann wenden wir eine indikationsübergreifende Methodik an, also die Effizienzgrenze oder die Ausmaßfestlegung, wie wir es bei der Dossierbewertung tun. Darin würde ich keinen Widerspruch sehen.

Moderator Jürgen Windeler: Ich sehe keine weitere Wortmeldung mehr, auch wenn ich - das ist nicht abwertend gemeint - in die Ecken schaue. Dann bedanke ich mich sehr herzlich, dass Sie sowohl Ihre Stellungnahmen abgegeben haben, sich die Mühe gemacht haben, hoffentlich zur Verbesserung des Methodenpapiers beizutragen. Vielen Dank, dass Sie hier waren. Vielen Dank für die überwiegend lebhafteste Diskussion. Ich wünsche Ihnen allen einen guten Heimweg. Wir werden uns jetzt an die Endfassung der Methoden machen. - Vielen Dank und gute Fahrt.

Anhang A – Dokumentation der Stellungnahmen

Inhaltsverzeichnis

	Seite
A.1 – Stellungnahmen von Organisationen, Institutionen und Firmen	A 3
A.1.1 – Allianz Chronischer Seltener Erkrankungen (ACHSE e. V.)	A 3
A.1.2 – Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF)	A 6
A.1.3 – AstraZeneca GmbH.....	A 12
A.1.4 – Boehringer Ingelheim Pharma GmbH & Co. KG.....	A 21
A.1.5 – Bundesarbeitsgemeinschaft Selbsthilfe von Menschen mit Behinderung und chronischer Erkrankung und ihren Angehörigen e. V. (BAG Selbsthilfe)	A 31
A.1.6 – Bundesverband der Pharmazeutischen Industrie e. V. (BPI)	A 36
A.1.7 – Bundesverband Medizintechnologie e. V. (BVMed)	A 44
A.1.8 – Deutsche Diabetes Gesellschaft e. V. (DDG) und Deutsche Gesellschaft für Kardiologie e. V. (DGK).....	A 50
A.1.9 – Deutsche Gesellschaft für Gesundheitsökonomie e. V. (DGGÖ)	A 53
A.1.10 – Deutsche Gesellschaft für Gynäkologie und Geburtshilfe e. V. (DGGG) ..	A 57
A.1.11 – Deutsche Gesellschaft für Hämatologie und medizinische Onkologie e. V. (DGHO) und Gesellschaft für Thrombose- und Hämostasenforschung e. V. (GTH).....	A 60
A.1.12 – Deutsche Gesellschaft für Hygiene und Mikrobiologie e. V. (DGHM).....	A 67
A.1.13 – Deutsche Gesellschaft für Innere Medizin e. V. (DGIM).....	A 71
A.1.14 – Deutsche Gesellschaft für Orthopädie und Unfallchirurgie e. V. (DGOU) ..	A 74
A.1.15 – Deutsche Gesellschaft für Psychoanalyse, Psychotherapie, Psychosomatik und Tiefenpsychologie e. V. (DGPT).....	A 88
A.1.16 – Deutsche Gesellschaft für Ultraschall in der Medizin e. V. (DEGUM)	A 93
A.1.17 – Deutsche Krankenhausgesellschaft e. V. (DKG)	A 96
A.1.18 – Deutscher Psoriasis Bund e. V. (DPB)	A 102
A.1.19 – Deutsches Netzwerk Versorgungsforschung e. V. (DNVF)	A 105
A.1.20 – Ecker + Ecker GmbH.....	A 109
A.1.21 – GlaxoSmithKline GmbH & Co. KG (GSK)	A 118
A.1.22 – Institut für Gesundheitsökonomik (IfG)	A 126
A.1.23 – MArS Market Access & Pricing Strategy GmbH und AHEAD GmbH ..	A 132
A.1.24 – MSD Sharp & Dohme GmbH.....	A 136
A.1.25 – Novartis Pharma GmbH.....	A 143
A.1.26 – Sanofi-Aventis Deutschland GmbH.....	A 153

A.1.27 – Verband Forschender Arzneimittelhersteller e. V. (vfa)	A 163
A.1.28 – YES Pharmaceutical Development Services GmbH	A 198
A.2 – Stellungnahmen von Privatpersonen	A 200
A.2.1 – Kohlmann, Thomas	A 200
A.2.2 – Mühlbacher, Axel	A 202
A.2.3 – Vach, Werner	A 239

A.1 – Stellungnahmen von Organisationen, Institutionen und Firmen

A.1.1 – Allianz Chronischer Seltener Erkrankungen (ACHSE e. V.)

ACHSE Lotse an der Charité | Universitätsmedizin Berlin | Campus
Virchow | Augustenburger Platz 1 | 13353 Berlin

Institut für Qualität und
Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8 (KölnTurm)
50670 Köln

ACHSE e.V.
c/o DRK Kliniken Berlin | Mitte
Drontheimer Straße 39
13359 Berlin

ACHSE Lotse an der Charité
Berlin Campus Virchow
Augustenburger Platz
13353 Berlin

Berlin, 6. August 2014

**Entwurf der „Allgemeinen Methoden Version 4.2.“ des
IQWiG v.18.6.2014 - Stellungnahme**

www.achse-online.de
www.achse.info

Schirmherrin:
Eva Luise Köhler

Ihr Ansprechpartner:
Dr. med. M.S. Christine Mundlos

Sehr geehrte Damen und Herren,
namens und im Auftrag der Allianz Chronischer Seltener Erkrankungen (ACHSE) e.V.
nehmen wir zu dem Entwurf der „Allgemeinen Methoden Version 4.2.“ des IQWiG
v.18.6.2014 wie folgt Stellung:

Als Dachverband der Patientenorganisationen von Menschen mit Seltene n Erkrankungen
und ihrer Angehörigen nehmen wir mit Sorge zu Kenntnis, dass in Abschnitt 3.1.1. die
Patientenzufriedenheit als ergänzende Zielgröße und im Abschnitt 3.2.4. die
Behandlungszufriedenheit als patientenberichteter Endpunkt (PRO) gegenüber der
Vorversion gestrichen wurden.

Diese wesentlichen Änderungen erfolgen ohne jegliche Begründung. Dies muss schon
deshalb erstaunen, weil das IQWiG sicher sonst im Rahmen eines HTA-Verfahrens eine
unbegründete Methodenänderung durch Dritte beispielsweise bei einer klinischen Studie
zu Recht als fehlerbehaftet ansehen würde. Wir fragen deshalb: Warum wird diese
Änderung vorgenommen und aufgrund welcher geänderten Evidenzlage erfolgt sie?

Wir konzedieren – im Einklang mit der Version 4.1. – dass die Patientenzufriedenheit alleine nicht als patientenrelevanter Endpunkt dienen kann. Sie kann aber sehr wohl eine ergänzende Information darstellen und insofern auch einen Hinweis auf eine möglicherweise höhere Lebensqualität geben.

Eine Veränderung der Behandlungszufriedenheit kann ebenfalls zusätzliche Hinweise auf die subjektive Akzeptanz der Behandlung geben. Gerade weil jede medizinische Intervention letztlich daran gemessen werden muss, „wie ein Patient fühlt, seine Funktionen und Aktivitäten wahrnehmen kann oder ob er überlebt“ (vgl. S.38 des Dokuments) ist es vollkommen unverständlich, dass dann Patienten- und Behandlungszufriedenheit als PRO wegfallen sollen.

Wir schlagen daher vor, diese Streichungen nicht vorzunehmen.

Mit freundlichen Grüßen

Christoph Nachtigäller
Vorsitzender ACHSE e.V.

Dr. Andreas Reimann
Stellv. Vorsitzender ACHSE e.V.

In Vertretung


Dr. Christine Mundlos

**A.1.2 – Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
e. V. (AWMF)**

Arbeitsgemeinschaft der	<i>Association of the</i>
Wissenschaftlichen	<i>Scientific</i>
Medizinischen	<i>Medical</i>
Fachgesellschaften e.V.	<i>Societies in Germany</i>



Stellungnahme der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften

zum Entwurf des IQWiG–Methodenpapiers:
„Allgemeine Methoden“ Version 4.2 vom 18.06.2014“

Die AWMF wurde am 10.07.2014 um eine Stellungnahme zum oben genannten Entwurf gebeten. Sie hat ihrerseits ihre Mitgliedsfachgesellschaften gebeten, bei gesehendem Bedarf eine eigene Stellungnahme zu verfassen. Die bis einschließlich 01.08.2014 bei der AWMF eingegangenen Stellungnahmen der DGGG, dggö, DGHO/GTH, DGOU sowie der DEGUM sind der Stellungnahme der AWMF als Anlage beigefügt. Es ist nicht auszuschliessen, dass sich andere Fachgesellschaften wegen der Kürze der Antwortzeit direkt an das IQWiG gewandt haben. Ebenso beigefügt ist die Stellungnahme des Deutschen Netzwerks für Versorgungsforschung (DNVF), mit dem eine Kooperation seitens der AWMF besteht. Die wichtigsten Aspekte der vorliegenden Stellungnahmen wurden in die Stellungnahme des AWMF e.V. eingearbeitet.

1. Allgemeine Anmerkungen

Die AWMF begrüßt die Fortschreibung des Methodenpapiers des IQWiG V. 4.2 als Maßnahme zur öffentlichen transparenten Darstellung der Arbeitsweise des Instituts und nimmt gerne die Möglichkeit der Stellungnahme war.

Wir möchten einfürend auf die bereits 2011 geäußerten Verbesserungsvorschläge zur frühzeitigen Einbindung des Sachverstands von Vertretern der Wissenschaftlichen Medizinischen Fachgesellschaften und der Patientenorganisationen hinweisen, die aus unserer Sicht in alle Verfahrensabläufe des IQWiG ohne produktspezifische Ausnahmen integriert werden sollten ¹.

Aus der Sicht der Fachgesellschaften ist, wie schon mehrmals daraufhingewiesen, eine Stellungnahmefrist von sechs Wochen erforderlich, um für beide Seiten zufriedenstellende, qualifizierte Stellungnahmen zu gewährleisten (siehe Stellungnahmen der DEGUM, DGGG, DGHO/GTH, DGOU).

Das ausschließliche Einholen externer Expertisen mit Hilfe spezieller Fragebogen, wie vom IQWiG angewandt, erscheint vielen Fachgesellschaften unzureichend und bedarf dringend einer Evaluierung. Die von vielen Fachgesellschaften präferierte mündliche Erörterung könnte sogar zu anderen Ergebnissen führen.

¹ Stellungnahme der AWMF zum Entwurf des Methodenpapiers Version 4 des IQWiG vom 08.03.2011 zum Aspekt: Produktspezifische Verfahrensabläufe. Verfügbar: http://www.awmf.org/fileadmin/user_upload/Stellungnahmen/Medizinische_Versorgung/Stellungnahme_IQWiG-Methoden_IV.pdf

Im Einzelnen bezieht die AWMF wie folgt Stellung:

2. Zu den geänderten Abschnitten „Kosten-Nutzen-Bewertung nach §35 SGBV“ (Abschnitt 2.1.4, 3.15 und neues Kapitel 4)

Die AWMF trägt hinsichtlich der Kosten-Nutzen-Bewertung ausdrücklich die Ansicht der dggö in Bezug auf den Auftrag des IQWiG mit. Aus Sicht der AWMF sollte das IQWiG über eine Bewertung der wissenschaftlichen Evidenz zur Kosten-Nutzenbewertung (Assessment) für ein Thema hinaus lediglich – optimal mehrere - Vorschläge zu deren Interpretation im Hinblick auf die Wirtschaftlichkeit des Ergebnisses (Appraisal) unterbreiten.

Die AWMF teilt auch die Bedenken der dggö im Hinblick auf die Eignung des vom IQWiG weiterhin favorisierten Effizienzgrenzenkonzepts und sieht die Auseinandersetzung und Präzisierung des Konstrukts „Lebensqualität“ als wichtigen Baustein für eine solide methodische Arbeit. Letzteres ist in der Version 4.2. des Methodenpapiers nur ansatzweise gegeben.

Eine Stellungnahmefrist von drei Wochen bei Kosten-Nutzen-Bewertungen kann aus der Sicht der Fachgesellschaften der Komplexität des Themas nicht gerecht werden - siehe allgemeine Anmerkungen.

3. Zu dem ergänzten Produkt „Potenzialbewertung“ gemäß § 137e SGB V (in den Abschnitten 1.1, 2.1, 2.2, 3.8, 7.2 (vorher 6.2), 8.1.3 (vorher 7.1.3) und 8.1.4 (vorher 7.1.4))

Entsprechend unserer früheren Stellungnahme halten wir auch für das neue Produkt „Potenzialbewertung“ des IQWiG eine regelhafte Hinzuziehung externen medizinisch-wissenschaftlichen Sachverständigen für erforderlich. Zudem sollten die Struktur und die Prozesse der internen Qualitätssicherung beschrieben werden.

Die AWMF sieht bei der Potenzialbewertung eine zusätzliche Verpflichtung des IQWiG gegenüber den Versicherten, höchste fachlich-wissenschaftliche und klinische Expertise sicherzustellen. Dies ist umso dringlicher geboten, als die Ergebnisse der Potenzialbewertungen vor der Bekanntgabe der Erprobungsrichtlinien durch den G-BA der Öffentlichkeit nicht bekannt sind und auch methodisch nicht kommentiert werden können. Zur Herstellung von Transparenz erachtet die AWMF auch eine Veröffentlichung der IQWiG-Berichte, die zu einer Ablehnung von Anträgen zur Erprobung im G-BA geführt haben, für äußerst wünschenswert. Hier muss definitiv von Transparenzlücken gesprochen werden.

4. Zu dem geänderten Abschnitten „Gesundheitsinformationen“ (2.1.7 und Kapitel 6, vorher Kapitel 5)

Die AWMF begrüßt die methodisch hochwertige Erstellung von Gesundheitsinformationen. Angesichts der vielfältigen Anbieter erscheint es sinnvoll, ein übergeordnetes Portal einzurichten, in dem die Gesundheits- bzw. Patienteninformationen der verschiedenen Anbieter abrufbar sind.

Weiterhin scheint eine thematische Abstimmung in Bezug auf Informationen aus weiteren Quellen (z.B. der externen Qualitätssicherung) erforderlich, wie sie z.B. durch das neue Institut für Qualitätssicherung nach §137a SGB V vorgesehen sind². Auch der Ansatz der zielgruppenspezifischen Umsetzung sollte weiter intensiviert werden.

5. Zu Änderungen des Abschnitts 2.2.1 „Auswahl externer Sachverständiger“

Die AWMF empfiehlt, die Auswahl und Beteiligung externer Sachverständiger im Hinblick auf klinisch-wissenschaftlichen und praktisch-medizinischen Sachverstand zu überdenken und verweist auf die Stellungnahmen der DGHO/GTH, DGOU und DEGUM. Der eingebrachte klinisch-wissenschaftliche und der praktisch-medizinische Sachverstand zum bearbeiteten Thema sollte einschließlich seiner Quellen zuverlässig sichergestellt sein.

Dass die Namen aller beteiligter Sachverständiger genannt werden, entspricht dem aktuellen Urteil des Oberverwaltungsgerichts Nordrhein-Westfalen vom 15.01.2014³ und ist zu begrüßen.

6. Änderungen zum Thema Patientenzufriedenheit in den Abschnitten 3.1.1 und 3.2.4

Die AWMF bedauert die fehlende Darlegung von inhaltlichen Gründen für das Streichen des Aspektes Patientenzufriedenheit bzw. Zufriedenheit mit der Behandlung und kann dieses daher in dem Methodenpapier nicht schlüssig nachvollziehen. Bekanntlich wird an anderen Stellen des SGB V und der Richtlinien des G-BA der Patientenzufriedenheit und –beteiligung eine zunehmend wichtigere Rolle zugeschrieben. Dass die methodischen Grundlagen für eine qualitative hochstehende Messung der Patientenzufriedenheit (noch) nicht weit verbreitet sind, ist kein Grund, von dem politischen Ziel der Wichtigkeit der Patientenzufriedenheit als Steuerungskriterium im Gesundheitswesen abzurücken. Hier ist das IQWiG gefragt.

Zur Bedeutung von Patientenzufriedenheit im Sinne der Beurteilung des Erlebten liegen positive Studien vor (siehe auch Stellungnahme des DNVF).

² siehe GKV-Finanzstruktur- und Qualitäts-Weiterentwicklungsgesetz, § 137a „Umsetzung der Qualitätssicherung und Darstellung der Qualität“ (3) Satz 4, http://www.gesetze-im-internet.de/sgb_5/_137a.html [31.07.2014]

³ Quelle: <http://openjur.de/u/676333.html> [31.07.2014]

7. Neufassung von Abschnitt 3.5 zu diagnostischen Verfahren mit Integration des alten Abschnitts 3.8 zu Prognosestudien

Die AWMF erkennt das Bemühen des IQWiG an, methodisch begründet das Durchführen nicht zielführender diagnostischer Tests zu vermeiden. Der im Methodenpapier im Abschnitt 3.5 einleitend formulierte Satz „Die alleinige Gewinnung diagnostischer Informationen (ohne therapeutische oder präventive Konsequenzen) hat regelhaft keinen sozialrechtlich relevanten Nutzen“ ist allerdings in dieser Verkürzung auch unter Berücksichtigung der angegebenen Literaturstelle (D. Roters, G-BA⁴) nicht ganz zutreffend.

Diagnostische Verfahren haben selbstverständlich einen sozialrechtlich relevanten Nutzen, wenn sie der Differentialdiagnostik und der Sicherung einer Diagnose nach der internationalen statistischen Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD) der WHO dienen.

Im Erkenntnisgewinn allein kann zudem ein individueller Nutzen für den Patienten liegen. Der Aufklärung der zu untersuchenden Person über Eigenschaften des diagnostischen Tests, Nebenwirkungen und klinische Konsequenzen kommt dabei höchste Priorität zu. Erklärtes Ziel muss auch aus der Sicht der AWMF die Vermeidung von Überdiagnostik sein.

Die Abhandlung diagnostischer und prognostischer Verfahren in Abschnitt 3.5 folgt in der Reihenfolge nicht der dort erläuterten Hierarchie der Studien und ist deshalb nicht gut nachvollziehbar. Bevor ein klinisch relevanter Nutznachweis geführt werden kann, sind die Testgüteeigenschaften des neuen Tests im Vergleich zum bisherigen Standard in geeigneten Studien an einem geeigneten Kollektiv zweifelsfrei festzustellen (Accuracy: Testgüteparameter und Nebenwirkungen sowie Zugewinn an Diagnosesicherheit bei Anwendung mehrerer Tests). Diese Voraussetzung wird im Kapitel erst nach der Besprechung von Studien im randomisierten Design zur klinischen Nutzenbewertung adressiert. Möglichkeiten und Grenzen von „Linked evidence“ sollten insgesamt ausführlicher dargestellt werden, da vielfach keine entsprechenden RCT vorliegen.

Gesundheitssystemrelevante Aspekte der Wirtschaftlichkeit eingesetzter diagnostischer und prognostischer Tests sollten zusätzlich und explizit adressiert werden.

8. Ergänzung zum Umgang mit unaufgefordert übermittelten Daten in Kapitel 7 (vorher Kapitel 6)

Die Eingabe von Daten zu unerwünschten Effekten bzw. Gefährdungen von Patienten muss möglich sein.

⁴ Roters D. "Recht auf Wissen": haben gesetzlich Krankenversicherte Anspruch auf diagnostische Leistungen ohne medizinische Konsequenzen? Z Evid Fortbild Qual Gesundheitswes 2011; 105(7): 526-530.

9. Ergänzung zu patientenrelevanten Endpunkten in Abschnitt 8.3.3 (vorher 7.3.3)

Eine substantielle Überarbeitung der Ausführungen zum Thema „Patientenrelevante Endpunkte“ im Methodenpapier steht aus Sicht der AWMF weiterhin aus. Wir verweisen hier auch auf die Stellungnahmen der DGHO/GTH und des DNVF. Insbesondere die Herausforderung multidimensionaler Entscheidungsfindung ist nicht ausreichend, nachvollziehbar und operational adressiert.

10. Im Methodenpapier nicht aktualisierte und zu kommentierende inhaltliche Aspekte

Validierung der Operationalisierungskriterien für das Ausmass des Nutzens

Für die in den beiden Tabellen 16 und 17 aufgeführten Kriterien steht eine Validierung z.B. hinsichtlich der Gewichtung aus. Äußerungen wie „Die in Tabelle 16 veranschlagte Rasterung der Schwellenwerte erscheint vernünftig und praktikabel“ sind nicht ausreichend belegt und entsprechen nicht dem selbstgesetzten Niveau der Argumentation des IQWiG.

Subgruppenanalysen

Die Berechnung des Zusatznutzens für Subgruppen von Patienten erfolgt durch das IQWiG auf der Basis von Ergebnissen der Verzerrungsanalyse, ohne dass die Power der jeweiligen Studien zur Detektion von Unterschieden in Subgruppen ausreicht. Zur Belegung von Subgruppeneffekten sollten ausreichend große Datenbasen vorliegen, ansonsten sind die Ergebnisse statistisch nicht aussagefähig.

Zur Diskussion und für eventuelle Rückfragen zu unseren Kommentaren stehen wir gern zur Verfügung.

Ansprechpartner/Kontakt:

Dr. med. Monika Nothacker, MPH, [REDACTED]

Prof. Dr. med. Ina Kopp, [REDACTED]

Prof. Dr. Hans-Konrad Selbmann, [REDACTED]

Anlagen:

Stellungnahmen von

- Deutsche Gesellschaft für Gynäkologie und Geburtshilfe (DGGG) vom 31.07.2014
- Deutsche Gesellschaft für Gesundheitsökonomie (dggö) vom 28.07.2014,
- Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie/Gesellschaft für Thrombose- und Hämostaseforschung (DGHO/GTH) vom 04.08.2014,
- Deutsche Gesellschaft für Orthopädie und Unfallchirurgie (DGOU) vom 21.07.2014
- Deutsche Gesellschaft für Ultraschall in der Medizin (DEGUM) vom 30.07.2014
- Deutsches Netzwerk Versorgungsforschung (DNVF) vom 29.07.2014

A.1.3 – AstraZeneca GmbH

Schriftliche Stellungnahme zu IQWiG Allgemeine Methoden, Entwurf für Version 4.2 vom 18.06.2014

1. Transparenz bei der Beratung durch den wissenschaftliche Beirat des Instituts

IQWiG Allgemeine Methoden

S.22 / 2.1.4 Kosten-Nutzen-Bewertung nach § 35b SGB V

Der Ablauf der Kosten-Nutzen-Bewertung nach § 35b SGB V ist in Abbildung 4 schematisch dargestellt. Alle Arbeitsschritte werden in Verantwortung des Instituts getätigt. Dabei werden regelhaft externe Sachverständige beteiligt. Bei Bedarf wird auch der wissenschaftliche Beirat des Instituts einbezogen. Der interne Qualitätssicherungsprozess ist in diesem Flussdiagramm nicht dargestellt.

Position AstraZeneca

Wiederum zur Erhöhung der Transparenz, sollte klarer definiert werden in welchen Fällen der wissenschaftliche Beirat des Instituts grundsätzlich einbezogen werden soll, kann respektive nicht soll und dargelegt werden welchen Einfluss diese Beratung auf die einzelne Bewertung hatte.

Vorgeschlagene Änderung:

Die vorangehend dargelegte methodische Darstellung bzw. Publikation der Umstände bzw. Inhalte der Beratung durch den wissenschaftlichen Beirat des Instituts sollte bei allen davon betroffenen Produkten des Instituts umgesetzt werden.

2. Mündliche Anhörung bei Kosten-Nutzen-Bewertungen

IQWiG Allgemeine Methoden

S.23 / 2.1.4 Kosten-Nutzen-Bewertung nach § 35b SGB V / Fußnote unter dem Flussdiagramm

Die Anhörung erfolgt mittels Einholung schriftlicher Stellungnahmen. Optional wird eine mündliche wissenschaftliche Erörterung zur Diskussion unklarer Aspekte in den schriftlichen Stellungnahmen durchgeführt.

Position AstraZeneca

Ungleich zum Prozess zur Dossierbewertung ist hier eine mündliche wissenschaftliche Erörterung im Rahmen des Stellungnahmeverfahrens nur optional vorgesehen. Gemäß der Verfahrensordnung des G-BA ist jeder, der gesetzlich berechtigt ist, zu einem Beschluss des Gemeinsamen Bundesausschusses Stellung zu nehmen, und eine schriftliche Stellungnahme abgegeben hat, auch zu einer mündlichen Stellungnahme berechtigt.

Vorgeschlagene Änderung:

Im Rahmen der Kosten-Nutzen-Bewertung wird regelmäßig eine mündliche wissenschaftliche Erörterung zur Diskussion unklarer Aspekte in den schriftlichen Stellungnahmen durchgeführt.

3. Transparenz bei der Beratung durch externe Sachverständige

IQWiG Allgemeine Methoden

S.25/ 2.1.5 Potenzialbewertung

Der Ablauf der Erstellung einer Potenzialbewertung ist in Abbildung 5 schematisch dargestellt. Alle Arbeitsschritte werden in Verantwortung des Instituts getätigt. Dabei können externe Sachverständige beteiligt werden (siehe Abschnitt 2.2.1). Der interne Qualitätssicherungsprozess ist in diesem Flussdiagramm nicht dargestellt.

Position AstraZeneca

Im Sinne der Transparenz und Nachvollziehbarkeit ist es erforderlich, im Methodenpapier die Sachverhalte genau zu konkretisieren, an denen externe Sachverständige beteiligt werden müssen, können bzw. nicht beteiligt werden. Darüber hinaus sind sowohl der Fragenkatalog, als auch die entsprechenden Antworten und Ausführungen des Sachverständigen zu veröffentlichen. Dabei sollte aus dem Abschlussbericht des Instituts der Einfluss dieser externen Beratung auf die Bewertung sowie die Begründung einer allfälligen Diskrepanz zwischen Bewertung und externer Beratung hervorgehen. Um die Unabhängigkeit der Sachverständigen zu wahren, ist die namentliche Nennung zweitrangig.

Vorgeschlagene Änderung:

Die vorangehend dargelegte methodische Darstellung bzw. Veröffentlichung der Inhalte der externen Beratung durch Sachverständige sollte sowohl bei den Produkten Kosten-Nutzen-Bewertung und Potenzialbewertung als auch bei den Institutsprodukten Berichte, Rapid Report und Dossierbewertung umgesetzt werden.

4. Potentialbewertung

IQWiG Allgemeine Methoden

S.26/ 2.1.5 Potenzialbewertung

Den Abschluss des Prozesses bildet die Erstellung der Potenzialbewertung. Nach § 137e SGB V muss der G-BA innerhalb von 3 Monaten eine Entscheidung zum Potenzial der beantragten Untersuchungs- oder Behandlungsmethode treffen. Daher werden Potenzialbewertungen durch das Institut regelmäßig binnen 6 Wochen fertiggestellt. Nach Fertigstellung wird die Potenzialbewertung dem G-BA zugeleitet. Eine Publikation der Potenzialbewertung erfolgt nicht, weil das Antragsverfahren gemäß § 137e SGB V der hohen Vertraulichkeit unterliegt. Erst wenn im weiteren Verlauf der G-BA eine Erprobungsrichtlinie erlässt, wird auch die Potenzialbewertung veröffentlicht.

S.74/3.8 Potenzialbewertung

Wenn aus Sicht des Instituts ein Potenzial feststellbar ist, wird die vom Antragsteller vorgeschlagene Erprobungsstudie geprüft; sofern der Antrag keinen solchen oder einen ungeeigneten Vorschlag enthält, konkretisiert das Institut die Eckpunkte einer möglichen Erprobungsstudie.

Position AstraZeneca

Wir stimmen zu, dass diese Art des Verfahrens einer hohen Vertraulichkeit unterliegt und gemäß § 137e SGB V die Publikation der gesamten Potentialbewertung nicht sogleich erfolgt. Nichtsdestotrotz ist eine Veröffentlichung der grundsätzlichen Methodik, der vordefinierten Kriterien für eine positive Potenzialbilanz, sowie die Bewertung der Erprobungsstudie in ihren Grundsätzen erforderlich.

Weiterhin ist ein aktiver mündlicher Austausch zwischen dem IQWiG und dem pharmazeutischen Unternehmer (Antragssteller) im Rahmen der Beurteilung und Konkretisierung der Erprobungsstudie unerlässlich. Diese wissenschaftlich geführte Diskussion soll die Basis schaffen für die zukünftige Optimierung der klinischen Forschung und die Bewertung des Nutzens des neuen Arzneimittels / der neuen Methode auf ein ausreichend sicheres Erkenntnisniveau heben.

Vorgeschlagene Änderung:

Publikation der Methodik und zugrundeliegende Kriterien der spezifischen klinischen Fragestellung unter Wahrung der hohen Vertraulichkeit.

Die Basis der Beurteilung und Konkretisierung der Erprobungsstudie bildet der aktive mündliche Austausch zwischen dem Institut und dem pharmazeutischen Unternehmen und dessen Protokollierung.

5. Regelhaftes Stellungnahmeverfahren zu Addenden

IQWiG Allgemeine Methoden

Das Verfahren zur Veröffentlichung eines Addendums orientiert sich an dem des ursprünglichen Institutsprodukts.

Position AstraZeneca

Es ist unverständlich, dass sich das Verfahren zur Veröffentlichung eines Addendums an dem des ursprünglichen Institutsproduktes orientiert, nicht jedoch an dem Verfahren zur Stellungnahme. Laut der Verfahrensordnung Kapitel 5 Abschnitt 3 § 19 wird mit der Veröffentlichung der der Bewertung einem definierten Personenkreis die Gelegenheit gegeben schriftlich Stellung zu nehmen. Im Anschluss an das schriftliche Stellungnahmeverfahren und vor einer Beschlussfassung dürfen die Stellungnahmeberechtigten welche von ihrem schriftlichen Recht Gebrauch gemacht haben auch mündlich Stellung nehmen.

Vorgeschlagene Änderung:

Das Verfahren zur Stellungnahme (schriftlich wie mündlich) eines Addendums soll sich an dem des ursprünglichen Institutsprodukts orientieren und somit z. B. auch bei einem

Addendum zur Nutzenbewertung für eine schriftliche sowie mündliche Stellungnahme Gelegenheit gegeben werden.

6. Patientenzufriedenheit soll valider patientenberichteter Endpunkt bleiben

IQWiG Allgemeine Methoden

S.39/ 3.1.1 Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens

Ergänzend ~~können~~ kann der interventions- und erkrankungsbezogene Aufwand ~~und die Zufriedenheit der Patientinnen und Patienten~~ mit der Behandlung berücksichtigt werden.

S. 54/ 3.2.4 Patientenberichtete Endpunkte

Neben der gesundheitsbezogenen Lebensqualität ~~und der Behandlungszufriedenheit~~ können auch andere Nutzendimensionen mittels PRO erfasst werden, wie z. B. Symptome der Erkrankung.

Position AstraZeneca

Die ersatzlose Streichung der Patientenzufriedenheit als patientenberichteter Endpunkt widerspricht dem patientenzentrierten Ansatz der Nutzenbewertung. Die Patientenzufriedenheit ist ein wesentlicher Aspekt für Therapie-Compliance respektive Adhärenz. Gerade bei bestimmten Patienten-Gruppen wie z. B. bei Kindern und älteren Patienten ist die Zufriedenheit mit der Therapie und der Anwendung einer Medikation - z. B. Verabreichungsform und Alltagstauglichkeit- ganz wichtige Aspekte. Bei Beachtung der methodischen Limitationen kann die Patientenzufriedenheit ein valider ergänzender patientenrelevanter Endpunkt darstellen.

Vorgeschlagene Änderung:

Die Patientenzufriedenheit als ergänzender patientenrelevanter Endpunkt bleibt bestehen.

7. Diagnostische Verfahren

IQWiG Allgemeine Methoden

S.66/ 3.5 Diagnostische Verfahren

Insgesamt entscheidend ist weniger, inwieweit eine diagnostische oder prognostische Information einen aktuellen oder zukünftigen Gesundheitszustand feststellen kann, sondern dass diese Information auch prädiktive Bedeutung hat, also den höheren (oder geringeren) Nutzen einer Folgebehandlung vorhersagen kann [184,498].

Position AstraZeneca

Die prädiktive Bedeutung einer diagnostischen oder prognostischen Information ist in Bezug auf den höheren (oder geringeren) Nutzen einer Folgebehandlung durchaus als wichtig zu erachten. Allerdings kann die Information über einen aktuellen oder zukünftigen Gesundheitszustand auch unabhängig von der Vorhersagbarkeit des Nutzens einer Folgebehandlung für den einzelnen Patienten alleinstehend von großer Relevanz sein. Eine

derart explizite Gewichtung der Relevanz sollte daher an dieser Stelle aus unserer Sicht nicht erfolgen.

Vorgeschlagene Änderung:

Auf die zwingende Notwendigkeit einer prädiktiven Bedeutung der diagnostischen oder prognostischen Information in Bezug auf den Nutzen einer Folgebehandlung ist zu verzichten.

8. Kosten-Nutzen-Bewertung medizinischer Interventionen - Rechtliche Grundlagen

IQWiG Allgemeine Methoden

S.76/4.1.1 Rechtliche Grundlagen für die Kosten-Nutzen-Bewertung nach SGB V

Das BMG nahm im Jahr 2008 zum Stellenwert von KNB in Berufung auf § 12 SGB V, dem sog. Wirtschaftlichkeitsgebot, folgendermaßen Stellung: „Maßgebend für die Verordnungsentscheidung des Arztes kann nur sein, ob das Arzneimittel zur Erreichung des Therapieziels, d. h. im jeweiligen Anwendungsgebiet, zweckmäßig und wirtschaftlich ist (§ 12 SGB V [78]).“ Daher muss die KNB einen Vergleich der Effizienz von Interventionen in einem Indikationsgebiet ermöglichen, ohne die Frage einer indikationsübergreifenden Priorisierung im Gesundheitssystem zu beantworten.

Position AstraZeneca

Eine Kosten-Nutzen-Bewertung sollte analog zur frühen Nutzenbewertung gemäß § 35a SGB V wirkstoff- und nicht indikationsbezogen erfolgen. Für Arzneimittel, die für mehrere Indikationen zugelassen sind, kann hierdurch eine indikationsübergreifende Bewertung gewährleistet werden.

Vorgeschlagene Änderung:

Auf eine ausschließliche Limitierung auf einzelne Indikationsgebiete sollte verzichtet werden.

9. Kosten-Nutzen-Bewertung medizinischer Interventionen - Datengrundlage

IQWiG Allgemeine Methoden

S.78/ 4.1.7 Datengrundlage

Werden Leitlinien verwendet, sollten diese aus dem deutschen Gesundheitssystem stammen und möglichst evidenzbasiert sein. Darunter werden Leitlinien verstanden, deren Empfehlungen auf einer systematischen Literaturrecherche beruhen, grundsätzlich mit einer Evidenz- und / oder Empfehlungseinstufung (Level of Evidence [LoE] und / oder Grade of Recommendation [GoR]) versehen sind und mit den Referenzen der ihnen zugrunde liegenden Primär- und / oder Sekundärliteratur verknüpft sind (modifiziert nach AGREE).

Position AstraZeneca

Es ist nicht nachvollziehbar, dass hier ausschließlich Leitlinien aus dem deutschen Gesundheitssystem und deren korrekte Verwendung berücksichtigt werden. Zum einen ist

die Möglichkeit zur Verwendung deutscher Leitlinien je nach Indikation stark eingeschränkt, da bspw. keine existieren, die vorhandenen Leitlinien keine Empfehlungseinstufung enthalten oder deren Empfehlungen nicht mehr dem aktuellen Versorgungskontext entsprechen. Dem gegenüber stellt die gleichwertige Berücksichtigung internationaler Leitlinien mit hoher Güte oft eine bessere Alternative dar. Zumal dies nicht ausschließlich für die frühe Nutzenbewertung so gehandhabt wird, sondern auch im ärztlichen Alltag.

Vorgeschlagene Änderung:

Das Institut nimmt die Möglichkeit der Berücksichtigung internationaler Leitlinien bei der Bestimmung der zu verwendenden Datengrundlage mit auf.

10. Kosten-Nutzen-Bewertung medizinischer Interventionen - Interpretation der Ergebnisse

IQWiG Allgemeine Methoden

S.81/ 4.1.9 Interpretation der Ergebnisse

Ist die Bestimmung mehrerer Effizienzgrenzen für die Bewertung einer Intervention erforderlich, so bleibt es dem Entscheidungsträger unter Beachtung der Relevanz der patientenrelevanten Endpunkte vorbehalten, eine Gewichtung vorzunehmen.

Position AstraZeneca

Es ist unklar, wie genau die hier genannte Gewichtung unter Beachtung der Relevanz der patientenrelevanten Endpunkte zu erfolgen hat. Nach dem derzeitigen Wortlaut sind keine definierten Rahmenbedingungen für die Durchführung einer Gewichtung gegeben, so dass die Art der Durchführung einzig dem Entscheidungsträger obliegt. Aus unserer Sicht sollte die Herangehensweise bei der Gewichtung klar geregelt sein und eine Durchführung nach einheitlichen Kriterien oder einem Algorithmus ermöglichen.

Vorgeschlagene Änderung:

Das Institut definiert für den Entscheidungsträger eine klare Vorgehensweise, anhand welcher genauen Kriterien wie zu gewichten ist.

11. Kosten-Nutzen-Bewertung medizinischer Interventionen - Maß des Gesamtnutzens

IQWiG Allgemeine Methoden

S.87/ 4.3.3 Maß des Gesamtnutzens

Die Ergebnisse sollten dem Entscheidungsträger zusammen mit der Endpunktgewichtung zur Verfügung gestellt werden. Hierdurch ergibt sich für den Entscheidungsträger die Option, einen aus mehreren zusatznutzenbereinigten Erstattungspreisen gewichteten Erstattungsbetrag zu verhandeln.

Position AstraZeneca

Wie in Punkt 10 bereits erläutert ist auch hier unklar, wie genau die hier genannte Gewichtung erfolgt ist. Zudem ist nicht klar geregelt, ob der Entscheidungsträger obligat, oder wenn nicht obligat, wann er von der Option, einen aus mehreren zusatznutzenbereinigten Erstattungspreisen gewichteten Erstattungsbetrag zu verhandeln, Gebrauch machen sollte.

Vorgeschlagene Änderung:

Das Institut definiert für den Entscheidungsträger eine klare Vorgabe, welche Optionen, einen aus mehreren zusatznutzenbereinigten Erstattungspreisen gewichteten Erstattungsbetrag zu verhandeln, zum Tragen kommen können.

12. Kosten-Nutzen-Bewertung medizinischer Interventionen - Datengrundlage

IQWiG Allgemeine Methoden

S.87/ 4.4.4 Datengrundlage

Bevorzugt sollten evidenzbasierte Leitlinien aus dem deutschen Gesundheitssystem verwendet werden (siehe Abschnitt 4.1.7). Sind diese in dem zu untersuchenden Indikationsgebiet nicht vorhanden, ist abzuwägen und transparent darzulegen, ob andere deutsche Leitlinien verwendet werden können oder ob auf Expertenbefragungen zurückgegriffen werden soll.

Position AstraZeneca

Wie in Punkt 9 bereits erläutert ist die ausschließliche Betrachtung deutscher Leitlinien nicht nachvollziehbar. Zwar wird hier bei nichtvorhandenen evidenzbasierten Leitlinien auf andere deutsche Leitlinien oder Expertenbefragungen verwiesen. Die Berücksichtigung internationaler evidenzbasierter Leitlinien bei Übertragbarkeit auf den deutschen Versorgungskontext, fließt hier nicht als weitere Option ein. Bei Berücksichtigung sollten die internationalen Leitlinien dann den Expertenmeinungen aufgrund der höheren Evidenzstufe vorzuziehen sein.

Vorgeschlagene Änderung:

Die Berücksichtigung internationaler evidenzbasierter Leitlinien stellt eine mögliche Option dar und ist hier mit aufzunehmen.

13. Kosten-Nutzen-Bewertung medizinischer Interventionen - Quelle der epidemiologischen Daten

IQWiG Allgemeine Methoden

S.99/4.5.2 Datengrundlage

Epidemiologische Daten können aus Sekundärdaten wie öffentlichen Datensammlungen und GKV-Routinedaten (siehe Abschnitt 4.4.4) sowie Registerdaten und ggf. wissenschaftlichen Publikationen kommen (siehe Abschnitt 4.1.7).

Position AstraZeneca

Epidemiologische Daten von kommerziellen Datenanbietern können, die methodische Qualität vorausgesetzt, genauso valide Quellen für Sekundärdaten darstellen wie öffentliche Datensammlungen und GKV-Routinedaten. Daher spricht nichts gegen die Verwendung dieser Daten von kommerziellen Anbietern. Im Gegenteil, gewisse benötigte Daten sind den vom Institut genannten Quellen nicht zu entnehmen oder nicht zugänglich.

Vorgeschlagene Änderung:

Explizite Nennung der kommerziellen Datenanbieter als valide Quelle für epidemiologische Daten.

14. Patientenpräferenzen

IQWiG Allgemeine Methoden

Patientenpräferenzen werden nur noch im Rahmen der Kosten-Nutzen-Bewertung betrachtet, und wurden in Kapitel 3 „Nutzen-Bewertung“ gestrichen.

Position AstraZeneca

In der Dossier-Vorlage zu Modul 3 im Kapitel 3.3.6 Angaben zu Versorgungsanteilen steht explizit *„Im Weiteren sollen bei dieser Abschätzung auch der Versorgungskontext und Patientenpräferenzen berücksichtigt werden.“* Die Patientenpräferenz ist ein wichtiger Aspekt der Nutzenbewertung, insbesondere bei der Beurteilung von Endpunkten nach derer Patientenrelevanz.

Vorgeschlagene Änderung

Das Institut begründet unter methodischen Gesichtspunkten, wieso Patientenpräferenzen für die Nutzen-Bewertung nicht mehr betrachtet werden, respektive nimmt diese wieder als relevante ergänzende Evidenzquelle in Kapitel 3 auf.

A.1.4 – Boehringer Ingelheim Pharma GmbH & Co. KG

**Stellungnahme der Firma Boehringer Ingelheim Pharma GmbH & Co. KG
(im Folgenden „Boehringer Ingelheim“) zum Entwurf
„Allgemeine Methoden Version 4.2.“ vom 18.06.2014**

Inhaltsverzeichnis

Abkürzungsverzeichnis	2
Einleitung	3
1. Streichung des Endpunktes „Patienten- bzw. Behandlungszufriedenheit“	4
2. Erhebung von Präferenzen zur Erstellung eines Maßes des Gesamtnutzens	5
3. Kosten-Nutzen-Bewertung medizinischer Interventionen	6
Literaturverzeichnis.....	8

**Stellungnahme der Firma Boehringer Ingelheim Pharma GmbH & Co. KG
(im Folgenden „Boehringer Ingelheim“) zum Entwurf
„Allgemeine Methoden Version 4.2.“ vom 18.06.2014**

Abkürzungsverzeichnis

COPD	Chronisch obstruktive Lungenerkrankung
EMA	European Medicines Agency
FDA	Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
ICS	Inhalierbare Corticosteroide
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
SGB V	Sozialgesetzbuch Fünftes Buch

**Stellungnahme der Firma Boehringer Ingelheim Pharma GmbH & Co. KG
(im Folgenden „Boehringer Ingelheim“) zum Entwurf
„Allgemeine Methoden Version 4.2.“ vom 18.06.2014**

Einleitung

Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) hat am 18. Juni 2014 den Entwurf der Allgemeinen Methoden Version 4.2 veröffentlicht und zur Anhörung gestellt.

Boehringer Ingelheim nimmt nachfolgend zum oben genannten Entwurf zu einigen ausgewählten Punkten Stellung.

**Stellungnahme der Firma Boehringer Ingelheim Pharma GmbH & Co. KG
(im Folgenden „Boehringer Ingelheim“) zum Entwurf
„Allgemeine Methoden Version 4.2.“ vom 18.06.2014**

1. Streichung des Endpunktes „Patienten- bzw. Behandlungszufriedenheit“

In den Abschnitten 3.1.1 und 3.2.4 der Entwurfsversion 4.2 des IQWiG Methodenpapiers wurden die patientenberichteten Endpunkte Patienten- bzw. Behandlungszufriedenheit ohne Begründung gestrichen.

Diese Änderung kann nicht nachvollzogen werden, da die Erhebung der Behandlungszufriedenheit einen wichtigen ergänzenden, aus Patientenperspektive erhobenen und für diesen relevanten Endpunkt in vielen Indikationsfeldern darstellen kann (Bradley und Gamsu, 1994; Atkinson et al., 2004).

Patienten- bzw. Behandlungszufriedenheit werden regelmäßig in Zulassungsstudien erhoben und ihre Ergebnisse wurden bereits in mehreren Verfahren der frühen Nutzenbewertung nach §35a SGB V dargestellt (13 von 90 Verfahren [Stand: Juli 2014]).

Darüber hinaus haben Patientenorientierung und –präferenzen in der Regel einen wichtigen Stellenwert im Bereich der Leitlinienentwicklung der Fachgesellschaften und bei den behandelnden Ärzten (Schubert et al., 2006).

Für die Messung der Patientenzufriedenheit stehen validierte Instrumente zur Verfügung, die auch zur Bewertung des Qualitätsmanagements im medizinischen Bereich eingesetzt werden können.

Das IQWiG hatte 2006 in der überarbeiteten Version des Methodenpapiers 2.0 die Aufzählung der nachstehend patientenrelevanten Zielgrößen aufgenommen und wie folgt kommentiert:

„Über klinisch relevante Ereignisse wie Mortalität, Morbidität und unerwünschte Wirkungen der medizinischen Intervention hinaus können auch andere Effekte oder Ereignisse für Patienten Bedeutung haben, wie z.B.

- *Krankheitsbezogene Lebensqualität, einschließlich der Auswirkung auf die Alltagsaktivitäten und das Wohlbefinden,*
- *Annehmlichkeiten, Umstände und Aufwand der Behandlung,*
- *Behandlungszufriedenheit und –präferenzen,*
- *Auswirkungen auf Angehörige.*

Ergebnisse, die für die Patienten wichtig sind, sollten berücksichtigt werden, bevor bewertende Schlussfolgerungen über Interventionen getroffen werden. Daten solcher Ergebnisse können die Aussagen einer systematischen Übersicht wesentlich ändern.“

Auch in vom IQWiG selbst durchgeführten Nutzenbewertungen (z.B. zu Insulinanaloga) wurden ebendiese Endpunkte als patientenberichtete Ergebnisparameter in die Bewertung eingeschlossen.

**Stellungnahme der Firma Boehringer Ingelheim Pharma GmbH & Co. KG
(im Folgenden „Boehringer Ingelheim“) zum Entwurf
„Allgemeine Methoden Version 4.2.“ vom 18.06.2014**

Fazit:

Aus Sicht von Boehringer Ingelheim ist es widersprüchlich, dass die patientenrelevanten Endpunkte Patienten- bzw. Behandlungszufriedenheit, welche die individuelle Wahrnehmung von Betroffenen hinsichtlich ihrer medizinisch-induzierten Intervention erfassen können, zukünftig nicht mehr berücksichtigt werden sollen.

Die Endpunkte Patienten- und Behandlungszufriedenheit sollten daher auch in Kapitel 3 wieder als patientenrelevante Endpunkte Aufnahme finden.

2. Erhebung von Präferenzen zur Erstellung eines Maßes des Gesamtnutzens

Patientenpräferenzen spielen international im Rahmen der Zulassung (beispielsweise EMA, FDA), der Bewertung von Gesundheitstechnologien sowie der anschließenden Erstattungsentscheidungsfindung neben den Methoden der evidenzbasierten Medizin eine zunehmend bedeutendere Rolle.

Es ist verwunderlich, weshalb es zwischen den Abschnitten 3 und 4 zu methodischen Unterschieden in der Nutzen- bzw. der Kosten-Nutzenbewertung kommen soll. In Kapitel 4 zur ‚Kosten-Nutzen-Bewertung medizinischer Interventionen‘ geht das IQWiG auf die Erhebung von Präferenzen zur Erstellung eines Maßes des Gesamtnutzens und mögliche Methoden der multikriteriellen Entscheidungsanalyse ein. Analogien zum Kapitel 3 werden zwar angedeutet, aber eine eindeutige methodische Stellungnahme zur Identifikation, Gewichtung und Aggregation von patientenrelevanten Endpunkten in der Nutzenbewertung medizinischer Interventionen fehlt. Auch in diesem Fall wird keine Rationale geliefert, warum Präferenzmessungen allein bei Kosten-Nutzen-Bewertungen zur Gewichtung von Endpunkten in Betracht gezogen werden jedoch nicht in der Nutzenbewertung, obwohl diese auch hier eine relevante Rolle spielen können. Beispielsweise wurde im Rahmen der frühen Nutzenbewertung von Afatinib (Giotrif®) eine Behandlungspräferenzstudie von Patienten mit nicht-kleinzelligem Lungenkarzinom im Dossier dargestellt und aufgeführt welchen Stellenwert Patienten einzelnen Endpunkten zuordnen (Boehringer Ingelheim, 2013a; G-BA, 2014).

Das IQWiG hat zwei Pilotprojekte zur Erhebung von Patientenpräferenzen bzw. zur Identifikation, Gewichtung und Priorisierung multipler Attribute durchgeführt. Dabei wurden das Analytical Hierarchy Process-Verfahren sowie die wahlbasierte Conjoint-Analyse untersucht (IQWiG, 2013; IQWiG, 2014a). Bei beiden Projekten kam das IQWiG zu dem Schluss, dass beide Verfahren geeignet sind um Patientenpräferenzen zu messen und die Ergebnisse solcher Präferenzenerhebungen für unterschiedliche Fragestellungen verwendet werden können (IQWiG, 2013).

Fazit:

Boehringer Ingelheim ist der Auffassung, dass die Ergebnisse von Patientenpräferenz-Studien bei verschiedenen Aspekten der Nutzenbewertung von medizinischen Interventionen (z.B. bei der Auswahl von patientenrelevanten Endpunkten oder der Saldierung des Zusatznutzens) eine wichtige Rollen spielen können und von daher zu berücksichtigen sind (G-BA, 2014a). Aus diesem Grund sollten die Inhalte von Abschnitt 4.3.3 B) ebenfalls in Kapitel 3 aufgeführt und in Kapitel 4 darauf verwiesen werden.

3. Kosten-Nutzen-Bewertung medizinischer Interventionen

Die vom IQWiG gestellten Anforderungen an die Datengrundlage für eine Kosten-Nutzen-Bewertung werden in Abschnitt 4.1.7 erläutert. Darin lautet es: *„werden Leitlinien verwendet, sollten diese aus dem deutschen Gesundheitssystem stammen und möglichst evidenzbasiert sein“* (IQWiG, 2014).

Aus den bisherigen Verfahren der frühen Nutzenbewertung nach §35a SGB V sind Beispiele bekannt, die zeigen, dass die Aktualität der nationalen Behandlungsleitlinien nicht immer den europäischen und/oder internationalen Leitlinien entspricht. So sind beispielsweise bei Lungenkrebs die europäischen und internationalen Leitlinien auf einem neueren Stand (Boehringer Ingelheim, 2013b).

Das Gleiche trifft auf die nationale Behandlungsleitlinie für chronischobstruktive Lungenerkrankungen zu (G-BA, 2014b), sodass bei der frühen Nutzenbewertung der Wirkstoffkombination Indacaterol/Glycopyrronium die internationale GOLD-Leitlinie 2014 anstelle der ungültigen Nationalen Versorgungsleitlinie COPD zur Beurteilung der spirometrischen COPD-Schweregrade herangezogen wurde.

Fazit:

Ein alleiniger Fokus auf deutsche Leitlinien kann dazu führen, dass Entscheidungen nicht basierend auf dem aktuellen Stand der Wissenschaft und der tatsächlichen Versorgungsrealität in Deutschland getroffen werden. Aus Sicht von Boehringer Ingelheim erscheint es sinnvoll, nationale und internationale Leitlinien bei der Bewertung zu berücksichtigen und darüber hinaus die Qualität und Aktualität der verwendeten Datengrundlage sowie die Relevanz für Deutschland zu bewerten.

**Stellungnahme der Firma Boehringer Ingelheim Pharma GmbH & Co. KG
(im Folgenden „Boehringer Ingelheim“) zum Entwurf
„Allgemeine Methoden Version 4.2.“ vom 18.06.2014**

Weitere Anmerkungen zur Kosten-Nutzen-Bewertung

Abschnitt 4.1.3: Grundsätzliche Aspekte der Modellentwicklung

Das IQWiG führt auf S. 82 aus: *„Da oft Daten zu einzelnen Aspekten fehlen, ist es wichtig, den Einfluss auf die Ergebnisse durch Sensitivitätsanalysen zu explorieren.“* Sensitivitätsanalysen stellen kein Mittel dar, Evidenzlücken bei der Modell-Parametrisierung zu schließen. Boehringer Ingelheim erachtet die gewählte Formulierung daher als verwirrend und schlägt folgende Alternative vor: Wenn verschiedene Annahmen und Modell-Inputs plausibel sind, ist es wichtig, deren Einfluss auf die Ergebnisse durch Sensitivitätsanalysen zu analysieren.

Abschnitt 4.1.4: Modelldokumentation und Modellvalidierung

Im Zuge der Darlegung der Transparenzanforderungen an das den Reviewern und der Öffentlichkeit zu überlassende elektronische Modell (S. 86) sollte eine orientierungsgebende Aufzählung von grundsätzlich akzeptierten Modellierungsumgebungen (z.B. Excel, DATA, R, WinBugs) ergänzt werden, gegebenenfalls erweitert durch ein Rückfrageerfordernis bei beabsichtigter Verwendung eines Nicht-Standard-Softwarepaketes.

Abschnitt 4.1.5: Unsicherheit und Verteilung von Nutzendaten

In der Entwurfsversion der Allgemeinen Methoden 4.2 heißt es: *„Für geschätzte Effekte im Rahmen einer Nutzenbewertung können im Allgemeinen Konfidenzintervalle bzw. Credible Intervals berechnet werden, die die Präzision bzw. Unsicherheit der Punktschätzer anzeigen.“* Boehringer Ingelheim schlägt vor, die Begriffe „Konfidenzintervall“ und „Credible Interval“ dem jeweiligen Wahrscheinlichkeitsbegriff zuzuordnen sowie eine kurze, prägnante Erläuterung dieser wetteifernden Philosophien der statistischen Analyse (mit entsprechenden Literaturverweisen) zu geben.

Abschnitt 4.1.6: Abgrenzung der Kosten

Das IQWiG schreibt auf S. 93: *„Die Lebensqualität von Angehörigen wird im Allgemeinen auf der Nutzenseite nicht berücksichtigt.“* Boehringer Ingelheim erachtet dies für problematisch, da für bestimmte Indikationen (z.B. der Alzheimer-Krankheit) hinlänglich bekannt ist, dass das Krankheitsgeschehen Einfluss auf die Lebensqualität (pflegender) Angehöriger haben kann. Lebensqualitätseffekte bei Angehörigen sollten daher nicht generell von der Berücksichtigung ausgeschlossen werden.

**Stellungnahme der Firma Boehringer Ingelheim Pharma GmbH & Co. KG
(im Folgenden „Boehringer Ingelheim“) zum Entwurf
„Allgemeine Methoden Version 4.2.“ vom 18.06.2014**

Literaturverzeichnis

- Atkinson MJ, Sinha A, Hass SL, Colman SS, Kumar RN, Brod M, Rowland CR, 2004.
Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease. Health Qual Life Outcomes;2:12.
- Boehringer Ingelheim, 2013a. Dossier zur Nutzenbewertung gemäß § 35a SGB V – Afatinib (Giotrif®). Modul 4A. Stand: 15.11.2013. Verfügbar: https://www.g-ba.de/downloads/92-975-402/2013-11-14_Modul4A_Afatinib.pdf [Aufgerufen am 01.08.2014].
- Boehringer Ingelheim, 2013b. Dossier zur Nutzenbewertung gemäß § 35a SGB V – Afatinib (Giotrif®). Modul 3A. Stand: 15.11.2013. Verfügbar: https://www.g-ba.de/downloads/92-975-401/2013-11-14_Modul3A_Afatinib.pdf [Aufgerufen am 01.08.2014].
- Bradley C, Gamsu S, 1994. Guidelines for Encouraging Psychological Well-being. Report of a Working Group of the World Health Organization Regional Office for Europe and International Diabetes Federation European Region St Vincent Declaration Action Programme for Diabetes. Diabet Med.;11(5):510-6.
- G-BA, 2014a. Mündliche Anhörung gemäß 5. Kapitel § 19 Abs. 2 Verfahrensordnung des G-BA - Wirkstoff Afatinib. Stenografisches Wortprotokoll. Stand: 25.03.2014. Verfügbar: https://www.g-ba.de/downloads/91-1031-87/2014-03-25_Wortprotokoll_end_Afatinib.pdf [Aufgerufen am 01.08.2014].
- G-BA, 2014b. Tragende Gründe zum Beschluss des G-BA - Wirkstoff Indacaterol/Glycopyrronium. Stand: 08.05.2014. Verfügbar: https://www.g-ba.de/downloads/40-268-2791/2014-05-08_AM-RL-XII_Indacaterol-Glycopyrronium_2013-11-15-D-081_TrG.pdf [Aufgerufen am 01.08.2014].
- IQWiG, 20006. Methoden: 2.0. Stand: 19.12.2006. Verfügbar: https://www.iqwig.de/download/Methoden_IQWiG_Version_2.0.pdf [Aufgerufen am 01.08.2014].
- IQWiG, 2013. IQWiG-Berichte – Nr163. Analytic Hierarchy Process (AHP) – Pilotprojekt zur Erhebung von Patienten-präferenzen in der Indikation Depression Stand: 08.05.2013. Verfügbar: https://www.iqwig.de/download/Arbeitspapier_Analytic-Hierarchy-Process_Pilotprojekt.pdf [Aufgerufen am 01.08.2014].

**Stellungnahme der Firma Boehringer Ingelheim Pharma GmbH & Co. KG
(im Folgenden „Boehringer Ingelheim“) zum Entwurf
„Allgemeine Methoden Version 4.2.“ vom 18.06.2014**

IQWiG, 2014a. IQWiG-Berichte – Nr. 227. Wahlbasierte Conjoint-Analyse – Pilotprojekt zur Identifikation, Gewichtung und Priorisierung multipler Attribute in der Indikation Hepatitis C. Stand: 23.07.2014. Verfügbar: https://www.iqwig.de/download/GA10-03_Arbeitspapier_Version_1-1_Conjoint-Analyse-Pilotprojekt.pdf [Aufgerufen am 01.08.2014].

IQWiG, 2014b. Allgemeine Methoden: Entwurf für Version 4.2. Stand: 18.06.2014. Verfügbar: https://www.iqwig.de/download/IQWiG_Methoden_Entwurf-fuer-Version-4-2.pdf [Aufgerufen am 01.08.2014].

Schubert I, Lelgemann M, Kirchner H, von Ferber C, von Ferber L, Ollenschläger G, 2006. Handbuch zur Entwicklung regionaler Leitlinien. Verfügbar: http://www.pmvforschungsgruppe.de/pdf/02_forschung/c_handbuch_aezq.pdf [Aufgerufen am 01.08.2014]

A.1.5 – Bundesarbeitsgemeinschaft Selbsthilfe von Menschen mit Behinderung und chronischer Erkrankung und ihren Angehörigen e. V. (BAG Selbsthilfe)



Stellungnahme der

Bundesarbeitsgemeinschaft SELBSTHILFE von Menschen mit Behinderung und chronischer Erkrankung und ihren Angehörigen (BAG SELBSTHILFE) e. V.

zum

Entwurf des Methodenpapiers 4.2 des IQWiG

Hintergrund

Als Dachverband von 121 Bundesverbänden der Selbsthilfe chronisch kranker und behinderter Menschen sowie von 13 Landesarbeitsgemeinschaften misst die BAG SELBSTHILFE der Neufassung des Methodenpapiers für die Ausgestaltung der Potenzial- und Nutzenbewertung sehr große Bedeutung bei. Wir nehmen zu folgenden Punkten Stellung: Ergänzung des neuen Produkts Potenzialbewertung gemäß § 137e SGB V sowie der entsprechenden Methoden in den Abschnitten 1.1, 2.1, 2.2, 3.8, 7.2 (vorher 6.2), 8.1.3 (vorher 7.1.3) und 8.1.4 (vorher 7.1.4)

I. Erprobungsregelung

1. Neue Regelungen

In Kapitel 3.8 sind die spezifischen Evidenzstufen (gering, sehr gering, minimal) zur Potenzialbewertung und deren Bedeutung nach § 137e nach Antragstellung durch externe Anbieter durch das IQWiG beschrieben.

So heißt es auf Seite 75 wie folgt:

„Daher ergibt sich ein Potenzial insbesondere dann, wenn Studien geringer Ergebnissicherheit mindestens kleine Effekte zeigen, wenn Studien sehr geringer Ergebnissicherheit mindestens mittlere Effekte zeigen oder wenn Studien minimaler Ergebnissicherheit mindestens große Effekte zeigen. Als ungefähre Grenzen zwischen kleinen, mittleren und großen Effekten können für das relative Risiko Werte von 0,8 und 0,5 dienen [147,419].“

2. Bewertung

Aus Sicht der BAG SELBSTHILFE stellt diese Differenzierung nach dem relativen Risiko eine große Hürde dar. Daher wird vorgeschlagen, dass auch Studien wie Fallserien nach dem Gesichtspunkt der Gleichgerichtetheit (beispielsweise 2 Fallserien mit statistisch signifikanten Ergebnissen in gleicher Richtung) positiv bewertet werden sollten. Aus Erfahrungen in der Patientenvertretung sehen wir hier durchaus Diskussionsbedarf. Aus Sicht der BAG SELBSTHILFE ist zu betonen, dass hier durch den Gesetzgeber offene Begriffe eingeführt worden sind, die den klassischen Kategorisierung der EBM nicht unbedingt entsprechen, die aber die Grundlage zur Bildung von Kriterien liefern sollen, wann Erprobungsverfahren durchgeführt werden sollen.

Aus Sicht der BAG SELBSTHILFE sollte dieser Prozess der Kriterienbildung behutsam erfolgen, so dass die Festlegung konkreter Grenzwerte noch als verfrüht erscheint.

II. Bewertung des Nutzens und des Schadens

1. Änderungen zum Thema Patientenzufriedenheit in den Abschnitten 3.1.1 und 3.2.4

a. Neue Regelung

An verschiedenen Stellen hat das IQWiG den Begriff „Patientenzufriedenheit“ bzw. „Zufriedenheit der Patienten“ gestrichen.

b. Bewertung

Im Kapitel 3.1. wird aus Sicht der BAG SELBSTHILFE eine Klarstellung der Nutzendimensionen durch das SGB V vorgenommen. Im Rahmen der Nutzenbewertung bei neuen Methoden oder im Rahmen der frühen Nutzenbewertung sind hier Morbidität, Mortalität und Lebensqualität genannt. Die Bedeutung von patientenrelevanten Endpunkten für Patienten wird bereits zu Beginn des Verfahrens durch Befragung der Patienten aus den maßgeblichen Organisationen berücksichtigt. Die BAG SELBSTHILFE hält es für sinnvoll, dass diese anhand von validierten Instrumenten klar bestimmt werden können; eine unspezifische Messung der Patientenzufriedenheit ist aus unserer Sicht nicht sinnvoll. Daher stimmen wir der Streichung zur Klarstellung zu. Aussagen zu Patientenzufriedenheit lassen keine Rückschlüsse über den nach SGB V definierten „Zusatznutzen“ im Sinne patientenrelevanter Endpunkte

zu. Die Patientenvertretung sieht hier eine mögliche Stärkung zur verbesserten Implementierung der Erfassung der Lebensqualität und würde sich wünschen, dass Studien diesen Endpunkt stärker in Zukunft berücksichtigen.

Andererseits können Anwendungserleichterungen wie eine Vergrößerung der Einnahmeintervalle oder Erleichterungen bei der Applikation von Arzneimitteln durch veränderte Darreichungsformen durchaus von Patientinnen und Patienten als relevante Verbesserungen der Versorgung angesehen werden. Solche Verbesserungen sind zwar nicht per se als Zusatznutzen im Sinne des AMNOG-Bewertungsprozesses anzusehen, werden aber auch unter dem Oberbegriff der „Patientenzufriedenheit“ diskutiert.

Aus Sicht der BAG SELBSTHILFE sind diese Aspekte durchaus von Relevanz. Zum einen können sich die Verbesserungen auf die Lebensqualität auswirken, was dann mit Hilfe der hierfür geeigneten Instrumente nachzuweisen ist. Zum anderen können diese Aspekte auch dann, wenn sie sich nicht im Endpunkt Lebensqualität niederschlagen als „Fortschritt“ in der Versorgung angesehen werden. Solche „Fortschritte“ betreffen dann aber nicht die Frage des „Zusatznutzens im Sinne des AMNOG, sondern Anwendungsverbesserungen, deren mögliche Vergütungsfähigkeit anderweitig zu diskutieren ist.

Im Ergebnis ist es daher auch aus Sicht der BAG SELBSTHILFE angemessen, neben der Verbesserung der Lebensqualität das Kriterium der Patientenzufriedenheit nicht noch zusätzlich im Methodenpapier zu verorten.

2. Streichung der Definition und Konkretisierung der patientenrelevanten Endpunkte (Kapitel 3.1.1.)

a. Neue Regelung

Auf Seite 39 wird die Streichung der Konkretisierung der Begriffe Mortalität, Morbidität und Lebensqualität aus dem § 35b SGB V vorgeschlagen.

b. Bewertung

Die BAG SELBSTHILFE erachtet die Streichung an diese Stelle als nicht sinnvoll, da in diesem Abschnitte nachvollziehbar, patientenrelevant und gesetzeskonform beschrieben wird, was unter diesen abstrakten Begriffen verstanden wird. Dies kann nur hilfreich sein.

III. Informationsbewertung- Beurteilung klinischer Relevanz - Ergänzung zu patientenrelevanten Endpunkten in Abschnitt 8.3.3

a. Regelung

In Abschnitt 8.3.3. wurde neu geregelt, dass patientenrelevante Endpunkte mit Hilfe (komplexer) Skalen erhoben werden können. Es wird darauf hingewiesen, dass Voraussetzung für die Erhebung solcher Endpunkte die Verwendung von validierten bzw. etablierten Instrumenten ist.

b. Bewertung

Auch die BAG SELBSTHILFE geht davon aus, dass die sogenannten PROs (patient relevant outcomes) mit validierten Instrumenten erhoben werden müssen. Allerdings würden wir die Änderungen zu patientenrelevanten Endpunkten um den folgenden Aspekt ergänzen, da durch die oft multizentrischen durchgeführten Studien in verschiedenen Kulturen eine validierte Übersetzung hilfreich ist:

„Die patientenrelevanten Endpunkte Morbidität und Lebensqualität sollen mithilfe von (komplexen) Skalen erhoben werden. Eine Voraussetzung für die Berücksichtigung solcher Endpunkte und deren Ergebnisse ist die Verwendung von in der Regel in der Indikation validierten und kultursensiblen Instrumenten.“

Dr. Martin Danner, BAG SELBSTHILFE

Susanne Teupen, Stabsstelle Patientenbeteiligung G-BA

Düsseldorf, 06.08.2014

A.1.6 – Bundesverband der Pharmazeutischen Industrie e. V. (BPI)

Stellungnahme

Stellungnahme des BPI zum IQWiG Methodenpapier Version 4.2 (Entwurf)

Allgemeine Anmerkungen:

Sowohl im Kapitel 3 „Nutzenbewertung medizinischer Interventionen“ als auch im Kapitel 4 „Kosten-Nutzen-Bewertung medizinischer Interventionen“ liegen der Entscheidung mehrere (auch multiple) Endpunkte zugrunde. Es muss daher ein Weg beschrieben werden, wie man diesen Umstand berücksichtigt. Lösungsansätze zur Priorisierung und Gewichtung von Endpunkten bei einem multidimensionalen Nutzen bieten die Methoden der Conjoint Analysis (CA) und Analytic Hierarchy Process (AHP). Beide Ansätze waren in der vorherigen Version des Methodenpapiers (Version 1.0 vom 12.10.2009) im Kapitel 3, speziell 3.2.6 „Endpunktgewichtung“ aufgeführt. Warum erfolgt nunmehr im vorliegenden Entwurf (Version 4.2) eine Verlagerung der Passagen? Anders gesagt, warum werden jetzt im Kapitel 3.1.5 „Zusammenfassende Bewertung“ die DCE und AHP gestrichen und warum erfolgt an dieser Stelle nur noch ein Verweis auf den Abschnitt 4.3.3 „Maß des Gesamtnutzens“, wo im Unterpunkt B) „Erhebung von Präferenzen zur Erstellung eines Maßes des Gesamtnutzens“ die beiden Methoden wieder angeführt werden? Sind somit die beiden Methoden auch bei der Nutzenbewertung im Kapitel 3 anwendbar?

Besondere Anmerkungen:

2.2.1. „Auswahl externer Sachverständiger“, 2.2.2 Gewährleistung der fachlichen Unabhängigkeit und 2.2.3 Review der Produkte des Instituts

Gemäß den Ausführungen im Abschnitt 2.2.1 (Seite 33) vergibt das IQWiG Aufträge an externe Sachverständige in „einem transparenten und nichtdiskriminierenden Wettbewerb“.

Hierzu können sich Interessierte in die Sachverständigen-Datenbank des IQWiG eintragen, aus der auftragsabhängig Sachverständige nach den im Abschnitt 2.2.1 (Seite 34) genannten Kriterien ausgewählt werden.

Zur Verbesserung der Verfahrenstransparenz sollte das Auswahlverfahren, die in Frage kommenden Sachverständigen und auch die konkret bei der Auswahl angewandten Kriterien veröffentlicht werden. Zudem sollten auch externe Reviews im Sinne des Kapitels 2.2.3 veröffentlicht werden.

3.1. „Patientenrelevanter medizinischer Nutzen und Schaden“ (Lebensqualität)

1. „Die Zielgrößen (Mortalität, Morbidität und gesundheitsbezogene Lebensqualität) sind auch im SGB V als vorrangig zu berücksichtigende Zielgrößen genannt, z. B. in § 35 Abs. 1b SGB V. Ergänzend kann der interventions- und erkrankungsbezogene Aufwand **und**

Stellungnahme

die Zufriedenheit der Patientinnen und Patienten mit der Behandlung berücksichtigt werden."

2. Die Streichung der „Zufriedenheit des Patienten“ erfolgt ohne Begründung. Offen bleibt, warum die Zufriedenheit der Patienten in Zukunft keine Rolle mehr spielen soll.
3. In diesem Zusammenhang sei hier auf die umfassende Definition der WHO Quality of Life Gruppe verwiesen, die Lebensqualität als „die Wahrnehmung von Individuen bezüglich ihrer Position im Leben im Kontext der Kultur und der Wertesysteme, in denen sie leben, und in Bezug auf ihre Ziele, Erwartungen, Standards und Interessen“¹ definiert. Inzwischen hat sich eine operationale Definition durchgesetzt, nach der sich Lebensqualität als multidimensionales Konstrukt aus Wohlbefinden und Funktionsfähigkeit aus Sicht der Patienten zusammensetzt.² Grundlegende Dimensionen betreffen die körperliche, emotionale, mentale, soziale und alltagsfunktionale Komponente. Darüber hinaus kann sich auch die Patientenzufriedenheit auf die gesundheitsbezogene Lebensqualität auswirken.
4. So zählen auch subjektive Wohlbefindensmodelle zu den zentralen konzeptionellen Ansätzen der Lebensqualitätsmessung.³ Als breit akzeptiertes Instrument sei hier beispielhaft der Diabetes Treatment Satisfaction Questionnaire (DTSQ) genannt. Der DTSQ (Status Version) besteht aus acht Positionen (Items). Sechs davon erfassen die Behandlungszufriedenheit und zwei weitere die wahrgenommene Frequenz von Hyper- und Hypoglykämien. Jedes Item wird auf einer Skala von 0 (sehr unzufrieden) bis 6 (sehr zufrieden) bewertet.⁴ Der DTSQ wurde von der Weltgesundheitsorganisation (WHO) sowie der International Diabetes Federation (IDF) als sehr wertvoll in der Erhebung von Outcomes der Diabetesbehandlung eingestuft.⁵ Mittlerweile liegen viele Publikationen vor, in denen der DTSQ zur Messung der Behandlungszufriedenheit eingesetzt wurde.⁶

¹ WHO. The WHO quality of life assessment (WHOQOL): Position paper from the WHO. Soc Sci Med 1995; 41: 1403-09

² LC Doward, SP McKenna. Defining Patient Reported Outcomes. Value in health 2004; 7/1: S4-S8

³ M. Bullinger. Das Konzept der Lebensqualität in der Medizin - Entwicklung und heutiger Stellenwert. Z. Evid. Fortbild. Qual. Gesundh. wesen 2014; 108: 97-103

⁴ C. Bradley. The Diabetes treatment satisfaction questionnaire: DTSQ. In Bradley C, ed. Handbook of Psychology and Diabetes. A guide to psychological measurement in diabetes research and practice. Chur, Switzerland: Harwood Academic Publishers 1994; 11-133

⁵ C. Bradley, DS Gamsu. Guidelines for encouraging psychological well-being: report of a working group of the WHO regional office for Europe and IDF European region St. Vincent declaration action programme for diabetes. Diabet Med 1994; 11: 510 - 516

⁶ C. Bradley, CJB Gilbride. Improving treatment satisfaction and other patient reported outcomes in people with type 2 diabetes: the role of once-daily insulin glargine. Diabetes, Obesity and Metabolism 2008; 10/1: 50 - 65

Stellungnahme

3.1.1 „Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens“ und 3.2.4. „Patientenberichtete Endpunkte“

Die Termini „Patientenzufriedenheit“ und „Behandlungszufriedenheit“ als Aspekte der patientenberichteten Endpunkte (PROs) sollen an verschiedenen Stellen im vorliegenden Entwurf gestrichen werden.

In einem aktuellen „Reflection Paper on the use of patient reported outcome (PRO) measures in oncology studies“ (EMA/CHMP/292464/2014) der Europäischen Arzneimittelagentur (EMA) wird die Bedeutung der verschiedenen Aspekte patientenberichteter Endpunkte betont.

Es wird dargestellt, dass PROs als Oberbegriff für Symptome, gesundheitsbezogene Lebensqualität, Gesundheitszustand, Behandlungsadhärenz und Behandlungszufriedenheit, die direkt vom Patienten berichtet werden, zu sehen sind. Zudem wird die Wertigkeit dieser Aspekte betont. Die mit der Streichung der Behandlungszufriedenheit offenbar verbundene Abwertung im vorliegenden Entwurf ist vor diesem Hintergrund nicht nachvollziehbar

Da das objektive und subjektive Wohlbefinden der Patienten im Mittelpunkt stehen sollten, stellen Patientenzufriedenheit bzw. Behandlungszufriedenheit mit der gewählten Therapie wichtige Indikatoren für den Nutzen der Therapie dar. Die Erfassung und Darstellung der Patientenzufriedenheit und Behandlungszufriedenheit sollten daher wichtige Bausteine in der Bewertung von Therapien bleiben.

3.8. „Potenzialbewertung“

Das IQWiG führt aus, dass im Hinblick auf die Potenzialbewertung und im Vergleich zur Nutzenbewertung von Arzneimitteln niedrigere Anforderungen an die Evidenz zu stellen sind. Als Bewertungsstufen für die Potenzialbewertung werden hinsichtlich der Ergebnissicherheit die Grade „gering“, „sehr gering“ und „minimal“ neu eingeführt. Für jeden dieser Grade werden zudem minimale Effektstärken definiert und gemäß Kapitel 3.8 (Seiten 73-75) „ungefähre Grenzen“ festgelegt.

Die neu eingeführten Effektstärken „gering“, „sehr gering“ und „minimal“ führen in Bezug auf die Potenzialbewertung dazu, dass selbst die bestmöglich verfügbare Evidenz bei der Testung von Methoden verbunden mit einer sprachlichen Herabwürdigung als „geringer“ Effekt“ beschrieben wird.

Die Nutzenbewertung und die Potenzialbewertung setzen auf völlig unterschiedlichen Ausgangssituationen auf, denen Rechnung zu tragen ist. Die sprachliche Beschreibung der Effektstärken hat dieser Situation Rechnung zu tragen. Eine ausschließliche Verwendung negativ konnotierter Begriffe zur Beschreibung der Effekte von Potenzialen in der Bewertung von Methoden ist abzulehnen.

Stellungnahme

4. „Kosten-Nutzen-Bewertung medizinischer Interventionen“

4.1 Einführung

„Gemäß SGB V bestimmt das Institut auftragsbezogen über die Methoden und Kriterien für die Erarbeitung von Kosten-Nutzen-Bewertungen (KNB) auf der Grundlage der in den jeweiligen Fachkreisen anerkannten internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie.“ „Alle Abweichungen von den hier vorgelegten Methoden sind im Einzelfall zu begründen.“

Diese Generalklausel eröffnet dem IQWiG einen unbegrenzten Spielraum für Einzelfallentscheidungen. Sie bedeutet gleichzeitig eine maximale Planungsunsicherheit für den pharmazeutischen Unternehmer. Für ein transparentes Verfahren einer Kosten-Nutzen-Bewertung bedarf der Klarstellung, was im Verfahren durch das pharmazeutische Unternehmen zu liefern, vorzubereiten bzw. zusammenzutragen ist und was durch das IQWiG erfolgt. Eine Kosten-Nutzen-Bewertung ist eine nicht unerheblich finanzielle und personelle Investition. Ein Austausch, im Sinne eines Scoping-Workshop, zu Beginn des Verfahrens, bspw. um essentielle Eckpunkte sowie gemeinsam eine verbindliche Vorgehensweise festzulegen, oder auch ein das Verfahren begleitender Austausch kann die Qualität und Akzeptanz des Bewertungsergebnisses fördern.

In der Kosten-Nutzen-Bewertung wird der Evidenzfokus zunächst auch auf RCTs liegen. Die Limitationen bzgl. der Aussagekraft und der Erkenntnisgewinnung, die RCTs in diesem Zusammenhang aufweisen, sind bekannt. Gerade vor diesem Hintergrund und insbesondere bei Arzneimitteln zur Behandlung von chronischen Erkrankungen ist es umso wichtiger, dass neben RCTs weitere Studienformen in die Bewertung einfließen und anerkannt werden.

Patientenpräferenzen (Patienten- und auch Ärztesicht) sind bei der Beurteilung einer Intervention essentiell und sollten daher einbezogen werden. Dies bildet gerade die Forderung nach dem Versorgungsalltag ab.

4.1.1 Rechtliche Grundlagen für die Kosten-Nutzenbewertung nach SGB V

Laut Methodenpapier: „Daher muss die KNB einen Vergleich der Effizienz von Interventionen in einem Indikationsgebiet ermöglichen, ohne die Frage einer indikationsübergreifenden Priorisierung im Gesundheitssystem zu beantworten.“

Die Beschränkung der KNB auf Interventionen in nur einem Indikationsgebiet wird einer gesamtwirtschaftlichen und gesundheitspolitischen Betrachtung nicht gerecht. Indikationsübergreifende Bewertungen dienen der Allokation von Ressourcen in diejenigen Bereiche, in denen sie den größten gesamtgesellschaftlichen Nutzen entfalten.

4.1.2 Perspektive

Der Unterschied zwischen der GKV-Perspektive sowie der Perspektive der GKV-Versichertengemeinschaft sollte präzisiert werden.

Stellungnahme

4.1.3 Zeithorizont

Die Abbildbarkeit von chronischen Erkrankungen bzw. deren Endpunkte in Modellen über einen meist jahrzehntelangen Zeithorizont ist komplex. Diese Komplexität gilt es zu berücksichtigen, ggf. mit Hinblick auf einen Scoping-Workshop zu diskutieren und gemeinsame Lösungen zu finden.

4.1.5 Versorgungspfad

Der Terminus „Versorgungspfad“ sollte definiert werden. Der Begriff „Versorgungskontext“ sollte generell durch „Versorgungspfad“ ersetzt werden, sofern damit das Gleiche gemeint ist. Ist dies nicht der Fall, so muss auch der Begriff „Versorgungskontext“ definiert werden.

4.1.7 Datengrundlage

„Für das Maß des Gesamtnutzens werden einzelne Studien oder Datenerhebungen (siehe auch Abschnitt 4.3.3) verwendet.“

Das Maß des Gesamtnutzens stellt den Kern jeder KNB dar und bedarf deshalb einer präzisen Darstellung. Insbesondere müssen methodische Hinweise zur Identifikation, Gewichtung und Aggregation mehrerer Endpunkte aus Patientenperspektive ergänzt werden (z. B. in Anlehnung an die Methode der multikriteriellen Entscheidungsanalyse „Multiple Criteria Decision Analysis (MCDA)“). So liefern innovative Studienformen wie z. B. die präferenzbasierte Nutzenmessung mittels Discrete Choice Experiment (DCE) oder Analytic-Hierarchy-Process (AHP) wertvolle Zusatzinformationen über die Bedeutung von Therapien aus Sicht der Betroffenen (Arzt und Patient). Einschlägig bekannte Pilotprojekte dazu sind bereits seit längerer Zeit abgeschlossen und liegen dem IQWiG vor.

Im Übrigen gibt es aus gesundheitsökonomischer Sicht keine inhaltlichen oder methodischen Anhaltspunkte, welche den Einsatz unterschiedlicher Instrumentarien im Kontext der Nutzen-Schaden Abwägung (Kapitel 3) und der Kosten-Nutzen Analyse (Kapitel 4) sinnvoll erscheinen lassen. Eine gesundheitsökonomisch sinnvolle Kosten-Nutzen-Bewertung sollte in beiden Anwendungsfeldern auf die gleichen Methoden und Instrumente zurückgreifen.

4.2.3 Einflussdiagramm und Modellkonzept

Das Einflussdiagramm sollte durch ein Beispiel veranschaulicht werden.

„Wenn mehrere patientenrelevante Endpunkte nebeneinander dargestellt werden, wird für jeden patientenrelevanten Endpunkt eine eigene Effizienzgrenze erstellt.“

Diese Vorgehensweise ist nicht mehr zeitgemäß. Es existiert ein ausreichend erprobtes methodisches Instrumentarium mit dem verschiedene Nutzenkomponenten zu einem einzigen Maß des Gesamtnutzens aggregiert werden können (siehe Ausführungen zu Pkt. 4.1.7).

Stellungnahme

7. „Informationsbeschaffung“

Auf Seite 164 wird ausgeführt, dass unaufgefordert übermittelte Studiendaten seitens des Institutes bei der Abfassung seiner Berichte inhaltlich nicht berücksichtigt werden. Begründet wird dies mit dem potentiellen Risiko einer selektiven Auswahl der übermittelten Daten, die zu einer Verzerrung der Berichtsergebnisse führen könnten.

Diese vorgesehene Einschränkung hinsichtlich der Erhebung der Datengrundlage kann dazu führen, dass relevante Informationen bzw. neue aktuelle Evidenz unberücksichtigt bleiben. Auch dies kann eine Verzerrung des Berichtsergebnisses nach sich ziehen.

In Kapitel 1.2.4 (Seite 7) des vorliegenden Entwurfs wird ausgeführt: „Eine weitere zentrale Strategie der EbM besteht darin, alle nach der Qualität ihres Designs und ihrer Durchführung angemessenen Studien zu einer Frage zu identifizieren und so den Stand des zuverlässigen Wissens zusammenzufassen.“ Dieser Aussage folgend ist der beabsichtigte Ausschluss nicht nachvollziehbar.

Es bleibt auch fraglich, warum ein kategorischer Ausschluss unaufgefordert eingereicherter Daten erfolgt, wenn es gemäß Kapitel 7.2 des vorliegenden Entwurfs des Methodenpapiers zur etablierten Arbeitsweise des Instituts gehört, die eingereichten Daten zu überprüfen.

8.1.3. „Rangordnung verschiedener Studienarten/ Evidenzgrade“

Im Entwurf wird zunächst ausgeführt, dass verschiedene Ansätze zur Hierarchisierung der Evidenzgrade von Studientypen bestehen, wobei sich das IQWiG weitgehend an der Evidenzklassifizierung in der aktuellen Verfahrensordnung des Gemeinsamen Bundesausschusses bzw. § 5 Abs. 6 der Arzneimittel-Nutzenbewertungsverordnung orientiert.

Bei den bisherigen Dossierbewertungen in Verfahren nach § 35a SGB V war des Öfteren eine Situation gegeben, in der die in internationalen Entwicklungsprogrammen verwandten Komparatoren nicht mit den vom Gemeinsamen Bundesausschuss vorgegebenen zweckmäßigen Vergleichstherapien in Übereinstimmung waren.

§ 5 Absatz 5 Arzneimittel-Nutzenbewertungsverordnung sagt hierzu:

„Liegen keine direkten Vergleichsstudien für das neue Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vor oder lassen diese keine ausreichenden Aussagen über einen Zusatznutzen zu, können verfügbare klinische Studien für die zweckmäßige Vergleichstherapie herangezogen werden, die sich für einen indirekten Vergleich mit dem Arzneimittel mit neuen Wirkstoffen eignen.“

Stellungnahme

Indirekte Vergleiche sind bislang in der Evidenzklassifizierung nicht genannt. Es wird daher darum gebeten, dass in Kapitel 8.1.3 ausgeführt wird, an welcher Stelle indirekte Vergleiche bzw. Netzwerk-Metaanalysen in die Evidenzhierarchie eingruppiert sind.

8.3.3. „Beurteilung klinischer Relevanz“

Eine auf Seite 189 vorgenommene Ergänzung zu patientenrelevanten Endpunkten besagt, dass eine Voraussetzung zur Erhebung patientenrelevanter Endpunkte mittels (komplexer) Skalen die Verwendung von validierten bzw. etablierten Instrumenten ist.

Es stellt sich die Frage, warum es im ergänzten Text „eine Voraussetzung“ und nicht „die Voraussetzung“ heißt? Welche weiteren Voraussetzungen werden seitens des IQWiG hier verlangt? Weiterhin stellt sich die Frage, wie seitens des Instituts beurteilt wird, inwiefern ein bestimmtes Instrument als „validiert“ und „etabliert“ anzusehen ist. Die gewählten Formulierungen erscheinen unspezifisch und gewähren einen breiten Interpretationsspielraum, der nicht dazu beiträgt, die bestehende Unsicherheit hinsichtlich der Einbeziehung entsprechender geeigneter Instrumente in die Planung und Gestaltung klinischer Prüfungen zu vermindern.

Es wird daher darum gebeten, den Begriff „etabliertes Instrument“ näher zu definieren, die aus Sicht des IQWiG bekannten „etablierten Instrumente“ zu benennen und darzustellen hinsichtlich welcher Indikationsgebiete diese verwendet werden können.

Dies auch vor dem Hintergrund, dass mehrere aus Sicht des betreffenden pharmazeutischen Unternehmers etablierte und standardisierte Instrumente wie EQ-5D (EuroQol; präferenzbasiertes Verfahren zur Messung der gesundheitsbezogenen Lebensqualität) und den SF-36 (Short Form (36) Gesundheitsfragebogen/ Erhebung der gesundheitsbezogenen Lebensqualität) seitens des IQWiG in Bewertungsverfahren nicht akzeptiert wurden.

Berlin, 7.08.2014

A.1.7 – Bundesverband Medizintechnologie e. V. (BVMed)

IQWiG - Institut für Qualität und
Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8 (KölnTurm)
50670 Köln

Bundesverband
Medizintechnologie e.V.
Reinhardtstraße 29b
10117 Berlin
Tel. +49 (0)30 246 255 - 0
Fax +49 (0)30 246 255 - 99
info@bvmed.de
www.bvmed.de

Berlin, 06.08.2014

**Stellungnahme zum Entwurf des allgemeinen Methodenpapiers
Version 4.2 des IQWiG**

Sehr geehrte Damen und Herren,

das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) hat den Entwurf für den nächsten Überarbeitungsschritt seiner Allgemeinen Methoden vorgelegt. Der Entwurf für das Methodenpapier 4.2 umfasst insbesondere Methoden zur Potenzialbewertung zu neuen Untersuchungs- und Behandlungsmethoden (§ 137e SGB V) sowie die überarbeiteten Methoden zur Kosten-Nutzen-Bewertung (KNB).

Nach Prüfung des vorliegenden Entwurfs, insbesondere der Kapitel, die sich mit der Potenzialbewertung befassen, gibt es aus unserer Sicht zu den im Methodenpapier 4.2 gemachten Vorschläge des IQWiG in rechtlicher und methodischer Hinsicht Bedenken, insbesondere genügen die Vorschläge nicht den Anforderungen an eine hinreichende Bestimmtheit und Klarheit einer Regelung und werfen vor allem in kompetenzrechtlicher Hinsicht die Frage auf, ob das IQWiG zur Bewertung in dieser Form befugt ist.

Zusammengefasst sind folgende wesentlichen Aspekte aus unserer Sicht zur Potenzialbewertung von wesentlicher Bedeutung und sind daher bei der Überarbeitung zu berücksichtigen:

1. Das Methodenpapier in seiner vorgelegten Entwurfsfassung beinhaltet eine Vielzahl unbestimmter Rechtsbegriffe, deren Auslegung mit erheblichen Schwierigkeiten verbunden ist und die daher in der Praxis zu Unsicherheiten in der Rechtsanwendung, aber auch im konkreten Nachweis eines entsprechenden Potenzials führen dürften.
2. Die Abstufungen der Ergebnissicherheit dürften in der Praxis ebenfalls für Umsetzungsschwierigkeiten sorgen, da nicht klar ist, wie die einzelnen Abstufungsgrade tatsächlich umzusetzen sind bzw. wie eine konkrete Bewertung diesbezüglich erfolgt. Gerade die unpräzise Formulierung der Evidenzgrade ist zudem nicht nachvollziehbar. Es bestehen erhebliche Bedenken im Hinblick auf die zu fordernde Rechtssicherheit, da insbesondere die Formulierung nur ungefährender Grenzen beträchtlichen Raum für Interpretationen gibt und damit nicht klar ist, wie ein entsprechendes Bewertungsverfahren tatsächlich gestaltet wird.

3. Darüber hinaus ist nicht klar, ob auch das IQWiG an dem bereits vom G-BA eingeführten sogenannten „doppelten Potenzialbegriff“ festhalten möchte. Eine diesbezügliche Umsetzung durch das IQWiG wäre mit der Rechtsgrundlage in § 137e SGB V nicht zu vereinbaren und daher unzulässig.
4. Hinsichtlich des Inhalts geht zudem die Prüfungskompetenz des IQWiG über den durch § 137e SGB V gezogenen Rahmen hinaus.

Ergänzend ist für uns nicht erklärlich, ob und inwiefern die in der Potenzialbewertung erwähnten Fragen in den Kapitel 3. und 4. zum Tragen kommen. Zudem bleibt bei der Potenzialbewertung unklar, wie die herangezogenen getroffenen Feststellungen für eine nachfolgende Nutzenbewertung verpflichtend sind und wie diese multiplen Zielkriterien zu einer indikationsübergreifenden Kennzahl für das „Potenzial“ aggregiert werden, um dann auf Basis dieser Information eine Entscheidung über die Durchführung einer Erprobungsstudie zu treffen. Es bleibt auch unklar, in welchem Umfang sich die Methode der Potenzialbewertung auf die methodischen Erläuterungen in Kapitel 3 und Kapitel 4 bezieht.

Im Folgenden gehen wir detailliert auf die einzelnen Kapitel des Methodenpapiers ein:

1. Antragsinhalte und Arbeitsauftrag

Kapitel 2.1.5 Seite 25 und 26

Auf Seite 25 des Methodenpapiers wird ausgeführt, dass die Antragsinhalte in gleicher Weise wie bei Dossierbewertungen erfolgen, allerdings unter „Berücksichtigung der spezifischen Anforderungen an die Wahrung der hohen Vertraulichkeit im Rahmen von Potenzialbewertungen“.

Hierbei ist nicht klar, was tatsächlich unter „spezifischen Anforderungen“ gemeint ist und wie diese im Einzelfall ausgelegt werden müssen. Vor allem wird nicht deutlich, in welcher Form diese spezifischen Anforderungen tatsächlich berücksichtigt werden bzw. berücksichtigt werden können und welche Unterschiede das IQWiG zwischen Potenzialbewertung einerseits und Nutzenbewertung andererseits sieht und welche Konsequenzen hieraus am Ende tatsächlich zu ziehen sind.

Zudem wird ausgeführt, dass das IQWiG optional zur Unterstützung der Bewertung eine eigene Recherche durchführen kann.

Unklar bleibt dabei, in welchen Fällen tatsächlich eine eigene Recherche durch das IQWiG vorgesehen ist und wie weit diese geht. Vor allem ist nicht geklärt, wie das Verhältnis zwischen eigener Recherche und Recherche des Antragstellers zu sehen ist und welche Grenzen dem IQWiG hierbei gezogen werden bzw. gezogen werden müssen. Dies gilt umso mehr, als das Methodenpapier selbst betont, dass Unterschiede zur Nutzenbewertung bestehen, so dass fraglich ist, ob auf die allgemeinen Grundsätze der Nutzenbewertung zurückgegriffen werden kann.

2. Vertraulichkeit

Weiter heißt es im Methodenpapier auf Seite 26, dass eine Publikation der Potenzialbewertung nicht erfolge, weil das Antragsverfahren gemäß § 137e SGB V der hohen Vertraulichkeit unterliege. Erst wenn im weiteren Verlauf der G-BA eine Erprobungsrichtlinie erlässt, wird auch die Potenzialbewertung veröffentlicht.

Die Frage des Geheimhaltungsbedürfnisses ist bislang noch unzureichend geklärt, zumal § 137e SGB V keine eindeutigen Aussagen hierzu trifft. In jedem Falle muss sichergestellt sein, dass die vertraulichen, durch die Unternehmen übermittelten Angaben auch vertraulich behandelt werden und keine Veröffentlichung durch das IQWiG selbst erfolgt – unabhängig davon, wann und in welcher Form dies geschieht und ob eine Erprobungsrichtlinie erlassen wurde oder nicht. Dies ist im Rahmen des Methodenpapiers ausdrücklich klarzustellen.

3. Potenzialbegriff:

Nach Auffassung des IQWiG zielen Potenzialbewertungen im Gegensatz zu Nutzenbewertungen darauf ab, zu prüfen, ob neue Untersuchungs- oder Behandlungsmethoden möglicherweise einen Nutzen aufweisen. Potenzial hierbei bedeutet, dass erstens die bisher vorliegenden Erkenntnisse einen möglichen Nutzen erkennen lassen und zweitens auf Basis dieser Erkenntnisse eine Studie geplant werden kann, die eine Bewertung des Nutzens der Methode auf einem ausreichend sicheren Erkenntnisniveau erlaubt.

Fraglich ist, wie weitgehend die Formulierung des Potenzialbegriffs zu verstehen ist. Eine über den Gesetzeswortlaut des § 137e SGB V hinausgehende Prüfungscompetenz hat der Gesetzgeber weder für den G-BA noch das IQWiG vorgesehen:

Mit dem zum 1. Januar 2012 in Kraft getretenen Gesetz zur Verbesserung der Versorgungsstrukturen in der gesetzlichen Krankenversicherung ("GKV-Versorgungsstrukturgesetz") hat der Gesetzgeber in § 137e SGB V das Erprobungsverfahren für neue Untersuchungs- und Behandlungsmethoden eingeführt. Dieses Verfahren betrifft solche neuen Untersuchungs- und Behandlungsmethoden, deren Nutzen noch nicht hinreichend belegt ist, die allerdings ein Potenzial bieten. Auf diese Weise hat der Gesetzgeber ein Instrument geschaffen, das einen schnelleren Zugang innovativer Untersuchungs- und Behandlungsmethoden in Deutschland ermöglichen soll.

Entscheidende Voraussetzung für das Erprobungsverfahren ist, ob das vom Gesetzgeber geforderte Potenzial der jeweiligen Methode besteht. Dieses Erfordernis ist eine zentrale Frage im Erprobungsverfahren. Bereits der G-BA geht aber (in Abweichung zu § 137e SGB V) davon aus, dass eine Erprobung einer neuen Methode nur dann in Betracht komme, wenn sich – neben dem vom Gesetz verlangten "Potenzial einer erforderlichen Behandlungsalternative" – zusätzlich das "Potenzial einer Erprobung" belegen lässt. Diese Auffassung aber ist rechtswidrig, da ein "doppeltes Potenzialerfordernis" von den gesetzlichen Vorgaben des § 137e SGB V abweicht. Nach diesseitiger Auffassung hängt die Durchführung eines Erprobungsverfahrens allein von einem einzigen Potenzialerfordernis ab.

Es sprechen allerdings gewichtige Gründe dafür, dass es sich hierbei lediglich um sprachlich verschiedene Ausprägungen ein und desselben Erfordernisses handelt, das einheitlich sowohl für selbständige beantragte wie auch im Laufe eines regulären Bewertungsverfahrens eingeleitete Erprobungsverfahren gilt. Ein darüber hinausgehendes Potenzialerfordernis sieht das Gesetz demgegenüber nicht vor.

In der Folge sind weder G-BA noch IQWiG befugt, zu Lasten der betroffenen Medizinproduktehersteller über das Gesetz hinausgehende restriktivere Anforderungen an die Durchführung des Erprobungsverfahrens zu stellen. Vor diesem Hintergrund bedarf es der expliziten Klarstellung der Reichweite der Beurteilung durch das IQWiG im Rahmen der bestehenden Kompetenzen. Dies gilt nicht zuletzt auch für die Frage der Definition des Potenzials selbst. Zwar ergeben sich hierbei gewichtige Anhaltspunkte aus der Gesetzesbegründung selbst, in methodischer Hinsicht sollte sich jedoch gerade das gesetzgeberische Anliegen ebenfalls widerspiegeln. Eine die Auffassung des G-BA bestätigende Methodik jedenfalls ließe sich mit der derzeit geltenden Regelung in § 137e SGB V nicht vereinbaren.

4. Potentialbewertung

Kapitel 3.8 Potenzialbewertung ab Seite 74 ff.

Im Entwurf des IQWiG heißt es auf Seite 74, dass eine Potenzialbewertung gemäß § 137 e Abs. 7 SGB V auf einem Antrag beruhe, für den der G-BA form- und notwendige Inhalte definiert hat.

Eine Erprobung einer neuen Methode wird entweder (selbständig) auf Antrag oder (unselbständig) im Rahmen eines laufenden regulären Bewertungsverfahrens gemäß §§ 135, 137c SGB V durchgeführt. Hier ist allerdings fraglich, in welchem Zusammenhang das IQWiG § 137e Abs. 7 SGB V zu § 137e Abs. 1 SGB V sieht. Insbesondere ist nicht klar, ob und inwiefern eine Differenzierung zwischen dem Antragsverfahren einerseits und dem allgemeinen Bewertungsverfahren gemäß § 137e Abs. 1 SGB V andererseits vorzunehmen ist. Weiterhin heißt es auf Seite 74, dass im Rahmen der Potenzialbewertung vom Institut die Plausibilität der Angaben des Antragstellers geprüft werden. Diese Prüfung bezieht sich insbesondere auf die Sinnhaftigkeit der im Antrag dargestellten medizinischen Fragestellungen und die Güte der vom Antragsteller durchgeführten Literatursuchen. Hier sollte klargestellt werden, dass die Hoheit über das gesamte Bewertungsverfahren beim G-BA verbleibt und auch verbleiben muss.

Allgemein nämlich heißt es in § 137e Abs. 5 SGB V, dass die wissenschaftliche Begleitung und Auswertung der Erprobung durch eine wissenschaftliche Institution erfolgen kann. Das heißt allerdings auch, dass es sich lediglich um eine begleitende Unterstützung, nicht aber um eine komplette Verfahrensgestaltung handelt und schlussendlich die Letztentscheidung und das Letztbestimmungsrecht ausschließlich beim G-BA verbleiben müssen. Insofern ist fraglich, ob insbesondere eine umfassende Bewertung der medizinischen Sinnhaftigkeit der vom Antragsteller aufgeworfenen Fragen in dieser Form noch zulässig ist, da möglicherweise das Entscheidungsrecht des G-BA beeinträchtigt sein könnte.

Darüber hinaus heißt es auf Seite 74, dass auf Grund der besonderen Zielsetzung bei der Potenzialbewertung im Vergleich zur Nutzenbewertung deutliche niedrigere Anforderungen an die Evidenz zu stellen sind. Das Ziel der Erprobung besteht darin, eine adäquate Datengrundlage für eine zukünftige Nutzenbewertung herzustellen. Dementsprechend lasse sich ein Potenzial insbesondere auch auf der Basis nicht randomisierter Studien begründen.

Die Verwendung einer Vielzahl unbestimmter Rechtsbegriffe und die Heranziehung unklarer Definitionen führen in der praktischen Umsetzung zu einer erheblichen Rechtsunsicherheit, die einem transparenten und zielgerichteten Verfahren abträglich sind. Trotz der bestehenden Definitionsansätze ist nicht klar, was etwa unter „deutlich niedrigeren“ Evidenzanforderungen zu verstehen ist. Auch wird nicht erläutert, was tatsächlich eine adäquate Datengrundlage ist, wie diese zu bestimmen ist und welche Faktoren hierfür maßgeblich sind. Generell ist daher zu fordern, dass auf die Verwendung unbestimmter Rechtsbegriffe soweit wie möglich verzichtet wird und statt dessen klare und hinreichend nachvollziehbare Definitionen und Formulierungen verwendet werden, die es insbesondere dem Anwender ermöglichen, die exakten Anforderungen des Verfahrens abzuschätzen, um damit die Anforderungen an einen hinreichenden Nutzen- bzw. Potenzialbeleg zu erfüllen. Dies gilt nicht zuletzt auch für die Frage, was unter „adäquater Kontrolle“ (vgl. S. 75) zu verstehen ist.

5. Unklare Definition des Verzerrungspotenzials

Auf Seite 76 wird ausgeführt, dass auch hochwertige, nicht randomisierte Studien ein erhebliches Verzerrungspotenzial bergen, so dass bei der Ableitung eines Potenzials aus solchen Studien geprüft werden müsse, ob die vorliegenden Studien bezüglich der interessierenden Intervention Unterschiede in einer Größenordnung zeigen, die erwarten lassen, dass durch geeignete zukünftige Studien ein Nutzen nachgewiesen werden kann, und nicht mehr allein durch einen durchschnittlich erwartbaren Einfluss von Verzerrung erklärt werden können. Daher soll sich ein Potenzial insbesondere dann ergeben, wenn Studien geringer Ergebnissicherheit mindestens kleine Effekte zeigen, wenn Studien sehr geringer Ergebnissicherheit mindestens mittlere Effekte zeigen oder wenn Studien minimaler Ergebnissicherheit mindestens große Effekte zeigen.

Auch wenn eine Anpassung des Bewertungsverfahrens an die Erfordernisse der Potenzialbewertung grundsätzlich zu begrüßen ist, so muss doch dafür Sorge getragen werden, dass das Verfahren mit hinreichend klaren Definitionen unterlegt ist. Gerade im Hinblick auf das zu erwartende Verzerrungspotenzial aber ist die Formulierung unscharf. Insbesondere ist sie nicht geeignet, klare Abgrenzungskriterien aufzustellen und die aufgestellten Behauptungen zu belegen.

Dies gilt nicht zuletzt auch für die Definition der Grenzen zwischen kleinen, mittleren und großen Effekten. Hier heißt es, dass ungefähre Grenzen für das relative Risiko Werte von 0,8 und 0,5 sein können. Woher sich diese Werte allerdings ableiten, ist nicht in der zu fordernden Schärfe und Eindeutigkeit belegt. Zudem ist die Formulierung einer „ungefähren“ Grenze nicht geeignet, Klarheit über das zu erwartende Verfahren herzustellen und damit Rechtssicherheit zu gewährleisten. Die Einschränkung bzw. die Bewertung der Frage der Ergebnissicherheit findet sich in dieser Form nicht im Gesetzestext wieder. Gerade die Abstufungen der verschiedenen Stufen der Ergebnissicherheit dürften daher in der Praxis für erhebliche Unsicherheit sorgen, so dass genaue und klare Abgrenzungen zu begrüßen wären. Dabei sollten auch die gesetzgeberischen Wertungen berücksichtigt werden, nachdem von einem Potenzial dann nicht ausgegangen werden kann, wenn Hinweise auf erhebliche Patientenschäden gegeben sind. Wann jedoch von einem Potenzial auszugehen ist, ergibt sich aus dem Methodenpapier nicht. Zugleich ist fraglich, ob die Graduierung der Ergebnissicherheit in Anlehnung an die bisherige Evidenzgüteklassifizierung tatsächlich noch vom Gesetzeswortlaut gedeckt wird. Zudem darf darauf hingewiesen werden, dass bereits die gewählte Abstufung und die dazu erfolgende Erläuterung den Eindruck einer Stigmatisierung der vorliegenden Daten hervorrufen, ohne auf die Besonderheiten der Methodenbewertung einerseits und jenen der Medizinprodukte und der Untersuchungs- und Behandlungsmethoden näher einzugehen.

6. Endpunkte

Problematisch ist der vorliegende Entwurf auch im Hinblick auf die Frage der relevanten Endpunkte bei der Potenzialbewertung.

Hier heißt es auf Seite 76, dass abweichend zum Vorgehen bei Nutzenbewertungen auch Surrogatendpunkte berücksichtigt werden können, für die noch keine hinreichende Validität gezeigt werden konnte. Surrogatendpunkte sollen jedoch etabliert und plausibel sein, um hierüber ein Potenzial begründen zu können.

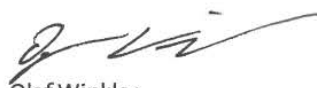
Zwar ist die Zulassung von Surrogatendpunkten grundsätzlich zu begrüßen. Jedoch ist es erforderlich, klare Erläuterungen dahingehend aufzustellen, inwiefern ein abweichendes Vorgehen zur Nutzenbewertung erforderlich ist und welche Surrogatendpunkte tatsächlich zugelassen sind. Aus der vorliegenden Formulierung lässt sich nicht ableiten, wann von einer Etablierung und Plausibilität im Sinne des Methodenpapiers ausgegangen werden kann. Die Problematik der Surrogatendpunkte erscheint damit im Ergebnis unklar und nicht geeignet, für Rechtssicherheit zu sorgen. Insofern bedarf es ebenfalls einer näheren Erläuterung dahingehend, wann von einer Etablierung und Plausibilität auszugehen ist und wie die hinreichende Validität der Surrogatendpunkte bestimmt werden kann. Hierbei ist insbesondere der aktuelle Stand der medizinischen Erkenntnisse maßgeblich, so dass nicht zuletzt auch einer entsprechenden Expertise der Fachgesellschaften und deren zwingende Einbindung bedarf.

Mit freundlichen Grüßen

BYMed - Bundesverband
Medizintechnologie e. V.



Joachim M. Schmitt
Geschäftsführer
Mitglied des Vorstands



Olaf Winkler
Leiter Referat Gesundheitssystem

**A.1.8 – Deutsche Diabetes Gesellschaft e. V. (DDG) und
Deutsche Gesellschaft für Kardiologie e. V. (DGK)**

Deutsche Diabetes Gesellschaft (DDG)
Deutsche Gesellschaft für Kardiologie (DGK)

Zum Entwurf

Allgemeine Methoden Version 4.2 des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) vom 18.06.2014.

Die Deutsche Diabetes Gesellschaft (DDG) und die Deutsche Gesellschaft für Kardiologie (DGK) sehen den sinnvollen und wichtigen Stellenwert des Arzneimittelmarktneuordnungsgesetzes (AMNOG) sowie des Verfahrens mit dem Gemeinsamen Bundesausschuss (G-BA) und dem Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Wir betrachten aber weiterhin mit großer Sorge die Entwicklungen, die bei der Nutzenbewertung bestehender und neuer Medikamente erkennbar wurden. Durch diese Entwicklungen wird die Versorgung der Patienten entscheidend beeinflusst und die ärztliche Therapiefreiheit teilweise eingeschränkt. Wir nehmen deshalb das Stellungnahmeverfahren des IQWiG zur Überarbeitung der „Allgemeinen Methoden Version 4.2“ zum Anlass, grundlegend zum Prozess der frühen Nutzenbewertung nach §35a SGBV Stellung zu nehmen. Die Stellungnahme richtet sich deshalb nicht nur an das IQWiG, sondern gleichermaßen an den G-BA.

Wir denken, dass bei diesem gesetzlich geregelten prozeduralen Verfahrensweg auch in der allgemeinen Methodik des IQWiG berücksichtigt werden muss, dass der vom G-BA potentiell Beauftragte (IQWiG) eine Verantwortlichkeit bezüglich der Plausibilitätsprüfung von Vergleichstherapie, Fragestellung und Ergebnis hat. Bei dieser Plausibilitätsprüfung würde sich ein sogenannter „strukturierter Dialog“ und eine Einbindung der wissenschaftlichen Fachgesellschaften oder bei übergreifenden Themen die Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) anbieten.

Die beiden Fachgesellschaften vertreten den allgemeinen internationalen Stand des medizinischen Wissens für die Therapie von Volkskrankheiten und Krankheitsbildern, die mit zu den häufigsten Todesursachen in Deutschland gehören. Beide Fachgesellschaften – in der Regel in Abstimmung mit europäischen Partnergesellschaften – geben, basierend auf der aktuellen wissenschaftlichen Evidenz, Leitlinien heraus.

Wie bereits in unserer Stellungnahme vom 22. Mai 2013 zur Version 4.1 dargelegt, die leider in keiner Weise in der Version 4.2 berücksichtigt worden ist, sind aus unserer Sicht bei der Nutzenbewertung durch das IQWiG häufiger Diskussionen, unterschiedliche Ansichten, Streitigkeiten bis hin zu Fehleinschätzungen entstanden. Wir sehen durch Beachtung der folgenden fünf Punkte ein Potential zur qualitativen Verbesserung des Evaluationsprozesses:

1. Festlegung der zweckmäßigen Vergleichstherapie
2. Formulierung der Fragestellung
3. Festlegung und Bewertung des Zusatznutzens
4. Einbindung von externen Gutachten
5. Begründung bei Abweichung von Leitlinienempfehlungen

Daher bitten wir bei der Überarbeitung des Methodenpapiers und der künftigen Arbeit des IQWiG um Berücksichtigung dieser schon damals angemahnten fünf Punkte. Eine Präzision dieser Punkte findet sich in unserer Stellungnahme vom 22.05.2013 (1).

Der Grund für die häufig vorkommende diskrepante Einschätzung eines eventuellen Zusatznutzens zwischen den Fachgesellschaften und dem IQWiG liegt aus unserer Sicht darin begründet, dass das IQWiG ausschließlich die Studienlage bzw. Evidenz bewertet, die mit dem Zulassungstext der Fachinformation konform geht. Dies ist kein plausibler und zielführender Umgang mit der Evidenz in Bezug auf die wissenschaftliche Bewertung einer Substanz. Zudem müssten die Bewertungen dann auch entsprechend formal richtig formuliert werden, d. h. Zusatznutzenbewertung bezogen auf den Text der Fachinformation.

Daher empfehlen wir, dass neben der bisherigen, allein auf die Fachinformation bezogenen Bewertung, in transparenter Weise auch eine allgemeine wissenschaftliche Beurteilung bzw. Stellungnahme durch das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) berücksichtigt wird. Das BfArM sollte auch systematisch bei der Anhörung eines Verfahrens zur Verfügung stehen.

Es ist international akzeptiert, dass bei Fehlen einer direkt auf den Zulassungstext bezogenen Evidenz, die bestmögliche Evidenz berücksichtigt werden sollte (2). Wir legen großen Wert darauf, dass sich das IQWiG diesen internationalen Wert-Maßstäben in der Beurteilung von neuen Substanzen im Sinne der Versorgung unserer Patienten anschließt.

Zudem halten wir es für erforderlich, dass die wesentlichen medizinischen Inhalte und Beurteilungen der „medizinisch-fachlichen Beratung“ transparent im Bericht der Nutzenbewertung dargelegt werden. Dies schließt auch die angeforderten externen Gutachten ein.

Wir sind der Überzeugung, dass nicht nur aus wissenschaftlichen und klinischen Gründen sowie Gründen der demokratischen Legitimation die wissenschaftlichen Fachgesellschaften durch ein transparentes Verfahren bei den o. a. Punkten berücksichtigt werden müssen. Eine entsprechende Optimierung des Prozesses unter Berücksichtigung der o.g. Punkte wird die politische Akzeptanz der Ergebnisse des AMNOG auch beim Gesetzgeber, den Kostenträgern, den Leistungserbringern sowie insbesondere auch bei den betroffenen Patienten, ihren Angehörigen und bei der Bevölkerung erhöhen.

Literatur

1. http://www.deutsche-diabetes-gesellschaft.de/fileadmin/Redakteur/Stellungnahmen/220513_Gemeinsame_Stellungnahme_zum_Methodenpapier_des_IQWiG.pdf
2. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfland M, et al. GRADE guideline: 8. Rating the quality of evidence-indirectness. J. Clin. Epidemiol 2011; 64 (12): 1303 – 1310

Für die Deutsche Diabetes Gesellschaft (DDG)

PD Dr. Erhard Siegel (Präsident)

Prof. Dr. med. Baptist Gallwitz (Vizepräsident)

Prof. Dr. med. Dirk Müller-Wieland (Vorstand und Sprecher der Kommission Wissenschaftliche Stellungnahmen)

Für die Deutsche Gesellschaft für Kardiologie (DGK)

Prof. Dr. Christian Hamm (Präsident)

Prof. Dr. Heribert Schunkert (Vorstand der Klinischen Kommission)

A.1.9 – Deutsche Gesellschaft für Gesundheitsökonomie e. V. (DGGÖ)

Stellungnahme der Deutschen Gesellschaft für Gesundheitsökonomie zu den inhaltlichen Aktualisierungen und Ergänzungen im Entwurf der Version 4.2 Allgemeine Methoden

28. Juli 2014

Vorbemerkung

Knapp fünf Jahre nach der Vorstellung der „Aktuellen Methoden der Kosten-Nutzen-Bewertung (Version 1.0)“ erneuert das IQWiG seine ökonomischen Methoden und integriert im Entwurf 4.2 des Grundlagenpapiers „Allgemeinen Methoden“ erstmals ein eigenes Kapitel zur „Kosten-Nutzen-Bewertung medizinischer Interventionen“. In den „Allgemeinen Methoden“ legt das Institut seine grundlegende Arbeitsweise und die zugehörigen wissenschaftlichen Grundsätze dar.

Eine Erweiterung des Grundlagenpapiers um Methoden der ökonomischen Evaluation ist sehr zu begrüßen, da sie ein markantes Defizit im evidenzbasierten Management im deutschen Gesundheitssystem angeht. Das neue Kapitel 4 bietet dafür einen Rahmen. Zu einigen Punkten nimmt die Deutsche Gesellschaft für Gesundheitsökonomie dggo aus Sicht der disziplinar zugehörigen Fachgesellschaft Stellung.

Rechtliche Zuständigkeit (Abschnitt 4.1.1)

Der Entwurf stellt in 4.1.1 die rechtlichen Grundlagen einer Beauftragung des IQWiG dar. Der Auftrag einer Kosten-Nutzen-Bewertung kann dabei unterschiedlich interpretiert werden:

- a) In der ersten Interpretation erarbeitet das Institut ausschließlich die wissenschaftliche Evidenz zu Kosten und Nutzen und bewertet deren wissenschaftliche Qualität („Assessment“). Die so bewertete Evidenz wird Entscheidungsträgern wie dem GBA zur Verfügung gestellt.
- b) In der zweiten Interpretation trifft das Institut zusätzlich zur Interpretation a) Aussagen zur Bewertung des Wirtschaftlichkeitsergebnisses, d.h. dazu, ob das Verhältnis der zusätzlichen Kosten zum zusätzlichen Nutzen als wirtschaftlich oder nicht wirtschaftlich beurteilt wird („Appraisal“).

Grundsätzlich wird international eine klare Trennung zwischen Assessment und Appraisal empfohlen, um die Beurteilung der methodischen Qualität der vorliegenden Evidenz nicht mit der Bewertung hinsichtlich einer angemessenen Kosten-Nutzen-Relation oder eines angemessenen Preises zu vermischen. Formal sollte es nicht in das Ermessen des IQWiGs gestellt werden, ob es zusätzlich zum „Assessment“ auch das „Appraisal“ übernimmt. Diese zentrale Entscheidung sollte entweder generell oder auftragsspezifisch klargestellt werden.

Bei einem reinen „Assessment“ überließe das IQWiG die Beurteilung, ob die Relation von Kosten und Nutzen der untersuchten Intervention akzeptabel ist, vollständig dem GBA oder den Partnern der Preisverhandlungen bei Arzneimitteln. Das Methodenpapier legt die zweite Interpretation nahe, bei der das IQWiG auch das „Appraisal“ übernimmt.

Generell wird mit einer Bewertung des Wirtschaftlichkeitsergebnisses und auch zur Zahlungsbereitschaft für den Zusatznutzen Stellung genommen. Um sozialrechtliche

Transparenz zu schaffen, müsste das IQWiG bei Übernahme eines „Appraisal“ Entscheidungsträger, Kontext und normative Grundlage, die mit Aussagen zur Zahlungsbereitschaft adressiert werden, genau benennen. Jedoch kann sich das IQWiG bei Aussagen zur Zahlungsbereitschaft weder auf normative Vorgaben der nach SGB zuständigen Entscheidungsträger wie GKV-Spitzenverband und GBA stützen, noch lässt sich deren Zahlungsbereitschaft wissenschaftlich errechnen, noch ist eine Beauftragung des Instituts durch den Gesetzgeber erkennbar, anstelle der Entscheidungsträger die Zahlungsbereitschaft zu formulieren. Daher sollte sich das IQWiG mit Ausnahme der Fälle, in denen eine Intervention ökonomisch dominiert oder durch eine andere dominiert wird, auf das „Assessment“ beschränken.

Interpretation der Ergebnisse (4.1.9) sowie Besonderheiten der Kosten-Nutzenbewertung nach § 35b SGB V (4.9.2-4)

Für die Beurteilung der Wirtschaftlichkeit einer neuen Intervention hält das IQWiG am Konzept der Effizienzgrenze fest. Ist ein Zusatznutzen gegeben, wird die Effizienzgrenze durch lineare Extrapolation der Steigung des letzten beobachteten Segments bestimmt. Der extrapolierte Teil dient der Bewertung der Wirtschaftlichkeit, also einem „Appraisal“; für ein „Assessment“ ist er nicht erforderlich. Für das „Appraisal“ der Entscheidungsträger ist eine Orientierung an der extrapolierten Effizienzgrenze aber aus gesundheitsökonomischer Perspektive in mehrfacher Hinsicht problematisch.

Das IQWiG-Konzept unterstellt erstens, dass die medizinischen Interventionen auf der Effizienzgrenze Zahlungsbereitschaften in einer effizienten Situation widerspiegeln. Im Arzneimittelbereich wird die Extrapolation der Effizienzgrenze vielfach auf Produkten aufbauen müssen, die bereits vor Einführung des AMNOG erstattet wurden. Damit wird missachtet, dass der Gesetzgeber das AMNOG geschaffen hat, weil er die Preisgestaltung neuer Arzneimittel verbessern wollte und eben nicht von einer effizienten Situation ausging. Zweitens kann man nicht unterstellen, dass die Effizienzentwicklung gleichmäßig über alle Indikationsgebiete verteilt ist. Bestehende Effizienzunterschiede zwischen Indikationsgebieten würden aber durch eine indikationsspezifische, historische Extrapolation weiter fortgeschrieben. Obige Argumente gelten auch für andere Bereiche der medizinischen Versorgung.

Wie seit langem bekannt, lässt sich ein auf der Extrapolation dieser Effizienzgrenze beruhendes Urteil über die Wirtschaftlichkeit einer neuen Intervention aus einer Reihe von Gründen ökonomisch nicht fundieren (vgl. deutsche und internationale Stellungnahmen 1-4). Auch gibt es keine empirischen Anwendungen zum Konzept der Effizienzgrenze außerhalb exemplarischer IQWiG Aktivitäten.

Die Extrapolation der Effizienzgrenze ist theoretisch nicht haltbar, praktisch unausgereift und in der gesundheitsökonomischen Evaluationsforschung nicht verbreitet. Sie kann die Anforderungen des §35b des SGBV nicht erfüllen, die Methoden und Kriterien für die Erarbeitung von Bewertungen auf die Grundlage der in den Fachkreisen anerkannten internationalen Standards zu stellen. Es wird daher empfohlen, auf Bewertungen zu verzichten, die auf einer Extrapolation der Effizienzgrenze beruhen. Dies gilt auch für Aussagen, die aus der Extrapolation der Effizienzgrenze abgeleitet werden wie die Darstellung der Unsicherheit in den Abschnitten 4.9.2-4. Die dggo wiederholt damit eine Forderung, die sie bereits 2009 im Rahmen ihrer Stellungnahme zur Version 2.0 der Methodik des IQWiG abgegeben hat.

Maß des Gesamtnutzens/Das QALY als Maß (4.3.3 A)

Die Darstellung wirkt stark geprägt von der kontroversen Perception des QALY-Konzepts und reflektiert den Forschungsstand nicht in einer Form, die von einer Methodenleitlinie zu erwarten ist. Ein wichtiger Bestandteil der QALYs, aber auch ein zentrales Nutzenmaß nach SGBV § 35b, ist die gesundheitsbezogene Lebensqualität. Deren Messung und Bewertung bilden ein wichtiges Arbeitsgebiet der gesundheitsökonomischen Evaluationsforschung. Hier werden – neben der bereits vom IQWiG geforderten Relevanz der Bewertungen für deutsche Patienten – nachfolgende Präzisierungen empfohlen. Diese Präzisierungen sollten, um einen einheitlichen Methodenstandard zu gewährleisten, auch bei der Beurteilung der gesundheitsbezogenen Lebensqualität als Endpunkt in reinen Nutzenbewertungsverfahren zur Anwendung kommen:

- Die beschreibenden und bewertenden Bestandteile einer Messung der gesundheitsbezogenen Lebensqualität sollten im Indikationsbereich der Patientenzielgruppe auf ihre psychometrischen Eigenschaften wie etwa Validität, Reliabilität und Reagibilität erfolgreich getestet worden sein.
- Das IQWiG sollte im Methodenpapier darlegen, welche Bezüge es bei der Aggregation und Bewertung der gesundheitsbezogenen Lebensqualität akzeptiert oder als Referenzdarstellung erwartet:
 - (i) Welche Population (Patienten, Allgemeinbevölkerung, Experten) soll bewerten?
 - (ii) Was soll bewertet werden (aktuell erlebte oder hypothetische Gesundheitszustände)?
 - (iii) Mit welchen Verfahren soll bewertet werden (wahlbasiert oder direkt)?

Diese Vorgaben können bei der Konzeption von Studien und bei der Anfertigung von Dossiers berücksichtigt werden und damit mittelfristig die Evidenzqualität verbessern.

- Bei der Verwendung bewerteter, gesundheitsbezogener Lebensqualität als Parameter in entscheidungstheoretischen Modellen ist genau anzugeben, bei welcher Population, mit welchen Verfahren und mit welchen Bewertungsbezügen die Erhebung erfolgte. Neben diesen Transparenzaspekten ist auf die Konsistenz der Parameter und deren Kompatibilität mit den vorgenannten Erwartungen des IQWiG zu achten.

Literatur

- (1) Gemeinsame Stellungnahme (2009) des Ausschusses für Gesundheitsökonomie im Verein für Socialpolitik vertreten durch den Vorsitzenden und der Deutschen Gesellschaft für Gesundheitsökonomie (DGGÖ) vertreten durch den erweiterten Vorstand zum IQWiG Entwurf einer Methodik für die Bewertung von Verhältnissen zwischen Nutzen und Kosten im System der deutschen gesetzlichen Krankenversicherung, Version 2.0.
- (2) Wasem J (2008) Kosten-Nutzen-Bewertung von Arzneimitteln. Eine unvermeidbare Abwägung. Deutsches Ärzteblatt 105(9):A438-40.
- (3) Greiner W, Kuhlmann A, Schwarzbach C (2010) Ökonomische Beurteilung des Effizienzgrenzenkonzeptes. Gesundheitsökonomie und Qualitätsmanagement 15: 241–50.
- (4) Sculpher M, Claxton K (2010) Sins of omission and obfuscation: IQWiG's guidelines on economic evaluation methods. Health Economics 19(10):1132-6.

A.1.10 – Deutsche Gesellschaft für Gynäkologie und Geburtshilfe e. V. (DGGG)



DGGG e.V. • Hausvogteiplatz 12 • 10117 Berlin

Herrn
Wolfgang Müller, M.A.
AWMF-Geschäftsstelle
Ublerstr. 20
40223 Düsseldorf

Präsident

Prof. Dr. med. Thomas Dimpfl
Klinikum Kassel GmbH
Frauenklinik
Mönchebergstraße 41-43
D-34125 Kassel

Kassel, den 31.07.2014

**Stellungnahme der DGGG zum Entwurf des IQWiG
Allgemeine Methoden in Version 4.2 vom 18.06.2014**

Sehr geehrte Damen und Herren,

wir möchten Ihnen seitens der DGGG eine Stellungnahme zum Entwurf des IQWiG

Allgemeine Methoden in Version 4.2 vom 18.06.2014 vorlegen.

Im Abschnitt „2 Produkte des Instituts“ wird im Rahmen des Stellungnahmeverfahrens

zu ihrem Produkt Bericht nach § 139a SGB V bzw. deren Vorprodukte Berichtsplan
und Vorbericht „eine Frist von mindestens vier Wochen“ gegeben (Seite 15 und 17).

Diese enge Fristsetzung kann eine Fachgesellschaft und deren vom Vorstand
ehrenamtlich beauftragte Autoren vor echte zeitliche Engpässe stellen. Die

Fristsetzung für das Prozedere der Anfrage/Erstellung/Veröffentlichung einer

Stellungnahme z.B. auf einen IQWiG-Bericht ist bereits sehr eng gesetzt. Der Vorstand

der DGGG hat ein Verfahren zu koordinierten Stellungnahmeverfahren innerhalb

unserer Fachgesellschaft abgestimmt, das einen Zeitrahmen von ca. sechs Wochen für
eine fundierte Stellungnahme (auch gegenüber anderen Institutionen) kalkuliert.

Nun erwähnt Ihr Entwurf für das neu implementierte Produkt Kosten-Nutzen-Bewertung

„eine Frist von drei Wochen“ (Seite 24). Es erscheint utopisch, innerhalb von drei

Wochen eine fachlich fundierte vom Vorstand der Fachgesellschaft befürwortete

Stellungnahme zu erstellen und an Ihr Institut fristgerecht mit allen Unterschriften

Repräsentanz der DGGG und Fachgesellschaften
Hausvogteiplatz 12 • D – 10117 Berlin
Telefon: +49 (0) 30/51488 33 40 • Telefax +49 (0) 30/514 88 344
info@dggg.de • www.dggg.de



einzureichen. Des Weiteren verschärfen Sie Ihren Wortlaut mit dem Wegfall des Worts „mindestens“. Wir finden keinen Abschnitt, wo das öffentliche Stellungnahmeverfahren – auch für die anderen Produkte des IQWiG – auf Anfrage verlängert werden kann. Wir wünschen hier ganz klar eine Entschärfung der engen Fristsetzung auf mindestens 6 Wochen.

Bei Ihrem neuen Produkt der Potenzialbewertung zur Erprobung von Untersuchungs- und Behandlungsmethoden gemäß § 137e SGB V wird abschnittsweise zur inhaltlichen Bewertung nach Übermittlung durch den G-BA von einer Beratung durch „medizinische Expertise“ gesprochen (Seite 24 und 25). Diese gilt gemäß § 137e Absatz 2 SGB V nur für Krankenhäuser, die nicht an der Erprobung teilnehmen. Diese Beratung wird mit dem Wort „kann“ im Entwurf nicht obligat vorgeschrieben (Seite 24 und 25). Im Interesse der untersuchten und behandelten Patienten erscheint es sinnvoll die jeweilige Fachgesellschaft bzw. externe Sachverständigen immer mit einzubeziehen, auch wenn das § 137e SGB V nicht zwingend vorschreibt.

Freundliche Grüße

Prof. Dr. med. Thomas Dimpfl
Präsident der DGGG

Stellungnahme erstellt von: Prof. Matthias W. Beckmann, Leitlinienbeauftragter der DGGG

A.1.11 – Deutsche Gesellschaft für Hämatologie und medizinische Onkologie e. V. (DGHO) und Gesellschaft für Thrombose- und Hämostaseforschung e. V. (GTH)



Gesellschaft für Thrombose- und
Hämostaseforschung e.V.

Geschäftsstelle Hannover

Feodor-Lynen-Str. 5 • 30625 Hannover
Tel. 0511.532 8488
Fax: 0511 532 4147
mail@gth-online.org



Hauptstadtbüro der DGHO

Alexanderplatz 1 • 10178 Berlin
Tel. 030.27 87 60 89 - 0
Fax: 030.27 87 60 89 - 18
info@dgho.de

Institut für Qualität und Wirtschaftlichkeit

im Gesundheitswesen (IQWiG)

Im Mediapark 8 (KölnTurm)

50670 Köln

4. August 2014

Stellungnahme zu IQWiG Allgemeine Methoden Entwurf Version 4.2

- 1. Zusammenfassung**
- 2. Einleitung**
- 3. Diagnostische Verfahren (Kapitel 3.5)**
- 4. Kosten-Nutzen-Bewertung (Kapitel 4)**
- 5. Gesundheitsinformationen (Kapitel 6)**
- 6. Was fehlt?**
- 7. Literatur**

1. Zusammenfassung

Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) legt nach der Veröffentlichung seines Methodenpapiers vom November 2013 jetzt den Entwurf einer weiteren Aktualisierung vor. Schwerpunkte sind die Integration der Kosten-Nutzen-Bewertung und die Gesundheitsinformationen. Dazu wurde eine Reihe weiterer inhaltlicher Änderungen vorgenommen. Kritikpunkte unsererseits betreffen:

- Diagnostische Verfahren

Das Thema ist hoch aktuell. In der Onkologie gibt es eine Fülle neuer Angebote, vor allem von sog. Biomarkern. Das Methodenpapier beschränkt den Nutzen neuer diagnostischer Verfahren auf die nachfolgende Therapie oder präventive Maßnahmen. Es fehlt hier das Kriterium der Diagnosesicherung.

Bei den Studienformen wird das zunehmend häufig eingesetzte Instrument der retrospektiven Analyse von Biomarkern in prospektiv geplanten Studien nicht diskutiert.

- Kosten-Nutzen-Bewertung

Wir begrüßen die Integration dieses wichtigen Themas in das Methodenpapier. Ein konzeptioneller Fortschritt des gesamten Prozesses der Kosten-Nutzen-Bewertung ist

insbesondere die Aufnahme einer ausführlicheren Darstellung von Methoden zur gewichteten Bewertung von Endpunkten (Analytic Hierarchy Process, Conjoint Analyse). Diese Methodik kann auch die Basis für eine Hierarchisierung von Endpunkten im Rahmen anderer Verfahren der Nutzenbewertung sein.

Die Instrumente zur Berechnung des Gesamtnutzens, wie QALY u. a., werden im Methodenpapier nur beschrieben. Eine tiefergehende Auseinandersetzung ist Aufgabe des gesundheitspolitischen Diskurses, kann in Einzelverfahren auch im G-BA stattfinden. Im Methodenpapier ist eine klarere Festlegung auf das jeweils anzuwendende Berechnungsmodell wünschenswert.

- **Gesundheitsinformationen**

Unabhängige Informationen sind wichtig. Das IQWiG orientiert sich bei der Auswahl der zu kommentierenden Krankheitsbilder an einer Prävalenz von 1%. Diese Gesamtzahl wird dem diagnostisch und therapeutisch relevanten Verständnis der biologischen Heterogenität von Erkrankungen und Krankheitsverläufen nicht gerecht.

Die größte Herausforderung bei Gesundheitsinformationen ist die Aktualität der Inhalte. Der Aufwand ist hoch. In der Onkologie erhebt sich angesichts zahlreicher, bereits bestehender Patienten-orientierter Angebote die Frage nach dem Bedarf an weiteren, arbeitsintensiven Angeboten. Hier sind Kooperationen sinnvoll.

- **Was fehlt im Methodenpapier 4.2?**

Im Rahmen der bisherigen Verfahren zur Nutzenbewertung neuer Arzneimittel haben sich zahlreiche, z. T. gravierende Defizite der IQWiG Methodik gezeigt. Sie werden in dieser Aktualisierung leider nicht korrigiert.

2. Einleitung

Das IQWiG wird vom G-BA bei Bewertungsverfahren mit der Erstellung von Berichten beauftragt. Die vom IQWiG angewandte Methodik ist im Schriftstück ‚Allgemeine Methoden‘ dargelegt, zuletzt aktualisiert in der Version 4.1. Der jetzt vorgelegte Entwurf für die Version 4.2 ergänzt das bisherige Methodenpapier vor allem um das Thema der Kosten-Nutzenbewertung und eine ausführlichere Darstellung der Gesundheitsinformationen. Wir fokussieren im Folgenden auf die neuen Inhalte und greifen darüber hinaus einige Punkte auf, bei denen die bisherige Methodik weiterhin Defizite aufweist [1]. Letztere Themen sind nicht neu, sondern auch Inhalt des in den letzten Jahren intensiveren Austausches zwischen dem IQWiG, unserer und anderen Fachgesellschaften.

3. Diagnostische Verfahren (Kapitel 3.5)

Das Methodenpapier geht von der Grundannahme aus, dass der Wert diagnostischer Verfahren durch sich anschließende therapeutische oder präventive Verfahren definiert wird. Dem stimmen wir inhaltlich zu, es fehlt bei dieser Festlegung aber der ganze Bereich der Diagnosesicherung. In den Empfehlungen der Fachgesellschaften zur Diagnose und Therapie von hämatologischen und onkologischen Erkrankungen werden diagnostische Verfahren aufgenommen, wenn sie folgende Kriterien erfüllen:

1. erforderlich zur Diagnosesicherung;
2. prädiktiv für die anschließende Therapie oder präventive Verfahren;
3. qualitätsgesichert.

Ad 1) Diagnosesicherung

Diagnosen in der Hämatologie, Onkologie und der Hämostaseologie werden auf der Basis der aktuellen WHO-Klassifikation erstellt. Deren Erfüllung erfordert z. B. bei akuten Leukämien oder myeloproliferativen Neoplasien die Durchführung immunologischer, zytogenetischer oder molekularbiologischer Tests. Sie werden als krankheitsdefinierend bezeichnet [2]. Nicht jedes dieser Testverfahren ist unmittelbar Therapie-relevant. Sie sichern aber die Diagnose, verhindern in der Differenzialdiagnose die Durchführung weitergehender (überflüssiger) Untersuchungen, sind Voraussetzung für die Kodierung von Erkrankungen nach dem ICD-Code, und haben sozialmedizinische Bedeutung.

Methoden zur Diagnosesicherung nach nationalen und internationalen Standards sollen in das Methodenpapier aufgenommen werden.

Ad 2) Weiterer Nutzen für Therapie und/oder präventive Maßnahmen

In der Hämatologie, Onkologie und der Hämostaseologie gibt es aktuell fast täglich neue Angebote potenzieller Biomarker für die Labordiagnostik. Dazu kommt die intuitive Attraktivität neuer bildgebender Verfahren. In den Empfehlungen der Fachgesellschaften wird bei Erkrankten mit gesicherter Diagnose der Nachweis eines prädiktiven Nutzens neuer diagnostischer Verfahren für die anschließende Therapie gefordert. Das entspricht dem Vorgehen der Zulassungsbehörden wie EMA und FDA. Beispiele für den prädiktiven Wert molekularbiologischer Marker bei anschließender Therapie sind die zulassungsrelevanten Analysen aktivierender Genmutationen zum Einsatz von Kinase-Inhibitoren (*ABL* bei der Chronischen Myeloischen Leukämie, *EGFR* beim nicht-kleinzelligen Lungenkarzinom, *BRAF* beim Melanom).

In der Onkologie hat sich in den letzten Jahren der hohe Wert von prospektiv geplanten, retrospektiv durchgeführten Studien bei der Etablierung neuer Diagnostikverfahren gezeigt. Dieses Design beschreibt die spätere Durchführung von z. B. molekularbiologischen Analysen an asserviertem Tumorgewebe der Teilnehmer prospektiv geplanter, randomisierter Studien. Voraussetzungen für die Akzeptanz der Ergebnisse dieser Studienform sind:

- Analyse einer hohen Rate von Studienpatienten, z. B. >90%;
- balancierte Verteilung der Marker-positiven und Marker-negativen Patienten;
- hohe Studienqualität; falls möglich Analyse in mehr als einer Studie;
- Aktualität der in der Studie/den Studien durchgeführten Therapie;
- qualitätsgesicherter Test.

Herausragendes Beispiel für die unmittelbar patientenbezogene Relevanz derartiger Analysen sind die Bestimmung des *RAS*-Mutationsstatus beim Einsatz der Anti-EGFR-Antikörper Cetuximab [3] und Panitumumab [4] bei Patienten mit metastasiertem kolorektalem Karzinom. Die Kombinationstherapie mit diesen Antikörpern führt nur bei Patienten mit *RAS*-Wildtyp zu höheren Remissionsraten, zur Verlängerung des progressionsfreien und des Gesamtüberlebens. Patienten mit Genmutationen von *KRAS* oder *NRAS* haben unter einer Behandlung mit Anti-EGFR-Antikörpern möglicherweise sogar einen ungünstigeren Krankheitsverlauf. Diese Erkenntnisse wurden in retrospektiven Analysen von asserviertem Tumorgewebe der Patienten von mehreren prospektiv randomisierten Studien gewonnen und haben zur Änderung der Zulassungsbestimmungen von EMA und FDA mit Aufnahme einer obligaten Bestimmung von *RAS*-Mutationen geführt.

Aus den Formulierungen des Methodenpapiers 4.2 in Kapitel 3.5. kann herausgelesen werden, dass prospektiv geplante, retrospektiv durchgeführte Studien die Kriterien einer Hochwertigkeit erfüllen. Explizit benannt wird dieser Studientyp nicht. Dieses Studiendesign wird aufgrund der schnellen

Verfügbarkeit von Daten und angesichts zunehmend umfangreicher Biobanken mit asservierten Tumorproben zunehmen. Eine explizite Aufnahme und Reflektion dieses Studientyps im Methodenpapier 4.2 ist sinnvoll.

Ad 3) Qualitätssicherung

Bei der Qualitätssicherung neuer Tests gibt es zum einen die Möglichkeit der Festschreibung eines bestimmten Tests oder einer bestimmten Methodik. Dieses Vorgehen wurde in der jüngeren Vergangenheit mehrfach von der FDA bei der Festlegung von einem oder wenigen Testverfahren zur molekularbiologischen Diagnostik zum Einsatz neuer Medikamente gewählt. Alternative ist eine Offenheit bei der Wahl der Methodik, aber die Verpflichtung zur Teilnahme an qualitätssichernden, Ergebnis-orientierten Ringversuchen. Wir ziehen dieses Vorgehen vor.

4. Kosten-Nutzen-Bewertung (Kapitel 4)

Die Aufnahme einer Methodik zur transparenten Teilnahme an der Kosten-Nutzen-Bewertung nach SGB V ist begrüßenswert. Grundsätzlich ist die vorgegebene Struktur der Kosten-Nutzen-Bewertung nach SGB V sinnvoll, beginnend mit der Erstellung von Versorgungspfaden, Wahl des geeigneten Bewertungsmodells, Feststellung eines Gesamtnutzens, umfassende Berechnung der Kosten unter Berücksichtigung der Krankheitsdauer bis zur transparenten Ergebnisdarstellung.

Wir begrüßen in diesem Kontext, dass das von wissenschaftlichen medizinischen Fachgesellschaften vorgeschlagene Instrument eines Scoping-Workshops in den Ablauf integriert wurde. Es kann helfen, bei der Frühen Nutzenbewertung mehrfach vorgekommene Fehler bei der Festlegung von Komparatoren (zweckmäßige Vergleichstherapie) zu vermeiden.

Die Erstellung von Versorgungspfaden ist sinnvoll. Ihre Voraussetzung im Verfahren der Kosten-Nutzen-Bewertung wird aber wohl dazu führen, dass nur sehr selektiv Verfahren initiiert und abschließend durchgeführt werden. Gleichzeitig ist sicherzustellen, dass die medizinischen Fachgesellschaften und Patienten in angemessener Weise und aktiv an der Erstellung der Versorgungspfade beteiligt sind. Es ist nicht sinnvoll, neben Evidenz-basierten Leitlinien eine virtuelle Parallelwelt aufzubauen.

Zentrales Element einer Kosten-Nutzen-Bewertung ist die Wahl des geeigneten Modells zur Berechnung und zur Darstellung des Gesamtnutzens. Das National Institute for Care and Health Excellence (NICE) in Großbritannien hatte frühzeitig das Instrument der Quality-Adjusted-Life-Years (QALY) zu diesem Jahr eingesetzt. Daran gibt es inhaltliche Kritik. Es ist fraglich, ob das QALY-Konzept mit dem in Deutschland geltenden Grundverständnis von Sozialstaat und Gesundheitswesen kompatibel ist, insbesondere in der normativen Festlegung eines finanziellen Durchschnittswertes ohne Berücksichtigung individuell unterschiedlicher Krankheitsbelastungen. Das IQWiG nennt Weiterentwicklungen und alternative Modelle, legt sich aber nicht auf ein Modell fest. Diese Restriktion des IQWiG auf Berechnungen mit ggf. unterschiedlichen Modellen ist sinnvoll und kann dem G-BA eine ausreichend breite Basis für Bewertungen und Entscheidungen geben. Eigene Kriterien des IQWiG für die Auswahl des jeweils bestgeeignet erscheinenden Modells fehlen im Methodenpapier Version 4.2.

Auch in der Kosten-Nutzen-Bewertung wird medizinische Expertise primär auf Basis eines Fragenkatalogs eingeholt. Das hat nach unserer Einschätzung nicht zur erforderlichen Qualitätssteigerung der Berichte mit dem Ziel der Abbildung oder der Annäherung an den aktuellen Stand des Wissens geführt. Hier sind offene Diskussionen erforderlich.

5. Gesundheitsinformationen (Kapitel 6)

In der Onkologie gibt es diesbezüglich zahlreiche Angebote. Anbieter reichen von Krankenkassen, regional oder national operierenden Krebsberatungsstellen und Krebsinformationsdiensten oder einem neu im Aufbau befindlichen Portal von Deutscher Krebsgesellschaft/Deutscher Krebshilfe, über Selbsthilfegruppen und Verbraucherschützer bis zu den wissenschaftlichen medizinischen Fachgesellschaften. Die Quantität ist hoch, die Qualität ist inzwischen oft ebenfalls, für einige Krankheitsentitäten sogar herausragend. Ressourcen sollten verantwortungsbewusst, d. h. bei den Gesundheitsinformationen kooperativ und nicht kompetitiv eingesetzt werden.

Als Kriterium für die Erstellung von Gesundheitsinformation hat das IQWiG eine Prävalenz von 1% gesetzt. Solche Zahlen sind heute fragwürdig. Rasch zunehmende Erkenntnisse der biomedizinischen Forschung haben zu einer Differenzierung vormals homogen erscheinender Krankheitsbilder in biologisch und klinisch distinkte Subgruppen geführt. Diese unterscheiden sich zunehmend sowohl in Diagnose als auch in Therapie und Management unterschiedlich. Die hämatologischen und onkologischen Erkrankungen sind Vorreiter in diesem Bereich. Wir halten offenere Kriterien für die Auswahl von Themen für Gesundheitsinformationen für angemessen.

Ein weiterer kritischer Punkt von Gesundheitsinformationen ist die Aktualisierung. Im Methodenpapier sind jetzt drei Jahre vorgesehen. Diese Vorgabe basiert auf einer Erhebung zu Änderungen im Stand des medizinischen Wissens aus den 90er Jahren des letzten Jahrhunderts. In der Hämatologie, Onkologie und der Hämostaseologie sind die Änderungen im Stand des Wissens nicht linear, sondern wellenförmig. Wenn das Konzept der Gesundheitsinformationen weiter verfolgt wird, erscheint eine variabelere Aktualisierung erforderlich.

6. Was fehlt im Methodenpapier 4. 2.?

Die Erfahrungen mit den IQWiG Berichten im Rahmen der Frühen Nutzenbewertung haben eine Reihe von Defiziten deutlich gemacht. Diese wurden in der Vergangenheit schon kritisiert, inzwischen z. T. auch in direkten Gesprächen diskutiert. Weiterhin bestehende Defizite betreffen:

- fehlende Validierung der Operationalisierungskriterien (Tabellen 16 und 17);
- fehlende Publikation des zugrundeliegenden Algorithmus in der verwendeten Software SAS;
- ungenügende Berücksichtigung anderer Endpunkte als Gesamtüberleben und fehlende Methodik zur Hierarchisierung von mehreren Endpunkten. Die jetzt im Methodenpapier bei der Kosten-Nutzen-Bewertung eingeführten Instrumente zur multikriteriellen Entscheidungsfindung oder Präferenzhebung könnten hier weiterentwickelt werden;
- ungenügende Methodik zur Berücksichtigung der Postprogressionstherapie einschl. des Cross-over in randomisierten Studien;
- flexiblerer Umgang mit Ergebnissen des Patient-Reported-Outcome;
- „Subgrupptis“ – Berechnung des Zusatznutzens für Subgruppen von Patienten auf der Basis von Ergebnissen der Verzerrungsanalyse, ohne dass die Power der jeweiligen Studien ausreichend zur Detektion von Unterschieden in diesen Subgruppen ist;
- unvollständige Beteiligung betroffener Patienten (nicht bei allen Berichten sind Patienten beteiligt);
- ungenügende Einbindung von Fachexperten für die jeweilige Krankheitsentität (niedrige Rate an Experten in leitenden Funktionen, niedrige Rate an Experten aus aktuellen Leitlinienprojekten der wissenschaftlichen medizinischen Fachgesellschaften).

7. Literatur (nach Erstellung der Konsensversion)

1. Karl-Anton Kreuzer, Peter Bettelheim, Andreas Rosenwald, Torsten Haferlach: Hämatologische Diagnostik. Leitlinien von DGHO, OeGHO, SGMO und SGH+SSH, Status August 2014. <https://www.dgho-onkopedia.de/de/onkopedia/leitlinien/haematologische-diagnostik>
2. Wörmann B et al.: Nutzenbewertung von Arzneimitteln der Onkologie und Hämatologie, 2013. http://www.dgho.de/informationen/gesundheitspolitische-schriftenreihe/dgho_gpsr_arzneimittelnutzenbewertung.pdf
3. Bokemeyer C, Van Cutsem E, Rougier P et al.: Addition of cetuximab to chemotherapy as first-line treatment for KRAS wild-type metastatic colorectal cancer: pooled analysis of the CRYSTAL and OPUS randomized clinical trials. Eur J Cancer 48:1466-1475, 2012. DOI: [10.1016/j.ejca.2012.02.057](https://doi.org/10.1016/j.ejca.2012.02.057)
4. Douillard JY, Oliner KS, Siena S et al.: Panitumumab-FOLFOX4 treatment and RAS mutations in colorectal cancer. N Engl J Med 369:1023-1034, 2013. DOI: [10.1056/NEJMoa1305275](https://doi.org/10.1056/NEJMoa1305275)

Die Stellungnahme wurde von Prof. Dr. Bernhard Wörmann in Kooperation mit Prof. Dr. Carsten Bokemeyer erarbeitet.

Mit freundlichen Grüßen

für die DGHO



Prof. Dr. med. Mathias Freund
Geschäftsführender Vorsitzender



Prof. Dr. med. Diana Lüftner
Vorsitzende



Prof. Dr. med. Martin Wilhelm
Mitglied im Vorstand - Sekretär

für die GTH



Prof. Dr. med. Christoph Bode
Präsident



Dr. med. Andreas Tiede, PhD
Mitglied im Vorstand – Sekretär

A.1.12 – Deutsche Gesellschaft für Hygiene und Mikrobiologie e. V. (DGHM)

Stellungnahme der Deutschen Gesellschaft für Hygiene und Mikrobiologie (DGHM) zum Entwurf „Allgemeine Methoden“ des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Die Arbeitsgemeinschaft Wissenschaftlich-Medizinischer Fachgesellschaften (AWMF) hat der DGHM die neue Version 4.2 der „Allgemeinen Methoden“ des IQWiG vom 18.06.2014 mit Bitte um Stellungnahme zugeschickt. –

Der Vorstand der DGHM kommt dieser Bitte gerne nach. In ihrer Bewertung fokussiert die DGHM dabei auf die Abschnitte des geänderten Dokuments, die im Vergleich zur vergangenen Version eine Änderung erfahren haben.

Aus Sicht der DGHM sind folgende Punkte für die Neufassung zu bedenken:

1. Abschnitt 3.5 Diagnostische Verfahren

Diagnostische Verfahren im Fach Medizinische Mikrobiologie, Virologie und Infektionsepidemiologie sind typischerweise durch den direkten oder indirekten Erregernachweis (letzterer durch Untersuchungen der Wirtsreaktion auf einen Erreger) sowie durch die Charakterisierung und Differenzierung des Erregers gekennzeichnet. Im Bereich der Infektionsepidemiologie kommen darüber hinaus Verfahren zum Einsatz, die u.a. als prognostischer Marker oder „Risikofaktoren“ insbesondere Eingang in Richtlinien und Empfehlungen zu Aspekten der Krankenhaushygiene finden.

Im Gegensatz zu diagnostischen Verfahren der übrigen (Human-)Medizin, welche klassischerweise die (Patho-)Physiologie, (Patho-)Biochemie, Funktionsprüfung, Bildgebung etc. menschlicher Organe und Organsysteme nachweisen, wird das medizinisch-mikrobiologische Untersuchungsverfahren daher **regelmäßig zum Nachweis und zur Bewertung eines externen, biologischen, häufig „schädigenden“ Faktors (des „Infektionserregers“)** eingesetzt. Erst aus der Interaktion dieses Faktors (Erreger, Erregerkonsortien, „Mikrobiom“) mit dem menschlichen Wirt resultiert die „Wirtsreaktion“, die dann in der Regel auch die Grundlage der „Krankheit“ ausmacht, welche dann wiederum zu einer Folgebehandlung führt oder führen kann.

In diesem Abschnitt wird durchgängig der besondere Wert eines diagnostischen/prognostischen Verfahrens dahingehend bewertet, dass diese einen gesundheitsbezogenen Nutzen aufweisen: *„Diagnostische Verfahren zeichnen sich dadurch aus, dass ihr gesundheitsbezogener Nutzen (oder Schaden) im Wesentlichen erst dadurch zustande kommt, dass sich ihnen therapeutische oder präventive Verfahren anschließen. Die alleinige Gewinnung diagnostischer Informationen (ohne therapeutische oder präventive Konsequenzen) hat regelmäßig keinen sozialrechtlich relevanten Nutzen [440].“* sowie *„Insgesamt entscheidend ist weniger, inwieweit eine diagnostische oder prognostische Information einen aktuellen oder zukünftigen Gesundheitszustand feststellen kann, sondern dass diese Information auch prädiktive Bedeutung hat, also den höheren (oder geringeren) Nutzen einer Folgebehandlung vorhersagen kann [184,498].“*

Für die Natur der medizinisch-mikrobiologischen Diagnostik, die eben nicht die Organfunktion, Bildgebung oder Pathophysiologie des menschlichen Körpers, sondern eben „nur“ den Nachweis bzw. Ausschluss des Schädigungsfaktors ermöglicht, macht dieser Umstand **die geforderte unmittelbare Auswirkung der Diagnostik auf die Folgebehandlung häufig schwierig, in vielen Fällen unmöglich.**

Hinzu kommt, dass der typische klinische Ablauf einer Infektionskrankheit dadurch charakterisiert ist, dass der Behandler bei Feststellen klinischen Zeichen und Symptomen, und regelmäßig **vor Erhalt des diagnostischen Untersuchungsergebnisses** die Behandlung (die in diesem Fall keine „Folgebehandlung“ sondern eine „kalkulierte“ oder „empirische“ therapeutische Erstmaßnahme darstellt) einsetzt. Anderes wäre auch gar nicht im Sinne des Stands der Heilkunst zu vertreten.

Die Qualitätsforderung an medizinisch-mikrobiologische Diagnostik-Verfahren, dass nämlich das Untersuchungsergebnis die nachfolgende therapeutische bzw. präventive Maßnahme nach sich zieht, ist daher für diesen Bereich **überzogen und nicht zielführend**. Die Bewertung medizinisch-mikrobiologischer Diagnostik-Verfahren muss sich daher primär an der **Zuverlässigkeit des Nachweises des zu untersuchenden Schädigungsfaktors (=des Erregers)** als Grundlage für eine dann gezielte therapeutische oder präventive Maßnahme ableiten.

Hingegen beschreibt dieser Abschnitt an mehreren Stellen die „Testgüte“ des diagnostischen Verfahrens (u.a. „Die bisherigen Ausführungen beziehen sich primär auf diagnostische Verfahren, die durch eine Erhöhung der Testgüte (also Sensitivität, Spezifität oder beides) mehr Patienten einer bestimmten therapeutischen Konsequenz zuführen.“). Hier ist jedoch anzumerken, dass „Sensitivität und Spezifität“ nur die Leistungsstärke des eigentlichen Tests bewertet. Für die medizinische Mikrobiologie ist in einer Testbewertung i.S. des gesundheitsbezogenen Nutzens jedoch die Bewertung der Prävalenz des Untersuchungsobjekts bezogen auf Population, Kohorte o.ä. entscheidend. Ein hoch, aber nicht 100% spezifischer Test ist als Screening-Verfahren nutzlos, sofern die Prävalenz des Infektionserregers extrem selten ist – auch ein solcher Test würde fälschlicherweise zahllose nicht infizierte Personen als positiv identifizieren und damit einer (nutzlosen bzw. schädigenden) Behandlung zuführen. Hier sind unbedingt weitere, populationsbezogene Merkmale wie positiver und negativer Vorhersagewert mit einzubeziehen.

2. Abschnitt 5.2.3. Strukturierte Aufbereitung und Auswertung von Empfehlungen

In diesem Abschnitt geht es u.a. um die Vereinheitlichung von Empfehlungsgraduierungssystemen. In dem (in der Version 4.2 nicht modifizierten) Text wird zutreffend festgestellt, dass „international ... eine konsentiertere Vereinheitlichung von Evidenz- und Empfehlungsgraduierungssystemen bislang nicht erfolgt ist“. Anscheinend soll auch diese neue Version für extrahierte Empfehlungen hier keine einheitliche Auswertung vorsehen. Angesichts der schwierigen Interpretierbarkeit der verschiedenen, international und bezogen auf Fächer, Fachgesellschaften, Behörden oder Institutionen unterschiedlichen Empfehlungsgraduierungssysteme wäre aus Sicht der DGHM jedoch ein einheitliches, von IQWiG verwendetes Graduierungssystem prinzipiell wünschenswert. Die Verfahrensstandards der DGHM, die sogenannten MiQ, orientieren sich dabei an dem **Empfehlungsgraduierungssystem der Amerikanischen Centers of Disease Control.**

3. Abschnitt 6.3.4 Aktualisierung

Dieser Abschnitt befasst sich mit der Aktualisierung von evidenzbasierten Gesundheitsinformationen für Bürgerinnen und Bürger. In diesem (neuen) Absatz wird ausgeführt, dass „das exponentielle Wachstum der Evidenz aber weiterhin anhält, ist davon auszugehen, dass die Halbwertszeit der Information sich generell verkürzt, d. h., dass die Informationen schneller veralten. Das Institut sieht daher üblicherweise drei Jahre als Zeitintervall für eine Aktualitätsprüfung an.“ Diese Aussage – auch wenn sie im Abschnitt „Gesundheitsinformationen“ steht, bezieht sich auf Leitlinienempfehlungen (und sollte möglicherweise auch dort diskutiert werden). Diese Zeitspanne steht in einem Missverhältnis zu der täglichen Erfahrung, u.a. der DGHM MiQ Qualitätssicherungskommission, zur Überarbeitung ihrer Verfahrensstandards. Die ebenfalls exponentiell komplexer werdende Notwendigkeit zur Literaturprüfung und Evidenzbewertung auch diagnostisch-mikrobiologischer Verfahren einerseits, der Erfordernis ausgewiesener Sachverständiger zur Erstellung solcher Verfahrensstandards (die unserer Erfahrung nach sämtlich über massive Zeitnot klagen) und der Kontext eines nach wie vor nicht oder nur marginal honorierten medizinisch-wissenschaftlichen Reputationsgewinns durch die Erstellung solcher Verfahrensstandards **macht eine dreijährige Überprüfung in aller Regel illusorisch.** Sofern IQWiG auch solcher medizinisch-mikrobiologischen Verfahrensstandards in ihre Gesundheitsinformationen bzw. Leitlinienbewertung einzubeziehen gedenkt, sind daher deutlich längere Überarbeitungsintervalle der Verfahrensstandards, die durch unsere Fachgesellschaft erstellt werden, zu berücksichtigen.

A.1.13 – Deutsche Gesellschaft für Innere Medizin e. V. (DGIM)



Generalsekretär

Prof. Dr. med. Dr. h.c. Ulrich R. Fölsch
Univ.-Klinikum Schl.-Holstein, Campus Kiel
Klinik für Innere Medizin I
Schittenhelmstraße 12
24105 Kiel



07.08.2014

Stellungnahme der Deutschen Gesellschaft für Innere Medizin (DGIM) zum Methodenpapier 4.2 des IQWiG

Potenzialbewertung (S.73 ff): Die Potenzialbewertung richtet sich offenbar auf die Einführung neuer Produkte. Das ist zu begrüßen. Soweit wir ersehen, geht aus dem Text nicht klar hervor, wie und in welcher Form medizinische Fachgesellschaften hier einbezogen werden sollen.

Patientenzufriedenheit(S. 38 ff): Warum ist in 3.1.1 der Begriff Patientenzufriedenheit gestrichen worden? Der Verweis auf die Literatur bleibt vage (S.39 u.).

Diagnostische Verfahren (3.5;S. 64) : Sind wirklich genügend Wirksamkeitsstudien für diagnostische Verfahren vorhanden? Die Verknüpfung von diagnostischer Information mit Nutzen für die Folgebehandlung (S.66) ist zwar wünschenswert und teilweise richtig, aber in der Praxis sehr schwer zu fassen und durch Studien zu belegen. Hier gehen sehr viele individuelle Faktoren des Patienten und seiner Krankheit, aber auch des diagnostizierenden sowie behandelnden Arztes ein (Ausbildung, Erfahrung etc.). Abhängig von diesen unterschiedlichen Konstellationen kann sich der Einsatz diagnostischer Verfahren sehr ändern. S. 67: Wer definiert das Verfahren, mit dem verglichen wird?

Kosten-Nutzen-Bewertung (S. 76 ff) : 4.1.3 bei der Bestimmung des Zeithorizontes (S.77) sollten die entsprechenden Fachgesellschaften einbezogen werden. 4.1.5 Die Erstellung eines Versorgungspfades bleibt unklar. Wer ist hier einbezogen? Was bedeutet Piggy-back-Studie? 4.1.7 Wie erfolgt die Auswahl der Experten zur Befragung?

Die Ausführung über die Methodik der KNB ist außerordentlich lang und nur für epidemiologisch geschulte Biometriker verständlich. Hat hier eine externe Prüfung des Textes stattgefunden? Bezüglich der Erhebung von Präferenzen zur Erstellung eines Maßes des Gesamtnutzens (4.2.5 B; S. 88 ff) ist sehr zu begrüßen, dass das IQWiG eine gewichtete Bewertung von patientenzentrierten Endpunkten methodisch implementiert. Bei der Planung des Analytic-Hierarchy-Process-Verfahrens und der Conjoint-Analyse zum Einsatz in der frühen Nutzenbewertung sollten die Fachgesellschaften eingebunden werden, insbesondere bei der Identifikation der relevanten Kriterien, Auswahl der zu Befragenden, Beurteilung der Patientenkollektive.

Gesundheitsinformationen (S.130 ff): Hier ist ein Testlesen durch die entsprechenden Zielgruppen vorgesehen. Ist geplant, auch Experten aus Fachgesellschaften beim Testlesen einzubeziehen? Wenn nein, warum nicht? Bei der Themenauswahl (S.141) sollten die entsprechenden Fachgesellschaften einbezogen werden. Es wird derzeit ein epidemiologisches Kriterium (Prävalenz der Erkrankung mindestens 1 % der Bevölkerung) herangezogen (S.142). Hier sollte ein entsprechender Katalog veröffentlicht werden, und er sollte durch die Fachgesellschaften überprüft werden können. Hinsicht-

lich der notwendigen Aktualisierung sollten auch die Vertreter der entsprechenden Fächer kontaktiert werden.

Informationsbeschaffung (S. 164 ff): Die Ablehnung der unaufgeforderten Übermittlung von Studiendaten (S. 164) ist zwar verständlich im Sinne der Vermeidung eines Informations-Bias, kann aber auch die Gefahr bergen, Informationen, die auf Patientengefährdung/-benachteiligung hinweist, nicht zu berücksichtigen. Wie wird das umgangen?

Informationsbewertung, Endpunkte (S. 183): Hier sollten unbedingt Experten aus den entsprechenden Fachgesellschaften mindestens einbezogen werden.

Allgemeiner Vorspann (S. 1 ff): Die Definition der Kriterien (S. 2) für die Kosten-Nutzen-Bewertung (Verbesserung Gesundheitszustand/Krankheitsdauer(Pleonasmus!?), Lebensdauer, Nebenwirkungen, Lebensqualität) ist schwierig und komplex, aber wie findet hier die Gewichtung statt, ist das ausreichend ausgeführt? Wer genau ist antragsberechtigt im Rahmen der Potenzialbewertung (S.3, siehe auch oben)? S. 8 Punkt 3 (Fehlen von ausreichenden qualitativen Studien): Wer wird zur Definition der „gegenwärtig besten Evidenz“ heran gezogen?

A.1.14 – Deutsche Gesellschaft für Orthopädie und Unfallchirurgie e. V. (DGOU)



DGOU Deutsche Gesellschaft für
Orthopädie und Unfallchirurgie

**Angeforderte Stellungnahme der Deutschen Gesellschaft für Orthopädie
und Unfallchirurgie (DGOU) e.V.**

**Methodenkritik
Entwurf für Version 4.2 vom 18.06.2014
„Allgemeine Methoden“
Institut für Qualität und Wirtschaftlichkeit im
Gesundheitswesen (IQWiG)**

Auftraggeber

Deutsche Gesellschaft für Orthopädie und Unfallchirurgie (DGOU) e.V.

Vertreten durch den Generalsekretär
Prof. Dr. med. Fritz-Uwe Niethard
per E-Mail am 11.07.2014

Auftragnehmer

Methodisches Beratungszentrum der Deutschen Gesellschaft für
Unfallchirurgie (DGO) e.V.

Vertreten durch
Priv.-Doz. Dr. med. Dirk Stengel, MSc(Epi)
Ärztlicher Leiter des Zentrums für Klinische Forschung,
Klinik für Unfallchirurgie und Orthopädie, Unfallkrankenhaus Berlin, Warener
Str. 7, 12683 Berlin,

Inhalt

1. Position der DGOU	3
2. Gegenstand des Kommentars.....	4
2.1. Allgemeines	4
2.2. Ergebnissicherheit	5
2.3. Fristen	5
2.4. Diagnostische Verfahren	7
2.5. Potenzialbewertung	9
2.6. Kosten-Nutzen-Bewertung.....	10
2.7. „Rangordnung verschiedener Studienarten / Evidenzgrade“	11
3. Empfehlungen der DGOU	12
4. Literatur	13

1. Position der DGOU

Die DGOU steht hinter einer transparenten und objektiven Nutzenbewertung im Einklang mit den Empfehlungen des G-BA und IQWiG

Das Methodenpapier des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) bildet (in allgemein verständlicher Auslegung der gesetzlichen Vorgaben) die Grundlage für die Bewertung des

- i. *Nutzens* etablierter und auf dem deutschen Markt zugelassener Diagnostika, Pharmaka, Medizinprodukte und sonstiger medizinischer Interventionen, welche der (frühzeitigen) Erkennung von Gesundheitsstörungen und deren Behandlung im Interesse der Betroffenen, der Behandler, aber auch der Solidargemeinschaft dienen sollen
- ii. *Zusatznutzens*, welcher sich aus einem (offensichtlich) günstigen Verhältnis zwischen Wirkung und Wirksamkeit, Sicherheit und Kosten neuer im Vergleich zu etablierten Methoden ergeben kann.

Nutzen und Zusatznutzen sind schwer zu trennen. Legt man den von Prof. Dr. Jürgen Windeler vorgeschlagenen Minimalkonsens zugrunde, ergibt sich der medizinische Nutzen einer Gesundheitsmaßnahme aus deren i. medizinischem, ii. positivem und iii. kausalem Effekt (i.S. einer günstigen Beeinflussung des Krankheitsverlaufes) in mehr als geringfügigem Ausmaß. Einfach kann man auch von Netto- und Brutto-Nutzen sprechen, je nachdem, ob unerwünschte Wirkungen in die Kalkulation einbezogen werden oder nicht.

Der Nutzenbegriff wird von Konsumenten, Anbietern und Kostenträgern unterschiedlich definiert. Hinzu kommen verschiedene klinische und methodische Perspektiven und Sozialisationen, unterschiedliche Definitionen von Relevanz, differierende Annahmen über akzeptable Belastungen des Gesundheitsbudgets, methodische Ansätze und neue methodische Strömungen, Umwälzungen im Sozial- und Gesellschaftsgefüge usw. Jeder Methodenvorschlag muss und sollte daher Debatten anregen, ohne dass als Endprodukt ein für alle Beteiligten akzeptables Gesamtwerk erwartet werden darf.

Die Deutsche Gesellschaft für Orthopädie und Unfallchirurgie (DGOU) e.V. geht mit den o.g. Definitionen des Nutzens bzw. Zusatznutzens medizinischer Maßnahmen grundsätzlich konform.

Die DGOU ist die Dachorganisation und berufspolitische Vertretung von Ärztinnen und Ärzten, welche sich der Prävention, nicht-operativen und operativen Therapie, Rehabilitation und Erforschung von Erkrankungen und Verletzungen des Bewegungsapparates verschrieben haben. Die DGOU sieht sich zunehmend mit Decken- und Bodeneffekten und den daraus resultierenden Schwierigkeiten in der Empfehlung bestimmter Diagnose- und / oder Therapieschritte (z.B. im Rahmen Evidenz- und Konsens-basierter Leitlinien) konfrontiert.

Typische Beispiele (willkürlich ausgewählt aus einer großen Bandbreite potentieller Innovationen) sind:

- i. Der (Zusatz-)Nutzen der minimal-invasiven und / oder Computer-assistierten (navigierten) Endoprothetik des Knie- oder Hüftgelenkes im Vergleich zum konventionellen Gelenkersatz
- ii. Der (Zusatz-)Nutzen einer bioaktiven Beschichtung extra- und intra-medullärer Implantate zur Stabilisierung diaphysärer Frakturen der langen Röhrenknochen im Vergleich zu unbehandelten Standardimplantaten

Die kritische Hinterfragung potenzieller und scheinbarer Innovationen unter dem Nutzenschirm, aber auch die Auseinandersetzung mit modernen Methoden der Nutzenbewertung gehören damit zu den vornehmlichen Aufgaben der DGOU.

2. Gegenstand des Kommentars

2.1. Allgemeines

Auf S. 3 (ii) des Methodenpapiers wird (in fetter Schrift) betont, dass „nur die (Anm.: mittels gelber Markierung bzw. Textlöschung) als Änderung gekennzeichneten Textteile Gegenstand des Stellungnahmeverfahrens sind.“ Obwohl dies auf einer rechtlichen Grundlage beruht und die Arbeit der methodischen Kommentatoren grundsätzlich erleichtern könnte, verbietet es gleichsam, antiquierten und überflüssigen Ballast (s. insb. die Ausführungen 1.2 Evidenzbasierte Medizin [EbM]) aus dem Methodenpapier zu entfernen. Dieser Kommentar wiederum bezieht sich ausschließlich auf *substanzielle* Änderungen und Ergänzungen im Methodenpapier.

Als solche werden definiert

- Streichungen, Änderungen, Ergänzungen oder der komplette Austausch von Passagen oder Kapiteln über mehrere Zeilen, Absätze oder Seiten
- Inhaltliche Änderungen, welche klar zu einer neuen oder andersartigen Interpretation von Tatsachen und / oder Konsequenzen im Vergleich zum vorherigen Report führen.

Nicht kommentiert werden Wort- oder einzelne Satzergänzungen, sofern sich hierdurch keine grundlegenden Änderungen der Aussagen ergeben.

Die DGOU akzeptiert ausschließlich durch wissenschaftliche Daten begründete Diagnose- und Therapiemaßnahmen und setzt damit ein klares Zeichen gegen Beliebigkeit und Willkür in der medizinischen Versorgung der Bevölkerung.

Gleichzeitig distanziert sich die DGOU von den apodiktischen, mittlerweile als veraltet und praxisfern geltenden Vorgaben der EbM, wie sie noch im kanadischen und britischen Raum, aber auch durch das deutsche Netzwerk EbM (DNEbM) gehandhabt, propagiert und gelehrt werden.

Die DGOU unterstützt und fördert hingegen Initiativen und Aktivitäten, welche dem Praxis- und damit Patienten-zentrierten Wissens- und Erkenntnisgewinn dienen. Dies umfasst explizit auch innovative biometrische bzw. biostatistische Ansätze, um das bisherige Arsenal klinisch-epidemiologischer Forschungsmethoden zu bereichern.

2.2. Ergebnissicherheit

Die kurze Ergänzung in 1.2.5 (Anforderungen an die Präzision und Verlässlichkeit), Punkt 4 (S. 8), dass der G-BA die „Möglichkeit hat, selbst Studien zu fordern oder zu initiieren, um die Evidenzlücken zu schließen“ ist von weitreichender Bedeutung. Wenn nämlich der G-BA Studien initiiert, muss er auch Sorge für deren Finanzierung tragen. Tatsächlich dürfen nicht i.S. einer selbsterfüllenden Prophezeiung klinische Studien gefordert werden (von denen es nie genug geben kann), ohne nicht auch deren finanziellen und logistischen Unterbau zu garantieren. *Die DGOU stützt daher diese kleine, aber wesentliche Ergänzung in jeder Hinsicht, erwartet aber konkrete und nachhaltige Entscheidungen über eine optimierte Finanzierung klinischer Studien aus öffentlicher Hand.*

Sowohl die Kosten-Nutzen-Bewertung nach §35b SGB V als auch die Potenzialbewertung sind gesetzlich vorgeschrieben und daher seitens der Fachgesellschaft kaum zu beeinflussen. Die Ergänzungen auf den Seiten 12 – 18 sind unkritisch und selbsterklärend.

In den in Abb. 4 (S. 23) illustrierten Ablaufplan der Kosten-Nutzen-Bewertung sollten speziell ausgewiesene und benannte Vertreter der betroffenen Fachgesellschaften integriert werden.

Der Passus in 2.1.4 (S. 22) „... Dabei werden regelhaft externe Sachverständige beteiligt ...“ ist zu vage. Erfahrungen der Vergangenheit zeigen, dass die Kriterien für die Auswahl der Sachverständigen und deren Qualifikation unbestimmt waren. Die DGOU fordert daher, im Falle einer Kosten-Nutzen-Bewertung nach §35b SGB V mindestens *zwei unabhängige, von einer Fachgesellschaft akkreditierte Sachverständige* in den Arbeitsprozess zu integrieren. Diese Forderung ergibt sich aus dem exponentiellen Wissenszuwachs und den Spezialisierungstendenzen innerhalb der verschiedenen Fachdisziplinen.

Der auf den Seiten 24 – 26 geschilderte Prozess, Auftragsinhalte zu konkretisieren und mit verschiedenen Stakeholdern abzustimmen, ist grundsätzlich klar.

Die Rolle der Vertreter der Fachgesellschaften sollte aber noch deutlicher herausgearbeitet werden. Eine allgemeine medizinische Expertise ist nicht mehr geeignet, um Empfehlungen des IQWiG und / oder des G-BA mit Einfluss auf die Versorgungslage zu beeinflussen.

Die Sachverständigendatenbank sollte in Abstimmung mit den Fachgesellschaften aktualisiert, angepasst und ggf. bereinigt werden.

2.3. Fristen

Die dreiwöchige Frist bis zur Abgabe einer Stellungnahme zum *Vorbericht* durch die „Öffentlichkeit“ (sind hiermit Konsumenten, Anbieter oder Kostenträger gemeint?) ist und bleibt indiskutabel. Aufgrund des üblichen Volumens der Vorberichte bei gleichzeitig nachhaltiger Bedeutsamkeit für die medizinische Versorgung muss aus Sicht der DGOU die Frist auf wenigstens sechs Wochen ausgedehnt werden, um adäquate Kommentare und Änderungsvorschläge zu unterbreiten.

Die auf S. 24 geschilderten Fristen sind etwas unklar. Der Öffentlichkeit soll eine Frist „von drei Wochen zur ... Stellungnahme (Anhörung) [zum Vorbericht] gegeben werden.“ Der Abschlussbericht wiederum soll dem G-BA „innerhalb von drei Monaten nach Einleitung des Stellungnahmeverfahrens zum Vorbericht“ übermittelt und muss innerhalb dieses Intervalls auch kommentiert werden.

Die DGOU empfiehlt dem IQWiG, ein zeitgenaues Ablaufschema der Einreichungs- und Kommentierungsintervalle zu entwerfen, um industriellen und nicht-industriellen Antragsstellern die notwendigen Zeithorizonte zu verdeutlichen.

Das Verfahren zur Einreichung des Abschlussberichtes (2.1.4) klammert die Fachgesellschaften bisher aus. Sollte dieses Verfahren gesetzeskonform sein, stimmt die DGOU dem Vorgehen zu. Eine unabhängige juristische Prüfung erscheint jedoch aufgrund der erheblichen Konsequenzen des Abschlussberichtes als Grundlage von und für Entscheidungen des G-BA zwingend notwendig.

Sollte es im Ablaufschema zwischen i. erster Vorberichtsphase, ii. überarbeitetem Vorbericht dank Stellungnahme der Fachgesellschaften, iii. erster Abschlussberichtsfassung und iv. Entscheidungsfindung des G-BA, des Stiftungsvorstandes, -rates und / oder -kuratoriums zu Verzögerungen oder diskrepanten Stellungnahmen kommen, muss erneut ein Vertreter / eine Vertreterin der jeweiligen Fachgesellschaft als Moderator/in, Schlichter/in oder Ombudsmann / Ombudsfrau benannt werden, um den Konflikt zu lösen.

Die institutseigene Sachverständigendatenbank sollte mit der Fachgesellschaft abgestimmt werden. Zu benennen sind vorzugsweise Sachverständige, welche sich durch aktuelle klinische oder sonstige Expertise auf dem entsprechenden Fachgebiet auszeichnen.

Die in 2.1.5 geschilderte Potenzialbewertung kommt im Vergleich zu anderen Passagen des Methodenpapiers etwas kurz. Sofern der interne Qualitätssicherungsprozess dargestellt werden darf, sollte er in das überarbeitete Dokument aufgenommen werden.

Die auf S. 39 genannten Zielgrößen Mortalität, Morbidität und gesundheitsbezogene Lebensqualität sind im SGB V verankert und werden durch die DGOU vertreten. Dennoch sind bei Verletzungen und Erkrankungen des Bewegungsapparates Surrogate wie radiologische Parameter (z.B. knöcherne Konsolidierung auf Röntgen- oder CT-Bildern) von erheblicher Relevanz. Die konditionale Wahrscheinlichkeit, dass z.B. im Falle einer inadäquaten Reposition von Gelenk-bildenden Fragmenten nach intra-artikulären Frakturen oder einer verzögerten bzw. ausbleibenden Knochenbruchheilung gute funktionelle Ergebnisse oder eine Wiederherstellung der gesundheitsbezogenen Lebensqualität auf das Ausgangsniveau erreicht werden können, ist gering.

Die Art der Gewichtung verschiedener patientenrelevanter Endpunkte (S. 49) ist unklar. Diese kann quantitativ, semi-quantitativ oder qualitativ sein, muss aber spezifiziert werden. Natürlich sind Nutzen und Schaden themenspezifisch zu definieren. Dennoch sollten die Gewichtungsprozeduren besser illustriert werden.

2.4. Diagnostische Verfahren

Die ab S. 64 ff unter 3.5 genannten Aspekte zu diagnostischen Verfahren sind inkonsistent und müssen überarbeitet werden.

Es sollte eine klare Trennung zwischen Efficacy (Accuracy), Effectiveness (Beeinflussung der ärztlichen Entscheidungsfindung) und Efficiency (günstige Beeinflussung des Krankheitsverlaufes) gemäß der Fryback-Thornbury-Hierarchie¹ erfolgen.

Die Aussage, dass „... die alleinige Gewinnung diagnostischer Informationen (ohne therapeutische oder präventive Konsequenzen) keinen sozialrechtlich relevanten Nutzen“ besitzt, ist schlichtweg *falsch*.

Ohne Kenntnisse über die diagnostische Genauigkeit eines Tests (d.h., ob der Test zwischen erkrankten und nicht-erkrankten Populationen unterscheiden kann) können keine Effectiveness- oder Efficiency-Studien geplant werden.

Für die meisten diagnostischen Tests liegen keine robusten Genauigkeitsinformationen vor. So sind z.B. Analysen des Traumaregisters, welche einen Überlebensvorteil von schwerverletzten Patienten durch die primäre Ganzkörper-CT suggerieren^{2, 3}, nur dann sinnvoll beurteilbar, wenn Informationen über die diagnostischen Genauigkeit der Methode⁴ und deren Konsequenz für ärztliche Entscheidungen berücksichtigt werden.

Die Gleichsetzung prognostischer und diagnostischer Indikatoren auf S. 64 des Papiers steht im eklatanten Widerspruch zu den Prinzipien des stellvertretenden Institutsdirektors, Herrn PD Dr. med. Stefan Lange, der zu Recht eine klare Abgrenzung zwischen diesen Informationen fordert. Sie steht allerdings auch im Widerspruch zu späteren differenzierten Äußerungen und muss dringend angepasst werden.

Im Folgenden sollen lediglich diagnostische, nicht prognostische Studien adressiert werden. *Randomisierte Studien stehen auf der finalen Stufe der diagnostischen Evaluationskaskade.* Sie sind *nur dann zulässig*, wenn sich der experimentelle Test im Vergleich zu einem akzeptierten oder offensichtlichen Referenzstandard als ausreichend genau erwiesen hat. Die Grenze, ab wann klassische Messgrößen wie Sensitivität, Spezifität, PPV, NPV, DOR, LR oder ROC des experimentellen Tests dem bisherigen Referenztest nicht-unterlegen oder äquivalent anzusehen sind, müsste durch das IQWiG festgelegt und benannt werden.

Die Aussage, dass Studien mit einer „idealerweise zufälligen Zuteilung von Patientinnen und Patienten zu einer Strategie mit bzw. ohne Anwendung der zu prüfenden diagnostischen Maßnahme oder zu Gruppen mit bzw. ohne Offenlegung der (diagnostischen) Testergebnisse“ der höchste „Evidenzgrad“ zugesprochen wird, wird durch die DGOU aus folgenden Gründen abgelehnt:

- i. Wie oben ausgeführt, erfordert eine randomisierte diagnostische Studie primär den Nachweis der diagnostischen Genauigkeit des interessierenden Tests. Dies ist häufig nicht gegeben. Ein inakkuratere Test kann und darf aber keine therapeutischen Konsequenzen nach sich ziehen. Zeigt also ein

- hypothetisches RCT den Vorteil eines individuellen Tests oder einer diagnostischen Strategie im Vergleich zu einer Kontrolle, ohne dass verlässliche Daten zu deren Genauigkeit vorliegen oder evtl. im Rahmen des RCT erhoben wurden, sind die Ergebnisse nicht zu interpretieren.
- ii. Die Protokoll-konforme Nicht-Offenlegung von Testergebnissen kann insb. im Fall von invasiven oder Strahlen-exponierenden Methoden zu ethisch grenzwertigen Situationen führen. Ein konkretes und aktuelles Beispiel ist das EF3X Trial.⁵ Dieses soll den Nutzen einer intra-operativen 3D-Bildgebung für die chirurgische Therapie intra-artikulärer Frakturen nachweisen. In der Kontrollgruppe werden Patienten einem 3D-Röntgen-Scan unterzogen - die Bilder werden den Operateuren jedoch nicht gezeigt. Obwohl die Studie formal einem „hohen Evidenzgrad“ genügt, sind die ethischen Probleme offensichtlich. Die Kontrollgruppe unterläuft eine Testprozedur mit methodisch unvermeidbarer additiver Strahlenexposition. Ihr wird dabei jeder mögliche Gewinn, der aus den Testergebnissen resultieren könnte (insb. der unmittelbare intra-operative Nachweis einer unzureichenden Fragmentreposition oder suboptimalen Implantatpositionierung), bewusst vorenthalten. Das IQWiG kann und darf derartigen Studien nicht Vorschub leisten.
 - iii. Sollten die o.g. Annahmen die Verfahrensordnung des G-BA repräsentieren, kann die DGOU nur dringend eine kritische Reflexion des bisherigen Verfahrens und eine Neubewertung anregen.

Die Aussage, dass „derartige Studien ... mit vergleichsweise moderatem Aufwand durchführbar sind“ (S. 65), ist jenseits der Realität.

Tatsächlich gehört die schrittweise und hierarchische Evaluation des Nutzens diagnostischer Tests (insb. im Falle fehlender definitiver Referenzstandards) zu den kompliziertesten und aufwändigsten Unterfangen der klinischen und der Versorgungsforschung.

Auf S. 66 werden erneut prognostische und diagnostische Indikatoren entgegen der IQWiG-Direktive vermengt bzw. unklar voneinander abgegrenzt.

Die Aussage „Insgesamt entscheidend ist weniger, inwieweit eine diagnostische oder prognostische Information einen aktuellen oder zukünftigen Gesundheitszustand feststellen kann, sondern dass diese Information auch prädiktive Bedeutung hat, also den höheren (oder geringeren) Nutzen einer Folgebehandlung vorhersagen kann“ ist inkorrekt. Dies wurde bereits detailliert im obigen Absatz begründet.

Einfach gesagt ist die diagnostische Information über den aktuellen Gesundheitszustand die unabdingbare Voraussetzung für alle weiteren medizinischen Maßnahmen.

Logisch fehlerhaft ist die Aussage, dass „ein Nutzen insgesamt sich in der Regel nur dann ergeben kann, wenn beide Interventionen ihr Ziel erfüllen: Sowohl bei unzureichender prädiktiver Diskriminationsfähigkeit der diagnostischen Intervention als auch bei fehlender Effektivität der therapeutischen Intervention wird eine Studie keinen Nutzen der diagnostischen Intervention zeigen können.“

Der Nutzen eines Tests kann sich nur dann ergeben, wenn dessen Diskriminationsfähigkeit nachgewiesen wurde. Es macht keinerlei Sinn, eine vergleichende Studie mit einem ungenauen Test durchzuführen. Auch der Nachweis

eines besseren Outcomes durch Anwendung eines experimentellen Tests ist bedeutungslos, sofern nicht vorher dessen Genauigkeit evaluiert wurde.

Unzutreffend ist auch die Aussage, dass „wenn der neue Test direkte patientenrelevante Vorteile aufweist, wie zum Beispiel geringere Invasivität oder Strahlungsfreiheit, es nicht immer notwendig sein (wird), erneut die gesamte diagnostisch-therapeutische Kette zu untersuchen, weil die therapeutischen Konsequenzen, die sich aus dem neuen Test ergeben, sich nicht von denen des bisherigen Tests unterscheiden.“

Gerade dann, wenn der neue Test mit einer geringeren Rate unerwünschter Ereignisse assoziiert ist, muss er beweisen, dass er genauso akkurat ist wie sein direkter Vergleichspartner.

Die auf S. 66 als „Anreicherungsdesign“ beschriebene Methode repräsentiert die Philosophie der individualisierten, personalisierten oder stratifizierten Medizin. Ihr sollte entweder ein eigenes Kapitel oder sogar ein separates Methodenpapier gewidmet werden. In der aktuellen Form und Fassung sind die Ausführungen schnell angreifbar. Beispielsweise wird das Problem eines *partial verification bias*, welches dann entsteht, wenn lediglich positive (nicht jedoch negative) Index-Test-Ergebnisse mittels eines akzeptierten Referenztests bestätigt werden, nur unzureichend und auch nicht im Hinblick auf therapeutische Konsequenzen adressiert.

Viele Referenztests sind entweder invasiv (z.B. Laparoskopie, Gewinnung biologischer Proben) oder exponieren Patient(inn)en gegenüber Röntgenstrahlung (z.B. CT). Die geforderten „Studien mit zufälliger Zuordnung der Reihenfolge der (voneinander unabhängigen und möglichst verblindeten) Testdurchführung bei denselben Patientinnen und Patienten“ wäre in diesem Falle aus ethischen Gründen unmöglich. Nach den Empfehlungen des IQWiG würden Patient(inn)en z.B. sowohl eine Röntgen- UND CT-Diagnostik, eine Normal- UND Niedrig-Dosis-CT oder eine CT UND eine invasive (operative) Diagnostik durchlaufen. Werden Patient(inn)en hingegen per Zufall der einen oder anderen Teststrategie (im Vergleich zu einem Referenzstandard?) zugeordnet, ist ein direkter Vergleich unmöglich. Das 1. Brückenprinzip nach Albert lautet „Sollen impliziert Können.“⁶ Das IQWiG fordert hier theoretisch ideale Studiendesigns, welche praktisch schlichtweg nicht umsetzbar sind.

Auf S. 69 wird erstmals eine Trennung zwischen prognostischen und prädiktiven Wertigkeiten vorgenommen.

Aufgrund der o.g. nicht überzeugenden Darlegungen und der speziellen Probleme und Methoden empfiehlt die DGOU, für diagnostische Verfahren ein separates Dossier zu verfassen.

2.5. Potenzialbewertung

Der wesentliche Aspekt der Potenzialbewertung (3.8.) erscheint etwas abrupt nach den Themenfeldern Früherkennung und Screening (3.6.) und Prävention (3.7.). Insgesamt lässt diese Gliederung Tendenzen erkennen, dass der aktuelle Methodenreport unter Zeitdruck und ohne die erforderliche Tiefe erstellt wurde.

Leider offenbaren sich ab S. 73 alle Konflikte, welche sich aus den bisherigen Methoden einer externen Leistungsbewertung von medizinischen Interventionen mit Allgemeingültigkeitsanspruch ergeben.

Auf S. 74 heißt es:

Aufgrund der besonderen Zielsetzung sind bei der Potenzialbewertung im Vergleich zur Nutzenbewertung deutlich niedrigere Anforderungen an die Evidenz zu stellen. Letztendlich besteht erst das Ziel der Erprobung darin, eine adäquate Datengrundlage für eine zukünftige Nutzenbewertung herzustellen. Dementsprechend lässt sich ein Potenzial insbesondere auch auf der Basis nicht randomisierter Studien begründen.

Es handelt sich hierbei um eine fundamentale Aussage. Es ist völlig unerheblich, zu welchem Zweck eine klinische Studie durchgeführt wird – keine einzige klinische Studie kann aus biostatistischen Gründen beweisführend sein.⁷ Jede klinische Studie trägt zur Annahme über die Wirksamkeit, Nicht-Wirksamkeit oder den Schaden einer bestimmten Intervention im Vergleich zu einer Kontrolle bei. Wenn nunmehr gemäß IQWiG-Empfehlungen ein *Potenzial* aufgrund einer nicht-randomisierten Studie begründet werden kann, gilt dies aus einfachen logischen Gründen auch für den möglichen *Nutzen* einer Intervention. *Potenzial und Wirksamkeit sind in der Erkenntnistheorie nicht zu trennen.*

Die Definitionen von „geringer“, „sehr geringer“ und „minimaler“ Ergebnissicherheit nicht randomisierter Studien sind aus methodischer Sicht inakzeptabel - auch die Begründungen für diese Kategorisierung sind nicht überzeugend. Der letzte Absatz auf S. 75 springt ohne Überleitung von einer therapeutischen zu einer diagnostischen Fragestellung.

Aus methodischer Sicht hinterlassen die o.g. Punkte allenfalls Ratlosigkeit.

2.6. Kosten-Nutzen-Bewertung

Das Kosten-Nutzen-Modell des IQWiG wurde bereits in der Vergangenheit kritisiert. Insb. der Terminus „Effizienzgrenze“ konnte den oberen Gremien der DGOU bisher nicht verständlich erläutert werden.

Leider beruht auch der aktuelle Bericht vornehmlich auf qualitativen, bestenfalls semi-quantitativen (S. 76 – 112) Aussagen, welche auch einem Lehrbuch der Gesundheitsökonomie entnommen werden können.

Eine einfache, beantwortbare Frage wäre

„Wie viel darf eine bestimmte (operative / nicht-operative / rehabilitative / sonstige) Maßnahme in Orthopädie und Unfallchirurgie pro QALY / DALY kosten, um in der Bundesrepublik als effektiv zu gelten?“

Diese Frage wird aus Sicht der DGOU im IQWiG-Bericht nicht ausreichend adressiert.

2.7. „Rangordnung verschiedener Studienarten / Evidenzgrade“

Das IQWiG ist unverändert den „Levels of Evidence“ des Centre for Evidence-Based Medicine in Oxford / UK verhaftet (<http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/>). Spätestens seit dem berühmten Zitat von Bruce Charlton “The idea of a hierarchy of methods is amazing nonsense ... the validity of a piece of science is not determined by its method”⁸ hätte jedoch ein Umdenkprozess einsetzen müssen, der sich in der klinisch-methodischen Wissenschaftsgemeinschaft langsam durchsetzt. Eine immer noch stattfindende Hierarchisierung von Studiendesigns lehnt die DGOU ab und unterstreicht die Bedeutung der *methodischen Qualität* einzelner Untersuchungen zur Beantwortung einer *a priori* definierten wissenschaftlichen Frage. Einfach gesagt kann die klinische Bedeutsamkeit einer exzellent geplanten, durchgeführten und berichteten Fallserie, Kohorten- oder Fall-Kontroll-Studie diejenige einer schlecht geplanten, durchgeführten und berichteten randomisierten kontrollierten Studie überbieten. Es ist daher müßig, sich über Studiendesigns und deren „Hierarchie“ in der klinisch-wissenschaftlichen Erkenntniskaskade“ zu verständigen, solange nicht Einigkeit über die methodische Qualität besteht.

Auf S. 174 wird betont, dass „sich diese [qualitative] Graduierung auf das Verzerrungspotenzial (siehe Abschnitt 8.1.4) und nicht auf den Evidenzgrad einer Studie bezieht.“ Somit besteht Einverständnis zwischen IQWiG und DGOU. Sollten jedoch zukünftig weiterhin Begriffe wie „Evidenzklassifizierungen“ oder „Evidenzhierarchien“ in Empfehlungen und / oder Vorgaben des IQWiG bzw. des GBA auftauchen, wird die DGOU auf die Notwendigkeit einer einheitlichen und konsentierten Terminologie hinweisen.

3. Empfehlungen der DGOU

In der Zusammenschau zieht die DGOU folgende Rückschlüsse und Konsequenzen aus der aktuellen Version des Methodenreports des IQWiG:

1. Insg. wirkt der aktualisierte Report nicht „aus einem Guss“. Es gibt deutliche Diskrepanzen in den verschiedenen Aussagen, Interpretationen und Forderungen.
2. Letztere resultieren möglicherweise aus der sensiblen Schnittstelle des IQWiG zwischen der übergeordneten Instanz des G-BA, den Kostenträgern, Fachgesellschaften, Leistungserbringern und Konsumenten.
3. Ein „one-fits-all“ Methodenbericht ist nicht mehr zeitgemäß. Es müssen separate Berichte für diagnostische, therapeutische, gesundheitsökonomische und sonstige Evaluationen erstellt werden – ohne das Zielpublikum durch schiereres Volumen abzuschrecken. Methodisch belastbare, inhaltlich verständliche, politisch umsetzbare und dabei rechtskräftige Dossiers können und sollten auf weniger als 30 Seiten abgehandelt werden.
4. Viele der genannten methodischen Standards sind entweder veraltet, vage, offensichtlich kompromissbemüht oder klinisch nicht umsetzbar. Die DGOU könnte besser mit klaren quantitativen Vorgaben oder Empfehlungen leben als mit Schaubildern ohne exakte Benennung der Einheiten auf x- und y-Achse.
5. Die DGOU bietet dem IQWiG und dem G-BA die klinische und methodische Expertise an, um für das gesundheitspolitisch und –ökonomisch relevante Aktionsfeld „Muskuloskeletale Erkrankungen und Verletzungen“ ein Methodenpapier mit unmittelbarer Versorgungsrelevanz zu erstellen.

4. Literatur

- (1) Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11(2):88-94.
- (2) Huber-Wagner S, Lefering R, Qvick LM et al. Effect of whole-body CT during trauma resuscitation on survival: a retrospective, multicentre study. *Lancet* 2009;373(9673):1455-1461.
- (3) Huber-Wagner S, Biberthaler P, Haberle S et al. Whole-body CT in haemodynamically unstable severely injured patients--a retrospective, multicentre study. *PLoS One* 2013;8(7):e68880.
- (4) Stengel D, Ottersbach C, Matthes G et al. Accuracy of single-pass whole-body computed tomography for detection of injuries in patients with major blunt trauma. *CMAJ* 2012;184(8):869-876.
- (5) Beerekamp MS, Ubbink DT, Maas M et al. Fracture surgery of the extremities with the intra-operative use of 3D-RX: a randomized multicenter trial (EF3X-trial). *BMC Musculoskelet Disord* 2011;12:151.
- (6) Albert H. *Traktat über kritische Vernunft*. 5 ed. Stuttgart: UTB; 1991.
- (7) Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;2(8):e124.
- (8) Charlton BG. Fundamental deficiencies in the megatrial methodology. *Curr Control Trials Cardiovasc Med* 2001;2(1):2-7.

A.1.15 – Deutsche Gesellschaft für Psychoanalyse, Psychotherapie, Psychosomatik und Tiefenpsychologie e. V. (DGPT)

Anmerkungen zu
Allgemeine Methoden des IQWiG
Version 4.2 vom 18.06.2014

A) Vorbemerkungen:

1) Die Allgemeinen Methoden des IQWiG sind mittlerweile zu einem sehr komplexen und umfangreichen (260 seitigem) Richtlinienwerk geworden. Das Institut schreibt dazu in der Präambel (Seite iii): „Die „Allgemeinen Methoden“ sollen die Vorgehensweise des Instituts allgemein beschreiben. Welche konkreten einzelnen Schritte das Institut bei der Bewertung einer bestimmten medizinischen Maßnahme unternimmt, hängt unter anderem von der jeweiligen Fragestellung und von der vorhandenen wissenschaftlichen Evidenz ab. Insofern sind die „Allgemeinen Methoden“ als eine Art Rahmen zu verstehen. Wie der Bewertungsprozess im konkreten Einzelfall ausgestaltet ist, wird projektspezifisch detailliert dargelegt.“ Der in dem Zitat genannte Rahmen ist weit gespannt und lässt viele verschiedene Möglichkeiten offen für die Planung und Realisierung projektspezifischer Beurteilungsprozesse von Therapien. Diese Offenheit ist generell begrüßenswert, bewirkt aber eine Unsicherheit in der Beantwortung der Frage, welche Verfahren im projektspezifischen Einzelfall zur Beurteilung einer oder mehrerer Therapien von dem Institut eingesetzt werden und ob innerhalb dieses Spielraums die für den Einzelfall adäquaten oder zumindest geeigneten Verfahren ausgewählt werden.

2) Der genannte Spielraum wirkt sich auch auf meine Stellungnahme aus: Worauf soll sich die Kritik beziehen, auf die im Einzelfall anzuwendenden strengsten Richtlinien, die in der Regel an die methodisch anspruchsvollsten Voraussetzungen gebunden sind, oder an die „liberalsten“ Richtlinien, deren Voraussetzungen möglichst geringe sind? In dieser Frage habe ich mich im Zweifelsfall an der ersten der beiden Antwortmöglichkeiten orientiert; die methodisch anspruchsvollsten Verfahren können nämlich bei Nichterfüllung ihrer Voraussetzungen am ehesten zu Fehlentscheidungen führen.

3) In der Präambel (Seite ii) steht: „Die Textteile, in denen Änderungen im Vergleich zur vorherigen Version der Allgemeinen Methoden vorgenommen wurden, sind gelb markiert bzw. als gelöschter Text sichtbar. Falls in Textabschnitten zahlreiche Änderungen auftraten, sind diese zur Verbesserung der Lesbarkeit in alter und neuer Fassung hintereinander dargestellt. Nur die als Änderung gekennzeichneten Textteile sind Gegenstand des Stellungnahmeverfahrens.“ An diese Vorgabe werde ich mich im Folgenden halten. Dabei spielt das (insgesamt gelb markierte) Kapitel 4 eine dominante Rolle: Auf dieses Kapitel entfällt der überwiegende Teil (geschätzte 80 – 90%) aller Textänderungen.

4) Meine Stellungnahme wurde von der DGPT in Auftrag gegeben; vereinbarungsgemäß sollte darin der IQWiG-Entwurf in erster Linie unter dem Aspekt von Psychotherapieverfahren betrachtet werden mit einem besonderen Gewicht auf der Frage, ob durch die einzelnen Richtlinien des IQWiG-Konzeptes tendenzielle Benachteiligungen in der Beurteilung von Psychotherapieverfahren oder Psychotherapieforschung impliziert werden. Das betrifft, wenn überhaupt, nur das neu aufgenommene (gelb markierte) Kapitel 4; die übrigen Kapitel enthalten keine gravierenden Benachteiligungen der Psychotherapie, höchstens solche, die generell für Therapiegebiete mit beabsichtigten Langzeitwirkungen, wie beispielsweise der Chirurgie, vorhanden sind.

5) Die unter 3) und 4) angeführten Argumente lassen es als angemessen erscheinen, mich in meiner Stellungnahme auf Kapitel 4 zu beschränken. Dabei stütze ich mich auf die gleichen Problemfelder wie in meiner letzten Stellungnahme zur Kosten-Nutzen-Analyse des IQWiG. Die darin von mir aufgeführten **sechs Probleme** behandle ich hier (unter den Punkten C bis H) in überarbeiteter, auf die neue Version bezogene Form, auch unter Berücksichtigung der Frage, wie weit meine damals geäußerten Bedenken und Kritiken beachtet und gegebenenfalls entkräftet wurden. Im Fokus meiner Stellungnahme zum IQWiG-Methodenpapier steht also die **Kosten-Nutzen-Bewertung medizinischer Interventionen (Kapitel 4)**

B) Ablauf einer Kosten-Nutzenbewertung des IQWiG

Die vorliegende Kosten-Nutzen-Bewertung (KNB) medizinischer Interventionen stellt eine gründlich überarbeitete und wesentlich erweiterte Fassung des bisherigen Entwurfs einer Kosten-Nutzen-Analyse des IQWiG dar. Ihr zentraler Teil besteht (wie bisher) in einer indikationsspezifischen Bewertung von Nutzen und Kosten aller für die Indikation in Frage kommenden Therapien und die daraus bestimmte **Effizienzgrenze**, ein in der betriebswirtschaftlichen Entscheidungstheorie wohlbekanntes Instrument (Kurve der „admissible decision rules“). Der Ablauf einer IQWiG-Kosten-Nutzen-Analyse besteht aus den Schritten:

- (1) Diagnosestellung, Indikationsfeld; die Kosten-Nutzen-Analyse einer Therapie bezieht sich auf ein festes Indikationsfeld.
- (2) Erfassung der vorhandenen konkurrierenden Therapien (A, B, C usw.) für das Indikationsfeld.
- (3) Konstruktion einer kardinalskalierten Bewertung des Nutzens der Therapien des Indikationsfeldes.
- (4) Aufstellung der Gesamtkosten für jede der in Betracht gezogenen Therapien (A, B, C usw.).
- (5) Bestimmung der (auf die Therapien A, B, C usw.) bezogenen Effizienzgrenze im Kosten-Nutzen-Koordinatensystem.
- (6) Bestimmung von Nutzen und Kosten der neu zu beurteilenden Therapie X; ihre Beurteilung durch ihre relative Lage zur Effizienzgrenze im Kosten-Nutzen-Diagramm: Liegt X oberhalb oder auf der Effizienzgrenze, so wird X prinzipiell zugelassen, anderenfalls nicht.
- (7) Aus der Effizienzgrenze bestimmte Richtlinien zur Festlegung von Fest- oder Höchstbeträgen für die Erstattung der Kosten von X (quantitative Spezifizierung des sechsten Schrittes).
- (8) Durchführung von Sensitivitätsanalysen zur Untersuchung der Abhängigkeit der Ergebnisse von Unsicherheiten innerhalb der benutzten Daten und Modellannahmen.

C) Erstes Problem: Kardinalskaliertes Nutzen

Auch in der neuen Version des IQWiG-Entwurfes wird die Voraussetzung, dass der Nutzen einer Therapie oder Intervention auf einer eindimensionalen Kardinalskala gemessen werden kann, als zentrale Forderung erhoben. Ihre Erfüllung ist zur Konstruktion einer Effizienzgrenze unverzichtbar. Zur Problematik einer solchen Voraussetzung die folgenden Punkte (teilweise in Übernahme aus meiner letzten Stellungnahme):

1. Der Nutzen einer Therapie kann auf vielen Behandlungsgebieten, insbesondere dem Gebiet der Psychotherapie, in den allermeisten Fällen nur auf einer Ordinalskala gemessen werden. In allen diesen Fällen stellt die Konstruktion einer Kardinalskala eine mehr oder weniger starke Willkür mit einer gewissen Beliebigkeit dar. Stichworte: Eine Ordinalskala ist nur unter sehr einschneidenden Voraussetzungen (Neumann-Morgenstern-Axiome) in eine Kardinalskala transformierbar. Eine Ordinalskala ist frei gegenüber monotonen Transformationen, eine Kardinalskala nur gegenüber linearen Transformationen. Die kardinale Skalierbarkeit stellt eine wesentlich stärkere Einschränkung an eine Variable dar als ihre ordinale Skalierbarkeit
2. Häufig wird man zu Nutzenmessung einer Intervention nur mehr oder weniger „gewaltsam“ eine Kardinalskala konstruieren können, oft nur, indem ein Kompromiss durch Einigung auf einen kleinsten gemeinsamen Nenner zustande kommt (z.B. der Summenskore verschiedener Messinstrumente, die Verweildauer stationärer Behandlungen, o.a.). Die so konstruierte Skala kann die Wirkung der zu vergleichenden Interventionen sehr unterschiedlich widerspiegeln, manche werden sehr valide, andere nur sehr inadäquat erfasst werden. Alle Interventionen werden aber mit dieser einen Skala verglichen, also quasi „über einen Kamm geschoren“. Diese gleichstellende Beurteilung der Interventionen trifft psychotherapeutische Behandlungen besonders empfindlich.
3. Die Voraussetzung eines kardinalskalierbaren Nutzens geht mit ihrem vollen Gewicht in die Konstruktion der Effizienzgrenze ein, dem entscheidenden Kriterium des Kosten-Nutzen-Vergleiches: *„Damit der Nutzen in die KNB mithilfe der Effizienzgrenze einfließen kann, muss er approximativ kardinalskaliert sein.“* (IQWiG-Methodenpapier Seite 86.) Die Einschränkung „approximativ“ ist hier unwesentlich; sie wird im Methodenpapier (Seite 86) wie folgt verstanden: *„Es genügt, wenn diese (die Nutzenskala) über den Bereich, der für die Definition des patientenrelevanten Zusatznutzens von Bedeutung ist, das Kriterium erfüllt, kardinalskaliert zu sein.“*

4. Die (von mir und anderen angeregte) Kritik über die Voraussetzung eines kardinalskalierbaren Nutzens enthält neben dem grundsätzlichen messtheoretischem Problem (siehe Punkt 1) auch die Frage, wie stark sich eine Nichterfüllbarkeit dieser Voraussetzung auf die Konstruktion einer Effizienzgrenze auswirkt, wird nicht behandelt. Zur Beantwortung dieser Frage wären sehr schwierige und tiefgehende Sensitivitätsanalysen erforderlich. Diese können meines Erachtens nach aber nicht oder nur in Ausnahmefällen die erforderliche Robustheit einer Effizienzgrenze ergeben, (Robustheit gegenüber einem vorhandenen Freiraum von Umskalierungen der Nutzenbewertung). Die hier genannten Sensitivitätsanalysen sind von ganz anderer Art als die im IQWiG-Methodenpapier in den Abschnitten 4.7 und 4.9 genannten.)

5. In der Regel gibt es für ein Indikationsfeld mehrere verschiedene Erfolgsmaße zur Bewertung von Behandlungen; darauf geht das IQWiG in Kapitel 3 ausführlich ein. Diese Maße führen dann zunächst einmal zu einer mehrfaktoriellen Bewertung mit mehreren (eindimensionalen) Nutzenfunktionen. Für diese komplexe realistische Situation einer mehrdimensionalen Nutzenwertung (mit mehreren Effizienzgrenzen) schlägt das IQWiG zwei Wege vor, die zu einer eindimensionalen Bewertungsskala führen sollen: Die Festlegung eines „übergeordneten“ Maßes für den Gesamtnutzen oder eine gewichtete Mittelbildung der einzelnen Nutzenfunktionen. Im ersten Fall kann und soll der G-BA maßgebend werden (die gängigen Maße dazu werden in Abschnitt 4.3.3 vorgestellt), im zweiten soll es dem Entscheidungsträger vorbehalten bleiben, eine Gewichtung vorzunehmen (siehe Abschnitt 4.1.9.). Beide Wege enthalten eine gewisse unvermeidliche Willkür, die in die Nutzenbewertung einer Therapie eingeht.

6. Mit der Einbeziehung mehrdimensionaler Nutzenbewertungen eröffnet das IQWiG erfreulicherweise das enge Feld kardinalskalierter (eindimensionaler!) Nutzenfunktionen. Die dabei unter 5. angegebene unvermeidliche Willkür kann und soll einer Sensitivitätsanalyse unterzogen werden, wie sie in Abschnitt 4.7 (insbesondere 4.7.2) beschrieben werden, um damit die Frage zu beantworten, wie stark die Bewertung einer Therapie von dem übergeordneten Maß oder der Gewichtung einzelner Erfolgsmaße abhängt. Das gleiche gilt für die in Abschnitt 4.2.1 angegebene Modellbildung zur Zusammenführung von Nutzen und Kosten; dazu heißt es auf Seite 81 des Methodenpapiers: „Zumeist werden in einer Studie nicht alle für die Entscheidung relevanten Variablen erhoben. Auch wird ein gesundheitsökonomisches Modell explizit dazu verwendet, Nutzen und Kosten über den studienbelegten Zeitraum hinaus zu extrapolieren. Gesundheitsökonomische Modelle sind somit, wie mathematisch-formalisierte Modelle, auch eine vereinfachte Abbildung der Realität. Durch eine bewusste Reduktion der Komplexität auf die für das Entscheidungsproblem relevanten Entscheidungsfaktoren und -variablen wird zudem analytische Klarheit geschaffen“.

7. Zusammenfassend kann ich feststellen: Wie bisher besteht das Kernstück einer Kosten-Nutzen-Bewertung des IQWiG in der Bestimmung einer Effizienzgrenze unter Verwendung einer kardinalskalierten (eindimensionalen) Nutzenfunktion. Darauf treffen die oben angeführten Kritikpunkte 1 bis 4 zu. Andererseits werden in der neuen Version des Methodenpapiers auch solche Richtlinien und Verfahren entwickelt, wie man die sehr komplexe Situation einer mehrdimensionalen oder auch nur ordinalskalierten Nutzenbewertung behandeln soll (um eine eindimensionale Kardinalskala zu erhalten); siehe dazu die oben angeführten Punkte 5 und 6. Damit wird in der neuen Version des IQWiG-Methodenpapiers auf die Problematik der Annahme einer eindimensionalen kardinalskalierten Nutzenbewertung ausführlich eingegangen, einige meiner früher geäußerten Kritikpunkte zu diese Thema wurden insoweit relativiert.

D) Zweites Problem: Maße des Gesamtnutzens (früher: Respondermaße)

Meine kritischen Bemerkungen zu Respondermaßen, die ich bezüglich früherer Versionen von IQWiG-Methodenpapieren vorgebracht habe, werden durch die neue Version entkräftet. Die ausführlichen Darstellungen verschiedener Maße für den Gesamtnutzen einer Therapie in Abschnitten 4.3.2 und 4.3.3 zeigen, wie problematisch es auch für das IQWiG-Konzept ist, den Gesamtnutzen (möglichst noch kardinal) zu messen. Es ist gut, dass nun dieses Problem in seiner vollen Bedeutung einbezogen wird. Dabei wird zu Recht keine allgemeine Lösung des Problems präjudiziert, sondern darauf verwiesen, dass nur fallweise Lösungen angestrebt werden können. Dazu die folgenden Zitate (Seite 87): „Hier wird keine spezielle Vorgehensweise zur Wertbestimmung des Nutzens auf einer Kardinalskala empfohlen, da jedes Indikationsgebiet unterschiedliche Möglichkeiten zur Bewertung des Nutzens bieten kann, die der Anforderung nach Kardinalskalierung genügen.“ Und: „Ein Maß des Gesamtnutzens ist in einer ganz allgemeinen Definition eine Aggregation der Bewertung von Nutzen und Schaden in einer Größe, wobei unterschiedliche patientenrelevante Endpunkte zu einem einzigen Maß zusammengefasst werden. Und weiterhin: „International gibt es unterschiedliche Maße, den Gesamtnutzen auszudrücken bzw. zu erfassen. Dazu gehören das qualitätsadjustierte Lebensjahr (QALY = Quality-Adjusted Life Year) und das behinderungsbereinigte Lebensjahr (DALY = Disability-Adjusted Life Year). Maße wie das „Saved Young Life Equivalent [404]“ oder das „Healthy Years Equivalent“ (HYE) [203] sind mit dem Anspruch angetreten, Schwächen des QALY zu korrigieren, das am weitesten verbreitet ist.“

E) Drittes Problem: Einbeziehung von Unsicherheiten

In meinen Stellungnahmen zu früheren Versionen des IQWiG-Methodenpapiers hatte ich kritisiert, dass bei der Konstruktion von Effizienzgrenzen die verschiedenen Unsicherheiten nicht berücksichtigt werden, die in dieser Grenze enthalten sind, und keine entsprechenden Sensitivitätsanalysen durchgeführt werden. Dieses Defizit wird in der vorliegenden Version (teilweise auch schon in ihrem Vorläufer) behoben: Die verschiedenen Ursachen von Unsicherheiten werden in Abschnitt 4.1.8 vorgestellt und in verschiedenen weiteren Abschnitten, vor allem in 4.7 und 4.9 verfahrensgemäß behandelt. Zitat aus Seite 106: *„Das Institut berücksichtigt sowohl univariate und multivariate deterministische als auch probabilistische Sensitivitätsanalysen und legt bei der Bearbeitung die Empfehlungen der gemeinsamen Modeling Good Research Practices Task Force Working Group der ISPOR und SMDM zugrunde“*. Auch die Auswirkung von Unsicherheiten auf die Konstruktion einer Effizienzgrenze und Beurteilung der Zulässigkeit einer Therapie wird berücksichtigt.

F) Viertes Problem: Abhängigkeit der Nutzenskala von einer Therapie

Dieses von mir früheren Stellungnahmen vorgetragene Problem wird in der neuen Version des IQWiG-Entwurfes nicht berücksichtigt, ich gebe daher hier noch einmal den Text aus meiner er letzten Stellungnahme (geringfügig verändert) wieder.

Das IQWiG Konzept geht davon aus, dass sich der Nutzen einer Therapie auf einer „allgemeingültigen“ (kardinalen) Skala messen lässt, die nur von dem betreffenden Indikationsgebiet, nicht aber von der eingesetzten Therapie abhängt. Ist dieses Postulat wirklich realistisch? Hängt nicht in vielen Fällen die Nutzenbetrachtung auch von der eingesetzten Therapie ab? Werden nicht in der gesamten modernen Medizin neue Therapien gerade in Hinblick auf neue (auch ethisch begründete) Heilungsziele entwickelt? Besonders betroffen von diesen Fragen sind die Bereiche der Psychiatrie und Psychotherapie, wo immer schon die Therapieform mit dem Heilungsziel und daher auch der Nutzenbewertung mehr oder weniger eng zusammenhing. Kann man z.B. einen Psychiater dazu „zwingen“, die OPD-Skala zu verwenden oder einen Psychoanalytiker, sich auf die BDI-Skala zu verlassen? Meine Kritik zielt nicht auf schulenspezifische Befindlichkeiten, sondern entstammt Grundsätzen der Entscheidungstheorie: Bei Abhängigkeit der Nutzenskala von den Handlungsalternativen gerät man vielmehr grundsätzlich in ein entscheidungstheoretisches Dilemma.

G) Fünftes Problem: Indikationsspezifische Bewertung von Nutzen und Kosten

Zu diesem Problem lautete meine Stellungnahme zu der letzten Version (in verkürzter Form): Nach dem IQWiG Konzept bezieht sich der Nutzen einer Therapie und die betreffende Skala stets auf eine bestimmte Krankheit (Diagnose) innerhalb eines festgelegten Indikationsgebietes. Diese Spezifizierung ist problematisch: Wie kann eine Therapie beurteilt werden, die bei verschiedenen Krankheiten (zwar mit unterschiedlichen Nutzen, aber doch sehr gut) wirkt? Und umgekehrt: Wie wird eine Therapie beurteilt, die bei einer bestimmten Krankheit einen sehr großen (maximalen) Nutzen unter allen anderen Therapien hat, aber nur auf diese eine Diagnose „zugeschnitten“ und anderenfalls untauglich ist?

Der indikationsspezifische Ansatz wird in der neuen Version im Wesentlichen beibehalten. Das hat sicherlich auch seine guten Gründe: Verschiedene Interventionen sind innerhalb eines Indikationsgebietes viel einfacher und klarer vergleichbar; außerdem kann eine Therapie in einem Indikationsfeld gut, in einem anderen schlecht abschneiden und deshalb auch diagnosespezifisch verordnet werden. Allerdings gibt es auch Passagen in der neuen Version, die den streng indikationsspezifischen Ansatz auflockern. Beispielsweise heißt es in Abschnitt 4.1.9: *„Wenn ein Maß des Gesamtnutzens festgelegt wird, ist dieses als primäres Ergebnis anzusehen. Ist die Bestimmung mehrerer Effizienzgrenzen für die Bewertung einer Intervention erforderlich, so bleibt es dem Entscheidungsträger unter Beachtung der Relevanz der patientenrelevanten Endpunkte vorbehalten, eine Gewichtung vorzunehmen. Eine ähnliche Vorgehensweise kann auch bei Zulassung einer Intervention in mehreren Indikationsgebieten erwogen werden.“*

H) Sechstes Problem: Lebensqualität und Salutogenese

In früheren Stellungnahmen hatte ich die mangelnde Einbeziehung der durch eine Therapie gewonnenen Lebensqualität kritisiert. Diese Kritik wird durch die neue IQWiG-Version entkräftet: In Abschnitt 4.3 werden verschiedenen Maße behandelt, durch die die gewonnene Lebensqualität als Nutzenbewertung einer Therapie einbezogen wird; siehe dazu oben unter Punkt D). Die von mir schon früher vorgebrachte Anregung, auch Konzepte der Salutogenese („was macht einen Menschen fähig, gesund zu bleiben?“) mit einzubeziehen, wurde nicht aufgegriffen.

A.1.16 – Deutsche Gesellschaft für Ultraschall in der Medizin e. V. (DEGUM)

Stellungnahme der Deutschen Gesellschaft für Ultraschall in der Medizin

zum Entwurf 4.2 vom 18.06.2014 des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen IQWiG

Die Deutsche Gesellschaft für Ultraschall in der Medizin begrüßt die ausführliche Beschreibung der Datenerfassung und Auswertung der IQWiG, die in den „allgemeinen Methoden“ festgehalten sind.

Auch die überarbeiteten bzw. neu verfassten Kapitel Kosten-Nutzen-Bewertung, Potenzialbewertung, evidenzbasierte Gesundheitsinformation für Bürgerinnen und Bürger und Informationsbeschaffung werden grundsätzlich begrüßt.

Folgende Aspekte sollten zusätzlich Berücksichtigung finden:

Kapitel 2.1.4 Kosten-Nutzen-Bewertung, Seite 24

Einholung der medizinische Expertise

Bei der Einholung der medizinischen Expertise sollten in jedem Fall immer auch die medizinischen Fachgesellschaften, in deren Fachbereich die Kosten-Nutzen-Bewertung fällt, mit einer Stellungnahme beauftragt werden. Dabei sollte die Fachgesellschaft einen Sachverständigen aus den eigenen Reihen benennen können, der als Ansprechpartner für das IQWiG-Institut zur Verfügung steht.

Anstelle, dass die externen Sachverständigen für spezielle Fragestellungen optional herangezogen werden können, sollte eine Stellungnahme der externen Sachverständigen und der Fachgesellschaften verpflichtend sein.

Kapitel 2.1.4 Kosten-Nutzen-Bewertung, Seite 24

Gelegenheit zur Stellungnahme

Für eine Frist von drei Wochen wird der Öffentlichkeit Gelegenheit zur Stellungnahme (Anhörung) gegeben. Hier sollte ein längeres Zeitfenster eingeräumt werden, um den medizinischen Fachgesellschaften auch in Urlaubszeiten die Benennung eines geeigneten Sachverständigen zu ermöglichen.

Kapitel 2.1.5 Potenzialbewertung, Seite 25

Einholung der medizinischen Expertise

Als medizinische Expertise sollten die medizinischen Fachgesellschaften, deren Fachbereich durch die Potenzialbewertung tangiert wird, nicht nur optional sondern verpflichtend miteingebunden werden.

2.2.1 Auswahl externer Sachverständiger, Seite 33

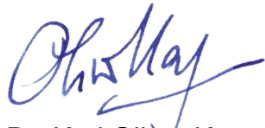
Als externe Sachverständige sollten immer auch die medizinischen Fachgesellschaften genannt werden, die aus ihren Reihen einen Sachverständigen benennen, der als Ansprechpartner für das IQWiG-Institut zur Verfügung steht.

Die Liste der externen Sachverständigen sollte offen online abrufbar sein.

3.8 Potenzialbewertung, Seite 74

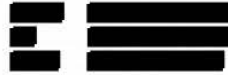
Antragsberechtigt sollten - neben dem Hersteller eines Medizinprodukts und einem Unternehmen mit wirtschaftlichen Interesse - auch Gruppen ohne wirtschaftliches Interesse, z.B. medizinische Fachgesellschaften sein.

Mit freundlichen Grüßen



Prof. Dr. Karl Oliver Kagan
für die
Deutsche Gesellschaft für Ultraschall in der Medizin

A.1.17 – Deutsche Krankenhausgesellschaft e. V. (DKG)



Datum 29.07.2014 

Institut für Qualität und
Wirtschaftlichkeit im Gesundheitswesen
Herr Prof. Dr. med. Jürgen Windeler
-Stellungnahme zu den „Allgemeinen
Methoden Version 4.2“ -
Im Mediapark 8 (Köln Turm)
50670 Köln

Stellungnahme der Deutschen Krankenhausgesellschaft zum Entwurf des IQWiG „Allgemeine Methoden Version 4.2“ vom 18.06.2014

Sehr geehrter Herr Prof. Windeler,
sehr geehrte Damen und Herren,

vielen Dank für die Zusendung des Entwurfs zu Ihren „Allgemeinen Methoden Version 4.2.“, zu dem wir wie folgt Stellung nehmen:

Zunächst möchten wir darauf hinweisen, dass wir die Kritikpunkte unserer ausführlichen Stellungnahmen aus den Vorjahren zu all denjenigen Punkten aufrechterhalten, die bisher unberücksichtigt blieben. Um Redundanzen zu vermeiden, werden sie in dieser Stellungnahme nicht nochmals aufgeführt. Es handelt sich hier teilweise um Aspekte, auf die wir bereits seit mehreren Jahren aufmerksam machen. Beispielsweise würden wir weiterhin eine obligatorische mündliche Anhörung zu den Stellungnahmen begrüßen, da hierbei Unklarheiten und Rückfragen zu den Stellungnahmen besser als auf dem Schriftwege geklärt werden könnten. Die mündlichen Anhörungen möchten wir als konstruktiven Dialog verstehen.

Zu Kapitel 2.1.5 Potenzialbewertung

Die Darstellung der einzelnen Arbeitsschritte im Rahmen der Potenzialbewertung wird von uns ausdrücklich begrüßt. Dennoch bleiben aus unserer Sicht noch einige wichtige Aspekte unklar. Bei der Darstellung der als optional bezeichneten eigenen Recherche zur Unterstützung der Bewertung in Kapitel 2.1.5 bleibt offen, in welchem Kontext diese Recherche konkret durchgeführt wird. Unklar bleiben aus unserer Sicht hier insbesondere Umfang und Inhalt, aber auch der Anlass, wann eine derartige Recherche als angezeigt erachtet wird. Unabhängig von der fehlenden Konkretisierung ist aus unserer Sicht eine Recherche des medizinischen Hintergrundes nicht nur optional durchzuführen, sondern bei jeder Potenzialbewertung zu fordern, um den Grundsatz der Gleichbehandlung zu gewährleisten. Eine Umfeldrecherche, die die Einordnung der zu bewertenden Untersuchungs- oder Behandlungsmethode in einem Gesamtversorgungskontext ermöglicht, ist aus unserer Sicht für eine differenzierte Einschätzung

des Potenzials unumgänglich. Eine derartige Recherche sollte die Grundlage der Bewertung aller vom G-BA an das IQWiG übermittelten Anträge sein und nicht nur optional in Einzelfällen durchgeführt werden.

Zu Kapitel 3.8 Potenzialbewertung

1.) „Potenzialbewertungen zielen im Gegensatz zu Nutzenbewertungen darauf ab zu prüfen, ob neue Untersuchungs- oder Behandlungsmethoden möglicherweise einen Nutzen aufweisen. Potenzial bedeutet hierbei, dass erstens die bisher vorliegenden Erkenntnisse einen möglichen Nutzen erkennen lassen...“

Die Beschreibung der Potenzialbewertung wird entgegen den gesetzlichen Vorgaben in diesem Abschnitt aus unserer Sicht sprachlich nicht deutlich genug von der Nutzenbewertung abgegrenzt. Im Rahmen der Potenzialbewertung wird ausschließlich untersucht, ob ein mögliches Potenzial einer Untersuchungs- und Behandlungsmethode vorliegt und keine Nutzenbewertung durchgeführt. Diese erfolgt im Rahmen einer Erprobung erst im Anschluss an die Durchführung einer Erprobungsstudie. Die Definition des Potenzials wird von Ihnen vielmehr im nachfolgenden Satz klar formuliert. Unserer Meinung nach bildet die folgende Formulierung die Abgrenzung der Potenzialbewertung von der Nutzenbewertung deutlicher ab:

„Potenzialbewertungen zielen im Gegensatz zu Nutzenbewertungen darauf ab zu prüfen, ob neue Untersuchungs- oder Behandlungsmethoden ein Potenzial aufweisen. Potenzial bedeutet hierbei, dass erstens die bisher vorliegenden Erkenntnisse Hinweise auf einen möglichen Nutzen erkennen lassen...“

2.) „Im Rahmen der Potenzialbewertung werden vom Institut die Plausibilität der Angaben des Antragstellers geprüft. Diese Prüfung bezieht sich insbesondere auf die Sinnhaftigkeit der im Antrag dargestellten medizinischen Fragestellung(en), die Güte der vom Antragsteller durchgeführten Literatursuchen (siehe Abschnitt 7.2), die Einschätzung der Ergebnissicherheit der relevanten Studien und die Korrektheit der im Antrag dargestellten Ergebnisse. Die Bewertung mündet in einer Aussage zum Potenzial der beantragten Untersuchungs- oder Behandlungsmethode“

In Kapitel 3.8 werden die einzelnen Arbeitsschritte im Rahmen der Potenzialbewertung und insbesondere methodische Aspekte zur Bewertung der Ergebnissicherheit nicht randomisierter Studien näher konkretisiert. Aus unserer Sicht bleibt jedoch unklar, welche Maßstäbe letztendlich in der Zusammenschau der vom Antragssteller eingereichten Unterlagen angelegt werden und so zu dem Ergebnis eines positiven bzw. negativen Potenzials führen.

Zu Kapitel 4.1.9

Interpretation der Ergebnisse zur Kosten-Nutzen-Bewertung

Das IQWiG hat zur Bestimmung eines angemessenen Preises für Innovationen die Methode der Effizienzgrenze entwickelt. Mit der Methode der Effizienzgrenze wird ermittelt, welche zusätzlichen Kosten ein festgestellter (Zusatz-) Nutzen eines Arzneimittels bzw. einer Gesundheitstechnologie verursachen darf. Daraus leitet das IQWiG im Verfahren nach § 35b SGB V eine Empfehlung an den G-BA zur Höhe des Erstattungsbetrages des Arzneimittels ab. Im Kern basiert die Methodik der

Effizienzgrenze darauf, dass die Kosten-Nutzen-Verhältnisse schon im Markt befindlicher Arzneimittel ermittelt werden und durch eine Extrapolation der bisher zu verzeichnenden Nutzenzuwächse bzw. Effizienzsteigerungen ein Preis für das zu bewertende neue Arzneimittel mathematisch abgeleitet wird.

Aus Sicht der DKG kann die Methodik einen wichtigen Beitrag zur Festlegung angemessener Erstattungsbeträge für Arzneimittel im Verfahren nach § 35b SGB V leisten. Eine zentrale Schwachstelle der Methodik der Effizienzgrenze ist aus unserer Sicht aber, dass die Effizienzgrenze regelhaft auf die bestehende Marktkonstellation der im jeweiligen Anwendungsgebiet verfügbaren Arzneimittel aufsetzt.

Dies führt dazu, dass die Anwendung der Effizienzgrenze in idealtypischen Marktkonstellation zwar grundsätzlich zu sachgerechten Ergebnissen führen kann (siehe u.s. Abbildung 1).

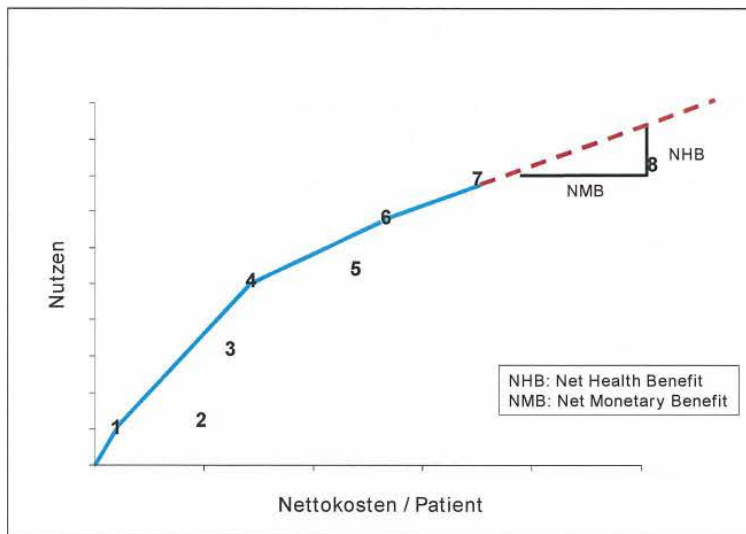


Abbildung 1 (aus: G-BA, „Dossier zur Kosten-Nutzen-Bewertung gemäß § 35b SGB V“, Modul K 4, S. 81)

Bei der Anwendung der Effizienzgrenze in Marktkonstellationen, die nicht dem idealtypischen Bild entsprechen, führt die Methodik der Effizienzgrenze allerdings zwangsläufig zu völlig anders gelagerten Ergebnissen (siehe u.s. Abbildung 2).

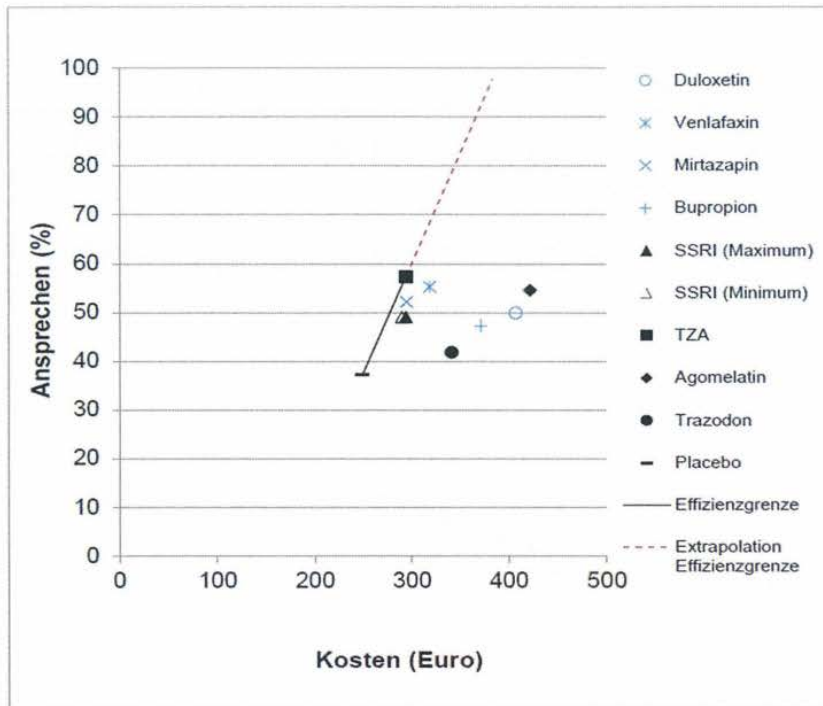


Abbildung 17: Effizienzgrenze für den Endpunkt Ansprechen, studienbelegter Zeithorizont, GKV-Versichertenperspektive

Abbildung 2 (Die Abbildung wurde exemplarisch aus Abschlussbericht des IQWiG vom 03.09.2013, „Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medikamentösen Behandlungen“, S. 207 entnommen.)

Dies liegt daran, dass mit der Methodik der Effizienzgrenze, technisch ausgedrückt, die Steigung des letzten Abschnitts der Effizienzgrenze extrapoliert wird um daraus den Preis für das zu bewertende Arzneimittel abzuleiten. Die Steigung der Effizienzgrenze wird aber durch die jeweilige Marktkonstellation der verfügbaren Präparate bestimmt und kann deswegen über verschiedene Anwendungsgebiete, aber auch im Zeitverlauf, völlig unterschiedlich sein. Die Extrapolation der Effizienzgrenze aus Abbildung 1 führt zu völlig anders gelagerten Ergebnissen, als die Extrapolation der Effizienzgrenze aus Abbildung 2 (siehe jeweils gestrichelte Linie). Das bedeutet im Ergebnis, dass bei einem Verlauf der Effizienzgrenze wie in Abbildung 2 ein Zusatznutzen eines Arzneimittels deutlich weniger kosten darf, wie ein vergleichbarer Zusatznutzen, bei einem Verlauf der Effizienzgrenze aus Abbildung 1.

Da die Methodik der Effizienzgrenze in erster Linie von der jeweiligen Marktkonstellation abhängig ist, führt sie in vielen Fällen zu inkongruenten Ergebnissen und damit zu unsachgerechten Arzneimittelpreisen. Dies wurde auch durch die bisher vorliegende Kosten-Nutzen-Bewertung zu Antidepressiva deutlich. Das IQWiG hatte

nach diesem Vorgehen für die zu bewertenden Präparate mit belegtem Nutzen erhebliche Preissenkungen von bis zu 98,5 Prozent abgeleitet. Beispielsweise liege der angemessene Preis für den Wirkstoff Bupropion bei 1,48 € (Herstellerpreis 104,88 €) und für Duloxetin bei 9,30 € (Herstellerpreis 241,18 €). Die Anwendung der Effizienzgrenze zeigte in diesem ersten Probelauf, dass diese Methode in bestimmten idealtypischen Marktkonstellationen unterstützend sinnvoll sein kann, dass aber eine ausschließliche Ableitung von Preisen durch Extrapolation historischer Nutzenzuwächse zu realitätsfernen Ergebnissen führt.

Aus Sicht der Deutschen Krankenhausgesellschaft sind Kosten-Nutzen-Bewertungen von Arzneimitteln perspektivisch sinnvoll, diese müssen aber nach realistischen Maßstäben erfolgen. Im Ergebnis ist aus unserer Sicht die Anwendung der Methodik der Effizienzgrenze nur eingeschränkt, in idealtypischen Marktkonstellationen, möglich. Da idealtypische Marktkonstellationen allerdings nicht die Regel, sondern die Ausnahme sind, ist zu befürchten, dass die Anwendung der Methodik der Effizienzgrenze nur sehr eingeschränkt möglich ist oder in vielen Fällen zu unsachgerechten Ergebnissen führt. Insgesamt ist aus Sicht der DKG deshalb eine grundlegende Weiterentwicklung der Methodik der Effizienzgrenze erforderlich.

Zu Kapitel 7.2 Überprüfung der Informationsbeschaffung

In Kapitel 7.2 legen Sie dar, wie Sie die Informationsbeschaffung in Dossiers bzw. Anträgen zur Erprobung überprüfen und verweisen auf das in Abschnitt 7.1.1 beschriebene Vorgehen zur Qualitätssicherung. Kapitel 7.1.1 bezieht sich jedoch auf die Informationsbeschaffung des IQWiG im Rahmen von Nutzenbewertungen. Unklar bleibt aus unserer Sicht daher, wie die formale Überprüfung der Literaturrecherche im Rahmen der Erprobung konkret erfolgt, da die Besonderheit der Potenzialbewertung u.a. in der kurzen, maximal sechswöchigen Bearbeitungsdauer, die dem IQWiG zur Verfügung steht, besteht. Vor dem Hintergrund, dass die Bewertung des Potenzials insbesondere auf Grundlage der vom Antragsteller eingereichten Unterlagen erfolgt, spielt die Beurteilung der Validität des Antrags eine entscheidende Rolle. Eine zentrale Funktion kommt dabei der Information zu, ob der für die jeweilige Fragestellung relevante Studienpool durch die Antragsunterlagen vollständig erfasst wird oder aber eine selektive Berichterstattung durch den Antragsteller erfolgt. Eine konkrete Beschreibung der einzelnen Schritte, wie die Überprüfung der Informationsbeschaffung bei der Potenzialbewertung erfolgt, würde aus unserer Sicht dazu beitragen, den Unsicherheitsfaktor der Validität von Erprobungsanträgen besser einschätzen zu können.

Diese Stellungnahme ist als institutionelle Meinungsäußerung aufzufassen, an der verschiedene Mitarbeiter der DKG beteiligt waren. Im Falle einer Anhörung gehen wir davon aus, dementsprechend Vertreter der DKG zu entsenden.

Mit freundlichen Grüßen
Der Hauptgeschäftsführer
In Vertretung



Dr. Nicole Schlottmann
Geschäftsführerin Dez. Medizin

A.1.18 – Deutscher Psoriasis Bund e. V. (DPB)



Deutscher Psoriasis Bund e.V. (DPB)

Selbsthilfe bei Schuppenflechte seit 1973

Seewartenstraße 10 · 20459 Hamburg

Telefon 040/22 33 99 0 · Telefax 040/22 33 99 22

E-Mail: info@psoriasis-bund.de

Internet: www.psoriasis-bund.de

DPB · Seewartenstraße 10 · 20459 Hamburg

Herrn
Prof. Dr. Jürgen Windeler
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Vorstand:
Horst von Zitzewitz (Vorsitzender)
Otffrid Hillmann (stellv. Vorsitzender)
Annette Behlau-Schnier
Marius Grosser



Hamburg, 30. Juli 2014

Sehr geehrter Herr Professor Windeler,

es ist sehr schade, dass der DPB erst über Dritte und sehr kurzfristig erfährt, dass das Methodenpapier des IQWiGs verändert werden soll und Stellungnahmen wohl nur bis Ende des Monats „angenommen“ werden. Eine sachgerechte Auseinandersetzung innerhalb des DPBs ist damit leider nicht möglich.

Einem Institut, das sich nur der Evidenz verschworen hat, sollte belegen und zwar mit evidenten Studien, warum patientenrelevante Endpunkte, zu denen auch Patientenzufriedenheit zu zählen sind, aus dem Methodenpapier eliminiert werden sollen (S. 39). Der Streichung ist keine Begründung zu entnehmen. Gleiches gilt für den gestrichenen Hinweis auf § 35b SGB V auf der gleichen Seite, der einen inhaltlichen Bezug bezüglich einer Betrachtung von „Nutzen“ herstellt.

Wenn das IQWiG sein Methodenpapier verändert, was durch neue, evidente Erkenntnisse sicherlich von Fall zu Fall notwendig sein kann, sollte das IQWiG mit der geforderten wissenschaftlichen Evidenz belegen, warum Änderungen vorgenommen werden müssen. Diese Darlegung der medizin-wissenschaftliche Notwendigkeit bedarf einer Basis. Ansonsten wirken Veränderungen als Willkür, auf der „Wissenschaftlichkeit“ aufgebaut wird.

Zwischen den Systembetrachtungen des IQWiGs und der Lebensrealität der gesetzlich Versicherten, die auch das IQWiG mit Beiträgen aus Arbeitseinkommen voll finanzieren, klafft eine Lücke. Aus Patientensicht hat das Versorgungssystem dafür Sorge zu tragen, dass Kranke und Behinderte trotz ihrer körperlichen Beschränkungen in ihrer individuellen Lebenswelt größtmögliche Lebenszufriedenheit als wesentliches Element der persönlichen Lebensqualität wieder erreichen oder bestmöglich erhalten können. Es ist nicht zu erkennen, wie das IQWiG die Lücke zur realen Lebenssituation individuelle Lebenswelten von kranken Menschen „wissenschaftlich-methodisch“ schließen will. Unbegründete Streichungen im angewandten Methodenpapier leisten dies wohl nicht.

Die IQWiG-Methoden-Überlegungen beschränken sich auf das weltweite Auffinden und Bewerten von Studien. Das heißt, dass Ergebnis bleibt dem Zufall der Existenz von Studien überlassen. Es kann absolut nicht der Schluss gezogen werden: Keine Studie = keine Evidenz.

Es gibt kaum Studien zur Lebensqualität bei Krankheit, wo die eingesetzten Instrumente maßgeblich gemeinsam mit Patienten entwickelt wurden. Der dermatologische Lebensqualitätsfragebogen

Mitgliedschaften:



Gedruckt auf
Umweltschutzpapier

(DLQI), der einen Weg zur Therapie mit Systemtherapeutika bei Psoriasis und Psoriasis-Arthritis in der GKV eröffnet, wurde vom Briten Finlay formuliert. Welche Lebenssituation von Kranken war damals die Blaupause für seinen Fragebogen? Haben sich die Verhältnisse selbst in Großbritannien nicht zwischenzeitlich geändert? „Patientinnen und Patienten sind gewissermaßen die „Endverbraucher“ von medizinischen Maßnahmen. Deshalb werden sie international in die Bewertungen von Nutzen und Kosten einbezogen. Allerdings geschieht dies bislang nicht systematisch, transparent und nachvollziehbar. Zudem werden dabei in der Regel rein qualitative Verfahren eingesetzt“ (Quelle IQWiG). Wie wurden die Bedingungen der individuellen Lebenswelten von Menschen mit Psoriasis in die Formulierung und Bewertung des Fragebogeninstruments einbezogen?

Es scheint, dass grundsätzlich wieder einmal Patienten im Mittelpunkt stehen, allerdings nur als Medium einer methodischen Wissenschaft die Systeminteressen bedient.

Was nicht auf Dauer zufrieden macht, nützt selbst bei bester Evidenz nichts, da es nicht wirkt (Nocebo-Effekt).

Mit freundlichem Gruß

DEUTSCHER PSORIASIS BUND e. V.


Hans-Detlev Kunz
Geschäftsführer

A.1.19 – Deutsches Netzwerk Versorgungsforschung e. V. (DNVF)



Stellungnahme zum IQWIG Methodenpapier 4.2

des Deutschen Netzwerks Versorgungsforschung e.V.

vom 29.7.2014

Zur Stellungnahme liegt die Version 4.2 des Methodenpapiers des IQWIG vor. Nur die bearbeiteten Stellen sind Gegenstand der Kommentierung.

Als wesentliche Neuerung sind Abschnitte zur neu eingeführten Potenzialbewertung anzusehen (Kapitel 2.1.5, 3.8 neu). Das geschilderte Vorgehen scheint prinzipiell angemessen; das betrifft insbesondere auch die niedrigeren Ansprüche an die Evidenzlage einschließlich der Verwendung von Surrogatendpunkten, sowie der Berücksichtigung der Effektstärke beim Anspruch an die Validität. Es sei erwähnt, dass die Potenzialbewertung komplett in Verantwortung des IQWIG erfolgt, Hinzuziehung externer Experten ist optional möglich, Stellungnahmeverfahren und abschließende Veröffentlichung entfallen. Es ist zu diskutieren, ob sich das IQWIG durch diese potentiell rein interne Bearbeitung angreifbar macht. Andererseits ist dies ein pragmatisches Herangehen, das angesichts der Vorläufigkeit der Einschätzungen und im Hinblick auf eine zügige Bearbeitung sinnvoll ist.

Im Kapitel zur Kostennutzenbewertung (Kapitel 4) scheint der Absatz 4.1.6. verzichtbar. Wie realistisch die intangiblen Kosten zu bewerten sind, ist fraglich (S. 94, D). Es sollte darauf hingewiesen werden, dass GKV-Daten, insbesondere wenn nur Daten einer einzelnen, möglicherweise auch nur regionalen GKV einbezogen werden, nicht repräsentativ für die Gesamtbevölkerung sind und der Ausschluss von PKV-Versicherten in jedem Fall zu bedenken ist (S. 97, S. 99). Auf die richtige Schreibweise von QALY sollte geachtet werden (Abkürzungsverzeichnis!).

Begrüßenswert ist die Aufnahme der ISOQOL Reportingstandards für patientenberichtete Endpunkte in die Liste der wichtigen Stellungnahmen (S. 175).

Sehr problematisch sieht das DNVF allerdings die Streichung des Zielkriteriums der Patientenzufriedenheit bei der Definition des patientenrelevanten Nutzens bzw. Schadens (z.B. S. 39, Kapitel 3.1.1). Zwar stimmen wir zu, dass es teilweise berechtigte Kritik an der Nutzung von Patientenzufriedenheit als Qualitätsindikator und damit auch als Maß für die Bewertung von Maßnahmen gibt: Deckeneffekte, hohe Subjektivität, erhebliche Abhängigkeit von nicht qualitätsrelevanten Faktoren. Einige Kritikpunkte lassen sich aber durch die Erfassung von Patientenerfahrungen anstatt der generellen Zufriedenheit umgehen. Patientenerfahrung bezieht sich dabei auf bewertungsfreie Berichte von Patienten zu Ereignissen, z.B. Information, Versorgungsorganisation etc. Folgende Punkte sprechen dringlich für die Beibehaltung bzw. Stärkung von Patientenzufriedenheit und – erfahrung bei der Bewertung medizinischer Interventionen:

1. Die Erfassung von Patientenerfahrung ermöglicht die Erhebung von Informationen, an die man durch andere Informationsquellen nicht gelangt: Menschlichkeit der Versorgung (Kommunikation, Privatsphäre, Informationsbedürfnisse etc.) und spürbare Wirksamkeit der Versorgung (Black et al. 2013, Manary et al. 2013). Diese Ereignisberichte sind weniger von Bias, Subjektivität und Erwartungen geprägt.



2. Zwar handelt es sich bei Patientenerfahrung vorwiegend um Prozessindikatoren. Es gibt jedoch viele Belege zu Zusammenhängen zwischen Patientenerfahrungen (auch Patientenzufriedenheit) und Outcomes (Leitlinienadhärenz, Mortalität, Wiedereinweisungsrate, Patientensicherheit, klinische Wirksamkeit etc.) (z. B. Doyle et al. 2013). Patientenzufriedenheit erhöht die Therapieadhärenz, von der eine hohe Outcomerelevanz angenommen wird (z. B. Hirsh et al. 2005)
3. Als Anforderungen an die Messung von Patientenerfahrungen könnte man Kriterien aufstellen, die erfüllt werden müssten, z. B. Entwicklung des Instruments nach Best Practice (inkl. Pretesting), gute psychometrische Eigenschaften, Validierung des Instrumentes, Änderungssensitivität, Kontrolle von Confoundern, Auswahl des richtigen Zeitpunktes der Erhebung (Black et al. 2013, Manary et al. 2013).
4. Bei der Übertragung von Ergebnisse aus RCTs in die Praxis ergeben sich oft Probleme, weil der Kontext der Versorgung außer Acht gelassen wurde (vgl. Pfaff et al. 2009). Die Qualität einer Behandlung ist in der Praxis eben auch abhängig von z. B. der kommunikativen Kompetenz des Personals. Patientenerfahrungen und Patientenzufriedenheit können den Kontext erfassen. Lebensqualität allein nicht. Die Ergebnisse können dazu genutzt werden, den Kontext umzugestalten, um das gewünschte Ergebnis zu erzielen.
5. Es wird postuliert, dass Patientenzufriedenheit Kosten durch Beschwerden, Zweitmeinungen und Nachuntersuchungen vermeiden kann.
6. Patientenzufriedenheit zu erreichen ist ein eigener wertvoller Endpunkt der Behandlung, wenn Ziel ist, eine hochwertige Versorgung zu erbringen, die die Bedarfe der Patienten erfüllt. Daher reicht Lebensqualität allein als Patient Reported Outcomes bei klinischen Studien nicht aus, denn sie erfasst nicht die Angemessenheit der Behandlung (Chow et al. 2009).

Es sei auch darauf hingewiesen, dass Patientenbefragungen incl. Zufriedenheit Gegenstand des aktuellen Koalitionsvertrages der Bundesregierung sind, die Versorgungszufriedenheit zunehmend in Leitlinien thematisiert und auch durch den G-BA hoch priorisiert wird. Das DNVF geht somit von einer zunehmenden Bedeutung dieses Konstruktes als Qualitätsindikator in der Bewertung medizinischer Maßnahmen aus.

Abgesehen von dieser Streichung findet sich leider keine Überarbeitung des Methodenpapiers im Papier im Hinblick auf patient reported outcomes und Lebensqualität, diese Ausführungen sind weiterhin wenig substantiell. Allgemein sei noch darauf hingewiesen, dass das DNVF aufgrund der höheren Relevanz psychosozialer (vs. gesundheitlicher) Faktoren auf die Lebensqualitätsscores die Verwendung des Suffixes „gesundheitsbezogen“ nicht mehr empfiehlt.

Es gibt einige weitere Streichungen (Hintergrundinformationen mit Begründung der jeweiligen Streichungen wären hilfreich gewesen):

S. 33: Zu vergebende Aufträge werden nicht mehr auf der Webseite des IQWiG veröffentlicht. Das ist bedauerlich im Sinne einer geringeren Transparenz und Flexibilität.

S. 35: Externe Sachverständige können nicht mehr Nichtnennung beantragen. Im Prinzip im Sinne erhöhter Transparenz begrüßenswert; inwieweit hiermit evtl. manch explizitere Stellungnahmen entfallen werden, ist zu diskutieren.



S. 174 – Die zuvor sehr formalistische Evidenzqualitätsbewertung nach Studientyp wurde erfreulicherweise abgemildert, zumal ja das GRADE System bereits integriert ist.

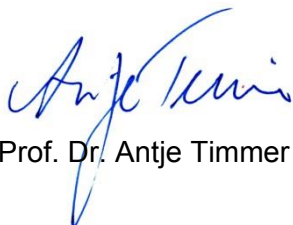
Zusammenfassend spricht sich das DNVF mit Nachdruck für die Beibehaltung bzw. Stärkung der Erhebung patientenrelevanter Endpunkte, und zwar einschließlich der Patientenzufriedenheit aus. Eine Überarbeitung der Ausführungen zum Thema patientenorientierte Endpunkte steht weiterhin aus. Darüber hinaus sind wenig relevante Änderungen erkennbar, die aus Seiten des DNVF problematisch erscheinen.

Literatur

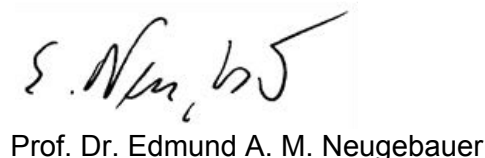
1. Black N, Jenkinson C: How can patients' views of their care enhance quality improvement? BMJ 2013; 339: 202–5.
2. Manary MP, Boulding W, Staelin R, et al.: The patient experience and health outcomes. N Engl J Med 2013; 368(3): 201–3.
3. Doyle C, Lennox L, Bell D: A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. BMJ Open 2013; 3: e001570.
4. Hirsh A, Atchison J, Berger J, et al: Patient satisfaction with treatment for chronic pain: predictors and relationship to compliance. Clin J Pain 2005;21(4): 302-10.
5. Pfaff H, Albert U, Bornemann R, et al.: Methoden für die organisationsbezogene Versorgungsforschung. Gesundheitswesen 2009; 71(11): 777–90.
6. Chow A, Mayer E, Darzi A, Athanasiou T: Patient-reported outcome measures: The importance of patient satisfaction in surgery. Surgery 2009; 146(3): 435-43.

Die Stellungnahme wurde im Auftrag des DNVF-Vorstands koordiniert von:

Prof. Dr. A. Timmer, Carl von Ossietzky Universität Oldenburg; Beiträge von Dr. L. Ansmann, Köln, Prof. A. Icks, Düsseldorf, Prof. H. Pfaff, Köln, Prof. E. Neugebauer, Köln; PD Dr. M. Klinkhammer-Schalke, Regensburg und Prof. M. Koller, Regensburg.



Prof. Dr. Antje Timmer



Prof. Dr. Edmund A. M. Neugebauer

A.1.20 – Ecker + Ecker GmbH

Ecker + Ecker GmbH, Stellungnahme zu den Allgemeinen Methoden – Entwurf 4.2 vom 18.06.2014

Vorgelegt von: Dr. Thomas Ecker am 05.08.2014

Einleitung

Ecker + Ecker GmbH unterstützt pharmazeutische Unternehmen regelmäßig im Verfahren der frühen Nutzenbewertung. Aufbauend auf den damit verbundenen Erfahrungen möchten wir heute zu dem vom IQWiG vorgelegten Allgemeinen Methoden – Entwurf 4.2 vom 18.06.2014 wie folgt Stellung nehmen.

Hintergrund

Die „Allgemeinen Methoden“ haben die Aufgabe, die gesetzlichen und wissenschaftlichen Grundlagen des IQWiG zu erläutern (Allgemeine Methoden, S iii). Die Aufgaben des Instituts werden hier ebenso dargelegt wie die wissenschaftlichen Werkzeuge, die für die Bearbeitung der Institutsprodukte verwendet werden.

Die Arbeit des IQWiG gliedert sich in Produkte. Im Rahmen der frühen Nutzenbewertung erstellt das IQWiG im Auftrag des G-BA primär das Produkt der Dossierbewertung (Allgemeine Methoden, S. 20). Zwei Formen der Dossierbewertung sind zu unterscheiden (siehe Tabelle 1):

- umfassende Bewertung (Bewertungsaspekt 1 – 6)
- Bewertung von Anzahl der Patienten und Kosten der Therapie (Bewertungsaspekt 3 – 4), sog. Bewertung als Orphan drug.

Tabelle 1: Bewertungsaspekte unterschiedlicher Nutzenbewertungen durch das IQWiG

	Bewertungsaspekt	Beauftragung bei umfassender Bewertung	Beauftragung bei einer Bewertung als Orphan drug	In Allgemeinen Methoden – Entwurf 4.2 beschrieben
1.	Zugelassene Anwendungsgebiete	X		nein
2.	Medizinischer Nutzen	X		nein
3.	Medizinischer Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie	X		ja
4.	Anzahl der Patienten in den für die Behandlung infrage kommenden Patientengruppen	X	X	nein
5.	Kosten der Therapie für die gesetzliche Krankenversicherung	X	X	nein
6.	Anforderungen an eine qualitätsgesicherte Anwendung	X		nein

Problemstellung

Der spezifische Prozess der Dossierbewertung ist in Abschnitt 2.1.3 der Allgemeinen Methoden beschrieben. Methodische Fragen der Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V, also der frühen Nutzenbewertung, sind in Abschnitt 3.3 der Allgemeinen Methoden beschrieben. Inhaltlich werden in diesem Abschnitt 3.3 im wesentlichen Fragen zur Operationalisierung der Feststellung des Ausmaßes des Zusatznutzens beschrieben, also Bewertungsaspekt 3 (siehe Tabelle 1).

Aussagen zur Methodik der anderen in Tabelle 1 genannten Bewertungsaspekte fehlen in Abschnitt 3.3. Sie finden sich auch nicht in einem anderen Abschnitt der Allgemeinen Methoden.

Trotz fehlender methodischer Vorgaben macht das IQWiG in seinem Produkt „Dossierbewertung“ Aussagen auch zu den anderen Bewertungsaspekten.

Dieses, bislang abstrakt geschilderte Problem soll im folgenden beispielhaft an den letzten drei jeweils am 01.8.2014 veröffentlichten Dossierbewertungen des IQWiG illustriert werden:

Beispiele zu Bewertungsaspekt 1 (Zugelassene Anwendungsgebiete)

Tabelle 2: Beispiele zu Bewertungsaspekt 1 (jeweils veröffentlicht am 01.08.2014)

Dossierbewertung	Angabe pharmazeutisches Unternehmen	Angabe IQWiG	Bewertung IQWiG
Insulin degludec ^a	Bei Diabetes mellitus Typ 1 muss Tresiba [®] mit kurz/schnell wirkendem Insulin kombiniert werden, um den mahlzeitenbezogenen Insulinbedarf zu decken.	Behandlung des Diabetes mellitus bei Erwachsenen. Das vorliegende Bewertungsmodul bezieht sich auf Diabetes mellitus Typ 1.	keine
Riociguat ^b	Adempas ist indiziert für die Behandlung erwachsener Patienten der WHO Funktionsklassen (FK) II bis III mit - inoperabler CTEPH, - persistierender oder rezidivierender CTEPH nach chirurgischer Behandlung, zur Verbesserung der körperlichen Leistungsfähigkeit.	keine	keine
Dimethylfumarat ^c	Behandlung von erwachsenen Patienten mit schubförmig remittierender Multipler Sklerose	Dimethylfumarat wird zur Behandlung von erwachsenen Patienten mit schubförmig remittierender Multipler Sklerose angewendet	keine
^a Quelle: Dossier des Herstellers (Abschnitt 1.3); IQWiG-Verfahren A14-13; Angaben beispielhaft aus Bewertungsmodul I ^b Quelle: Dossier des Herstellers (Abschnitt 1.3); IQWiG-Verfahren G14-02 (keine Aussage, da nicht Gegenstand der Beauftragung des IQWiG) ^c Quelle: Dossier des Herstellers (Abschnitt 1.3); IQWiG-Verfahren A14-14 (S. 39)			

In zwei der hier dargestellten Beispiele zu Dossierbewertungen trifft das IQWiG zum Bewertungsaspekt 1 (Zugelassene Anwendungsgebiete) inhaltliche Aussagen. Diese Aussagen weichen von denen der jeweiligen Hersteller ab. Eine Bewertung nimmt das IQWiG jedoch nicht vor. Die Ursache für die unterschiedlichen Aussagen werden vom IQWiG nicht diskutiert. Methodische Grundlagen des IQWiG für das Vorgehen des IQWiG werden nicht gegeben. Die hier bestehenden Unterschiede würden aber eine Feststellung und Bewertung erfordern. Der methodische Maßstab hierfür ist in den Allgemeinen Methoden jedoch nicht erläutert.

Beispiele zu Bewertungsaspekt 2 (Medizinischer Nutzen)

Tabelle 3: Beispiele zu Bewertungsaspekt 2 (jeweils veröffentlicht am 01.08.2014)

Dossierbewertung	Angabe pharmazeutisches Unternehmen	Angabe IQWiG	Bewertung IQWiG
Insulin degludec ^a	Insulin degludec führte sowohl als Monotherapie als auch in Kombination mit anderen Antidiabetika (OAD und/oder kurz wirkendes Insulin) zu einer Senkung des HbA1c-Wertes.	keine	keine
Riociguat ^b	Riociguat zeigte in CHEST-1 statistisch signifikante und klinisch relevante Vorteile (Angaben jeweils für die ITT-Gesamtpopulation)	keine	keine
Dimethylfumarat ^c	Die Ergebnisse zum medizinischen Nutzen (Tabelle 1-A) basieren auf der Meta-Analyse der zwei DMF-Zulassungsstudien 109MS301 und 109MS302 für DMF vs. Placebo.	keine	keine
^a Quelle: Dossier des Herstellers (Abschnitt 1.5); IQWiG-Verfahren A14-13; Angaben beispielhaft aus Bewertungsmodul I ^b Quelle: Dossier des Herstellers (Abschnitt 1.5); IQWiG-Verfahren G14-02 (keine Aussage, da nicht Gegenstand der Beauftragung des IQWiG) ^c Quelle: Dossier des Herstellers (Abschnitt 1.5); IQWiG-Verfahren A14-14			

In keinem hier dargestellten Beispiel zu Dossierbewertungen trifft das IQWiG zum Bewertungsaspekt 2 (Medizinischer Nutzen) inhaltliche Aussagen. Methodische Grundlagen für das Vorgehen des IQWiG werden in den Allgemeinen Methoden jedoch nicht gegeben.

Beispiele zu Bewertungsaspekt 4 (Anzahl der Patienten in den für die Behandlung infrage kommenden Patientengruppen)

Tabelle 4: Beispiele zu Bewertungsaspekt 4 (jeweils veröffentlicht am 01.08.2014)

Dossierbewertung	Angabe pharmazeutisches Unternehmen	Angabe IQWiG	Bewertung IQWiG
Insulin degludec ^a	161.737	keine	„nicht nachvollziehbar“
Riociguat ^b	3.094	924 – 5.458	„mit Unsicherheit behaftet und tendenziell unterschätzt“
Dimethylfumarat ^c	87.771 – 107.275	keine	„plausibel“
^a IQWiG-Verfahren A14-13; Angaben beispielhaft aus Bewertungsmodul I (S. I.20) ^b IQWiG-Verfahren G14-02; Angaben beispielhaft aus Anwendungsgebiet A (S. 14) ^c IQWiG-Verfahren A14-14 (S. 39)			

In jeder der hier dargestellten Beispiele zu Dossierbewertungen gibt das IQWiG zum Bewertungsaspekt 3 (Anzahl der Patienten in den für die Behandlung infrage kommenden Patientengruppen) eine Bewertung ab. Methodische Grundlagen des IQWiG, wieso Angaben des Herstellers „nicht nachvollziehbar“, „mit Unsicherheit behaftet“ oder „plausibel“ sein sollen, bleiben aber unklar. Die hier zitierten Bewertungen erfordern einen methodischen Maßstab, der in den Allgemeinen Methoden jedoch nicht erläutert ist.

Beispiele zu Bewertungsaspekt 5 (Kosten der Therapie für die gesetzliche Krankenversicherung)

Tabelle 5: Beispiele zu Bewertungsaspekt 5 – (jeweils veröffentlicht am 01.08.2014)

Dossierbewertung	Angabe pharmazeutisches Unternehmen (zweckmäßige Vergleichstherapie)	Kostenangabe IQWiG (zweckmäßige Vergleichstherapie)	Bewertung IQWiG
Insulin degludec ^a	1.410,07 € bis 1.493,26 €	keine	„etwas zu hoch bewertet“
Riociguat ^b			
Dimethylfumarat ^c	22.124,70 €	keine	„leichte Überschätzung“
^a IQWiG-Verfahren A14-13; Angaben beispielhaft aus Bewertungsmodul I (S. I.21) ^b IQWiG-Verfahren G14-02; keine Angaben zur zweckmäßigen Vergleichstherapie, da Orphan Drug ^c IQWiG-Verfahren A14-14 (S. 40)			

In zwei der hier dargestellten Beispiele zu Dossierbewertungen gibt das IQWiG zum Bewertungsaspekt 4 (Kosten der Therapie für die gesetzliche Krankenversicherung) eine Bewertung ab. Methodische Grundlagen des IQWiG, wieso Angaben des Herstellers „etwas zu hoch bewertet“ oder „leichte Überschätzung“ sein sollen, bleiben aber unklar. Die hier zitierten Bewertungen erfordern einen methodischen Maßstab, der in den Allgemeinen Methoden jedoch nicht erläutert ist.

Beispiele zu Bewertungsaspekt 6 (Anforderungen an eine qualitätsgesicherte Anwendung)

Tabelle 6: Beispiele zu Bewertungsaspekt 6 (jeweils veröffentlicht am 01.08.2014)

Dossierbewertung	Angabe pharmazeutisches Unternehmen	Angabe IQWiG	Bewertung IQWiG
Insulin degludec ^a	Anforderungen, die über die oben beschriebenen und in der Fachinformation, Etikettierung, Patienteninformation, dem Risk-Management-Plan oder im Annex des EPAR (European Public Assessment Report) beschrieben sind, hinausgehen, bestehen nicht.	Bei der Beschreibung zur Dosierung fehlt die Angabe der Fachinformation aus Abschnitt 4.2, dass bei Diabetes mellitus Typ 1 Insulin degludec mit kurz / schnell wirkendem Insulin kombiniert werden muss, um den mahlzeitenbezogenen Insulinbedarf zu decken. Die Hinweise auf Inkompatibilitäten (Abschnitt 6.2 der Fachinformation) und Aufbewahrung (Abschnitte 6.3 und 6.4 der Fachinformation) fehlen.	Darüber hinaus sind keine Mängel in der Darstellung der qualitätsgesicherten Anwendung im Dossier festzustellen.
Riociguat ^b	Zusätzliche Maßnahmen zur Risikominimierung sind nicht beabsichtigt.	keine	keine
Dimethylfumarat ^c	Gemäß Fach- und Gebrauchsinformation muss die Behandlung unter Aufsicht eines Arztes eingeleitet werden, der Erfahrung in der Behandlung der Erkrankung besitzt.	In Abschnitt 3.4.1 des Dossiers zitiert der pU die relevanten Abschnitte (4.2 bis 4.7) aus der Fachinformation von Dimethylfumarat [35].	Die Angaben sind vollständig.
^a Quelle: Dossier des Herstellers (Abschnitt 1.8); IQWiG-Verfahren A14-13; Angaben beispielhaft aus Bewertungsmodul I ^b Quelle: Dossier des Herstellers (Abschnitt 1.8); IQWiG-Verfahren G14-02 (keine Aussage, da nicht Gegenstand der Beauftragung des IQWiG) ^c Quelle: Dossier des Herstellers (Abschnitt 1.8); IQWiG-Verfahren A14-14 (S. 39)			

In zwei der hier dargestellten Beispiele zu Dossierbewertungen nimmt das IQWiG zum Bewertungsaspekt 6 (qualitätsgesicherte Anwendung) eine Bewertung vor. Methodische Grundlagen für das Vorgehen des IQWiG werden in den Allgemeinen Methoden jedoch nicht gegeben.

Ergebnis

Der IQWiG nimmt in seiner Dossierbewertung zu bis zu sechs Bewertungsaspekten Stellung. Methodische Grundlagen der Dossierbewertung werden aber nur für den dritten Bewertungsaspekt (medizinischer Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie) in den Allgemeinen Methoden erläutert. Methodische Grundlagen für die anderen fünf Bewertungsaspekte werden in den Allgemeinen Methoden in ihrem vorliegenden Entwurf nicht gegeben, obwohl das IQWiG auch zu allen anderen Bewertungsaspekten Aussagen trifft und zumindest teilweise auch Bewertungen vornimmt.

Konsequenz der nicht transparenten methodischen Grundlage ist es, dass

- wertende Aussagen des IQWiG nicht oder nur unvollständig nachvollziehbar sind und
- eine methodische Konsistenz über unterschiedliche Dossierbewertungen nur schwer gewährleistet werden kann.

Die in den Allgemeinen Methoden dargestellten wissenschaftlichen Grundlagen des IQWiG bleiben somit unvollständig.

Empfehlung

Vor dem Hintergrund der anhand dieser Stellungnahme beschriebenen Defizite wäre es wünschenswert, wenn das IQWiG insoweit seine Allgemeinen Methoden ergänzt, damit die im Rahmen der Dossierbewertung vorgenommenen Bewertungen in allen Bewertungsaspekten nachvollziehbar werden.

A.1.21 – GlaxoSmithKline GmbH & Co. KG (GSK)

Stellungnahme von GlaxoSmithKline GmbH & Co. KG (GSK) zum Entwurf des IQWiG „Allgemeinen Methoden Version 4.2“ vom 18.06.2014

GSK begrüßt die kontinuierliche Weiterentwicklung der Allgemeinen Methoden des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Dies ermöglicht eine Methodengrundlage, die auf aktuellen Entwicklungen im Bereich der evidenzbasierten Medizin sowie der Gesundheitsökonomie beruht.

Die transparente Darstellung der Änderungen durch das IQWiG, insbesondere die nun erstmals vorgenommene Markierung der geänderten bzw. gelöschten Textteile, vereinfacht das Nachvollziehen konkreter Änderungen im Methodenpapier. An einigen Stellen wird auch vom IQWiG begründet, warum entsprechende Änderungen vorgenommen wurden. Allerdings wurde dieses Prinzip leider nicht durchgehend angewendet, wodurch sich zahlreiche Interpretationsspielräume ergeben.

Deshalb appelliert GSK im Rahmen des Transparenzgebotes an das IQWiG, eine Rationale für alle vorgenommenen inhaltlichen Änderungen darzulegen. Des Weiteren wäre es wünschenswert, eine transparente Veröffentlichung einer solchen Aktualisierung der Allgemeinen Methoden vorzunehmen, bspw. durch Anschreiben relevanter Patienten- und Industrieorganisationen und anderen Stakeholdern, um ein ordnungsgemäßes und umfassendes Stellungnahmeverfahren zu gewährleisten.

Inhaltlich wäre es von Seiten GSKs wünschenswert, neben der kontinuierlichen Anpassung der Allgemeinen Methoden an aktuelle Änderungen auch den Aspekt der soziologisch-medizinischen und politischen Relevanz, sowie die Auswirkungen auf Forschung & Entwicklung zu beachten und eine Konzentration auf die wesentlichen Aspekte vorzunehmen. Basierend darauf wäre insbesondere eine Verschlankung der methodischen Anforderungen - und somit indirekt des Dossiers zur frühen Nutzenbewertung gemäß § 35a SGB V - zu begrüßen, bspw. durch das Wegstreichen zahlreicher Subgruppenanalysen, die nur selten entscheidungsrelevant sind.

Im Detail möchte GSK zu folgenden Punkten Stellung nehmen:

1. Stellungnahme zu den Änderungen zum Thema Patientenzufriedenheit in den Abschnitten 3.1.1 und 3.2.4

In Abschnitt 3.1.1 des Entwurfs der Allgemeinen Methoden Version 4.2 (¹Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2014) definiert das IQWiG den patientenrelevanten medizinischen Nutzen bzw. Schaden einer medizinischen Intervention. In Abschnitt 3.2.4 führt das IQWiG weitere regelhaft in die Nutzenbewertung einzubeziehende, patientenberichtete Studienendpunkte (Patient Reported Outcomes, PRO) auf. Im Entwurf der Allgemeinen Methoden wurden in beiden Abschnitten die Patienten- und Behandlungszufriedenheit als ergänzende Zielgrößen zu Mortalität, Morbidität und gesundheitsbezogener Lebensqualität gestrichen, ohne dazu eine Rationale zu liefern. Das IQWiG schließt somit künftig gleich zwei PRO aus seiner Betrachtung aus, d.h. solche Studienendpunkte, bei denen der Patient selbst zu Wort kommt und eine Bewertung entsprechend seiner Bedürfnisse und der für ihn maßgeblichen Faktoren der Behandlung und des Therapieerfolgs abgeben kann.

Entgegen dieser Entscheidung legt das IQWiG in Abschnitt 1.2.1 die allgemeine Anerkennung der Wichtigkeit von PRO dar: „Im Idealfall basieren Entscheidungen außer auf der Evidenz auch auf dem klinischen Zustand und den Umständen, in denen sich eine Person befindet, und auf Ihren Präferenzen und Handlungen“ (¹Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2014). Des Weiteren spielen die Bedürfnisse der Patientinnen und Patienten im Rahmen der praktischen Anwendung der evidenzbasierten Medizin ebenfalls eine wichtige Rolle.

In der Tat stellt die Patienten- und Behandlungszufriedenheit ein recht junges Forschungsfeld dar. Die Beantwortung der Frage „Was macht eine gute Behandlung aus?“ ist durchaus komplex. Jedoch handelt es sich dabei um eine der zentralsten Fragen im Rahmen der Entscheidung für oder gegen eine Behandlung. Die Patientenzufriedenheit gilt hierbei als wichtiger Indikator für die Messung der medizinischen Versorgungsqualität aus Sicht der Patienten (²Donabedian 1980). Weiterhin liefert sie wichtige, komplementäre Informationen zur Vervollständigung der Qualitätsmessung.

Um ein möglichst vollständiges Bild von Nutzen und Schaden einer medizinischen Intervention zu erhalten, sollten in die Bewertung neben objektiven Indikatoren auch subjektive Zielgrößen, wie Lebensqualität und Patientenzufriedenheit, einbezogen werden.

Subjektive Zielgrößen wie die Patienten- bzw. Behandlungszufriedenheit sind mittlerweile gut messbare, international anerkannte Instrumente und finden immer häufiger Anwendung in klinischen Studien und auch im Rahmen des Authorisierungsprozesses der EMA (European Medicines Agency) (³Szende et al. 2005). Insbesondere finden der Diabetes Treatment Satisfaction Questionnaire (⁴Clare Bradley 1993), der Asthma Treatment Satisfaction Measure (⁵Martin et al. 2009) und der Treatment Satisfaction Questionnaire for Medication (⁶Atkinson et al. 2004) häufig Anwendung.

Abweichend davon plant das IQWiG nun Daten zur Behandlungs- und Patientenzufriedenheit, die von pharmazeutischen Unternehmen vorgelegt werden, künftig nicht mehr zu bewerten. Dieses Vorgehen bedeutet aus Sicht von GSK einen erheblichen Ausschluss patientenrelevanter Evidenz und steht im Widerspruch zum offiziellen Bestreben des IQWiG, Ergebnisse der aktuellen wissenschaftlichen Forschung in ihre Therapieempfehlungen mit einzubeziehen. Dieses Vorgehen widerspricht dem Grundsatz der evidenzbasierten Medizin, die bestverfügbare Evidenz zu

**Stellungnahme von GlaxoSmithKline (GSK) zum Entwurf des IQWiG
„Allgemeine Methoden Version 4.2“ vom 18.06.2014**

berücksichtigen. Ein Streichen der Patienten- bzw. Behandlungszufriedenheit führt weiterhin dazu, dass der Versorgungskontext und die Alltagsnähe aus Sicht des Patienten immer mehr in den Hintergrund rücken. Die Patientenperspektive wird somit durch die Perspektive der Nutzenbewerter ersetzt (⁷Häussler et al. 2014). Da die Patientenperspektive im deutschen Gesundheitswesen jedoch an Bedeutung gewinnt und zunehmend eine patientenorientierte Gesundheitsversorgung gefordert wird, erscheint die Überarbeitung der Allgemeinen Methoden des IQWiG aus Sicht von GSK somit nicht nachvollziehbar. Weder zur Streichung der Patienten-, noch der Behandlungszufriedenheit findet sich eine Begründung seitens des IQWiG, die diese Entscheidung transparent und nachvollziehbar machen.

2. Stellungnahme zu der Ergänzung zum Umgang mit unaufgefordert übermittelten Daten in Kapitel 7 (vorher Kapitel 6)

Das IQWiG plant, unaufgefordert übermittelte Daten in Zukunft inhaltlich nicht zu berücksichtigen und begründet dies mit dem Risiko eines Selektions-Bias.

GSK stimmt dem IQWiG dahingehend zu, dass es zu einem Selektions-Bias kommen könnte. Allerdings ist GSK der Auffassung, dass eine pauschale Nichtberücksichtigung von vorhandener Evidenz mit dem Grundprinzip der evidenzbasierten Medizin, die gesamte vorhandene Evidenz zu berücksichtigen, nicht vereinbar ist. Vielmehr sollte bei der Beurteilung der Evidenzgüte mögliche Verzerrungsaspekte berücksichtigt werden – analog zur Vorgehensweise innerhalb eines Dossiers zur frühen Nutzenbewertung gemäß § 35a SGB V.

3. Stellungnahme zur Überarbeitung der Ausführungen zu Gesundheitsinformationen nach Änderung der Formate infolge des Relaunches der Website Gesundheitsinformation.de vom 13.02.2014 in Abschnitt 2.1.7 und Kapitel 6 (vorher Kapitel 5)

In Abschnitt 2.1.7, Seite 30, wird auf die Einbeziehung von externen Reviews im Rahmen der Erstellung von Gesundheitsinformationen eingegangen. Diese aufbereiteten Informationen (in der Version 4.1 noch als „Produkte“ bezeichnet) werden vom IQWiG primär über die Website www.gesundheitsinformation.de der Öffentlichkeit zugänglich gemacht. Es werden hierbei, wie auch an anderen Stellen zu weiteren IQWiG-Produkten, keine Gründe bzw. Regeln für die Beteiligung bzw. das Einbeziehen von externen Reviews angeführt und eher abstrakt auf eine Beteiligung externer Sachverständiger hingewiesen. Hier wäre es wünschenswert, im Rahmen des Transparenzgebotes klare Regeln für das Einbeziehen von externen Reviews zu bestimmen und offen darzulegen.

4. Synopse

GSK begrüßt die kontinuierliche Weiterentwicklung der Allgemeinen Methoden sowie die transparente Darstellung der inhaltlichen Änderungen durch das IQWiG und sieht darin die Chance, das Verfahren der frühen Nutzenbewertung auf die für den Patienten relevanten Informationen zu fokussieren und somit insgesamt zu verschlanken.

Dieser Prozess sollte jedoch nicht losgelöst, sondern im Kontext und im offenen Dialog mit allen Beteiligten erfolgen. Im Sinne des Transparenzgebotes sieht GSK eine durchgängige Begründung sowie Kommunikation der inhaltlichen Aktualisierung der Allgemeinen Methoden geboten. In diesem Zusammenhang sollte das IQWiG alle relevanten Organisationen, wie bspw. die Gremien der Gemeinsamen Selbstverwaltung und Patientenorganisationen, identifizieren und gezielt über die Möglichkeit einer Stellungnahme informieren. Die darüber hinausgehende Einräumung eines generellen Stellungnahmerechts an interessierte Personen und Institutionen ist begrüßenswert.

Inhaltlich kritisch an der Aktualisierung erscheinen GSK insbesondere die Selektion und der Ausschluss patientenrelevanter Evidenz. Dem Grundprinzip der evidenzbasierten Medizin folgend sollte zunächst die gesamte verfügbare Evidenz in der Nutzenbewertung berücksichtigt werden. GSK spricht sich daher ausdrücklich für eine Beibehaltung der Patienten- und Behandlungszufriedenheit als patientenrelevante Zielgrößen in den Abschnitten 3.1.1 und 3.2.4 der Version 4.2 der Allgemeinen Methoden des IQWiG aus.

Um die Forschungsaktivitäten im Bereich der PRO voranzutreiben und zu einer besseren Validierung beizutragen, sollten entsprechende Daten weiterhin regelhaft durch das IQWiG bewertet und korrespondierendes Optimierungspotenzial in der Evidenzgenerierung aufgezeigt werden. Die Beibehaltung der Patienten- und Behandlungszufriedenheit als ergänzende Zielgrößen setzt aus Sicht von GSK die richtigen Anreize für eine patientenorientierte Arzneimittelversorgung.

5. Literatur

Referenzliste

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 18-6-2014, "Allgemeine Methoden - Entwurf für Version 4.2"
https://www.iqwig.de/download/IQWiG_Methoden_Entwurf-fuer-Version-4-2.pdf (eingesehen am 23-6-2014)
2. Donabedian, A. 1980, Explorations in quality assessment and monitoring: Vol. 1 The definition of quality and approaches to its assessment. Ann Arbor, Health Administration Press.
3. Szende, A., Leidy, N. K., Revicki, D. 2005, "Health-Related Quality of Life and Other Patient-Reported Outcomes in the European Centralized Drug Regulatory Process: A Review of Guidance Documents and Performed Authorizations of Medicinal Products 1995 to 2003", *Value in health*, vol. 8, no. 5, pp. 534-548.
4. Clare Bradley 1993, *Zufriedenheit mit der Diabetes-Behandlung: DTQs (Fragebogen)*, Health Psychology Research, Dept of Psychology, Royal Holloway, University of London. (eingesehen am 30-7-2014)
5. Martin, M. L., Patrick, D. L., Bushnell, D. M., Meltzer, E. O., Gutierrez, B., Parasuraman, B. 2009, "Development of the asthma treatment satisfaction measure", *Current Medical Research & Opinion*, vol. 25, no. 10, pp. 2495-2506.
6. Atkinson, M. J., Sinha, A., Hass, S. L., Colman, S. S., Kumar, R. N., Brod, M., Rowland, C. R. 2004, "Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease", *Health and quality of life outcomes*, vol. 2, no. 1, p. 12.
7. Häussler, B., Nolting, H.-D., and Wettich, A. 2014, "Wer "Patientenrelevanten Nutzen" fordert, muss nach "patient-reported outcomes" rufen"
<https://load.doccheck.com/de/digitale-zeitschriften/medizin-load/dzkgf-deutsche-zeitschrift-fuer-klinische-forschung-heft-01-2014.html/listtype/vendor> (eingesehen am 2-7-2014)

A.1.22 – Institut für Gesundheitsökonomik (IfG)

***Kommentierung zu „Allgemeine Methoden: Entwurf für Version 4.2“,
Entwurf des IQWiG vom 18.06.2014***

Mit Datum vom 18.06.2014 hat das IQWiG die Aktualisierung weiterer Abschnitte seiner Allgemeinen Methoden (Version 4.2.) zur Diskussion gestellt. Die Kommentierungsfrist endet am 07. August 2014. Das Institut für Gesundheitsökonomik (IfG) möchte die Möglichkeit nutzen, zur Weiterentwicklung der Methoden mit beizutragen und deshalb einige ausgewählte Punkte kommentieren. Grundsätzlich wird die Aktualisierung gerade der Methoden zur gesundheitsökonomischen Bewertung (Kosten-Nutzen-Bewertung) durch das IQWiG begrüßt, die an einigen Stellen Klarstellungen und methodische Verbesserungen gebracht haben.

Auf Beiträge zu methodischen Aspekten möchten wir verzichten und verweisen hierzu auf die bereits bekannten gesundheitsökonomischen Stellungnahmen zur Effizienzgrenze und die aktuellen Stellungnahmen der entsprechenden Fachgesellschaften wie der deutschen Gesellschaft für Gesundheitsökonomie. Als privates Auftrags- und Forschungsinstitut möchten wir jedoch zu zwei praktischen Umsetzungs-Punkten der Methoden Version 4.2 kurz Stellung nehmen. Diese erscheinen uns entscheidend dafür, dass die Kosten-Nutzen-Bewertung in Deutschland praktikabel umsetzbar wird:

1.) Abschnitt 4.2.5 C), S.86: „Eine elektronische Version des Modells muss zur Verfügung gestellt werden mit dem Einverständnis, dass das Modell öffentlich zugänglich gemacht wird und ggf. für zukünftige Evaluationen adaptiert werden kann. Die elektronische Modellversion muss vollständig zugänglich sein und den Reviewern sowie der Öffentlichkeit ermöglichen, alle in der Analyse eingesetzten Formeln und Beziehungen einzusehen und das Modell mit unterschiedlichen Inputdaten laufen zu lassen. Um das Review des Modells zu erleichtern, sollte der elektronischen Version eine Anwenderanleitung beigelegt werden, die beschreibt, welche Software und Hardware benötigt werden, wie Modellinputs verändert werden können, wo diese Inputs im Modell gefunden werden können,

wie das Modell ausgeführt werden kann und wie Ergebnisse extrahiert werden können.“

Im genannten Abschnitt wird für das gesundheitsökonomische Modell eine sehr weitgehende Veröffentlichungspflicht vorgegeben. Transparenz ist gerade auch in der Modellierung nicht nur wünschenswert sondern bei der Beurteilung und Bewertung obligat (1). Andererseits muss Transparenz immer auch gegen den Schutz des geistigen Eigentums der durchführenden Firmen und Personen abgewogen werden. Eine vollständige Veröffentlichungspflicht der verwendeten Modelle wäre international nach unserem Kenntnisstand für HTA Prozesse einzigartig. Und in der Begutachtungspraxis der meisten Fachzeitschriftenartikel erhält der Reviewer in der Regel nicht einmal beschränkten Zugriff auf das jeweilige Modell, auch wenn dies zur Bewertung häufig wünschenswert wäre.

Natürlich sollte und muss der bewertenden Institution, hier dem IQWiG, zur Bewertung eine vollständig transparente Version des jeweiligen Modells elektronisch vorliegen. Diese jedoch vollständig elektronisch zu veröffentlichen wirft unseres Erachtens erhebliche wettbewerbliche und Fragen des geistigen Eigentums auf. Hinsichtlich der praktischen Umsetzung und Anwendbarkeit könnte dies bedeuten, dass zunächst längere Diskussionen um die Veröffentlichung, gerade bei internationalen „Standard“ Ansätzen geführt werden müssten. Die Kosten für Modellierungen könnten initial vergleichsweise hoch sein, da jeder hierdurch erreichte Wissensvorsprung durch die komplette Veröffentlichung sofort an alle Wettbewerber weitergegeben würde. Anders als bei klinischen Studien – wo ja für die vollständigen Daten eine Kennzeichnung als Betriebs- und Geschäftsgeheimnis möglich ist- könnten Modellkomponenten direkt und von jedem für nicht nachvollziehbare Zwecke verwendet werden. Das könnte zur Folge haben, dass beispielsweise viele international bereits bestehende „Standard Krankheits-Modelle“ deshalb nicht oder nur eingeschränkt eingesetzt werden könnten. Insgesamt könnte das in unnötiger Weise die praktische Anwendung der Kosten-Nutzen-Bewertung behindern. Ob durch die mit den Modell-Veröffentlichungen zunehmend gewonnene, öffentliche Informationsbasis dieses Anfangshemmnis überkompensiert werden kann, erscheint uns fraglich.

Die entsprechende international ISPOR Task force schreibt zu der Veröffentlichungs-Frage: „To address the conflicting needs of transparency versus feasibility and intellectual property, information put in the public domain without restriction (“public documentation”) should be distinguished from information made available under agreements that protect intellectual property (“confidential documentation”). For public documentation, the nontechnical description should be available to all who ask. In addition, at their discretion, modelers can choose to make technical documentation, or a working copy of the model, publicly available.” (1) Weiterhin heißt es in der ISPOR Empfehlung: “Given the documentation’s size and technical nature, need to protect intellectual property, and because journals can gain full access to all documentation during the review process, journals should not require that it be included in the published report of an analysis.” Als “best practice” wird eine nicht-technische und technische Dokumentation gefordert, nicht jedoch die Veröffentlichung des eigentlichen Modells. (1)

Analog diesen Vorgehens und -analog wie bereits bei den klinischen Studiendaten im Prozess der frühen Nutzenbewertung praktiziert (Ablage der klinischen Studienberichte (CSRs) in Modul 5 mit ggf. Kennzeichnung als Betriebs-/Geschäftsgeheimnis)- schlagen wir vor, dass dem IQWiG (und ggf. G-BA) die jeweiligen Modelle komplett und ungeschützt vorzulegen sind. Dies könnte auch die Verwendung zur Gegenvalidierung bei mehreren Einreichungen im selben Therapiegebiet beinhalten und ermöglichen. Verpflichtend veröffentlicht werden sollte jedoch nur die Bewertung des IQWiGs, sowie die Modellstruktur und die Teile der Modelldokumentation, die keine wettbewerbliche Relevanz haben („public documentation“). Damit könnte die notwendige Balance zwischen Transparenz mit Nachvollziehbarkeit der Entscheidungen und Schutz geistigen Eigentums erreicht werden. Eine komplette elektronische Modellveröffentlichung erscheint uns nicht zielführend.

2. Abschnitt 4.9.4, S.111-112: „Interquartilsregion als Maß der Streuung für Preisverhandlungen“. [...] „Die IQR erlaubt es, unter Berücksichtigung der gesamten Unsicherheit (umgesetzt durch PSAs) einen Spielraum für mögliche Verhandlungen um Erstattungsbeträge zu eröffnen, in dessen Grenzen auch die Unsicherheit der Effektschätzer und der Kosten berücksichtigt wird.“

Der Vorschlag, aus dem Net Health Benefit (NHB) und probabilistischen Sensitivitätsanalyse (PSA) einen Erstattungsbetrag analytisch abzuleiten ist grundsätzlich interessant. Damit würde insbesondere eine Art Unter- und Obergrenze des Erstattungsbetrages analytisch definiert, was bisher in Deutschland auf der Basis einer Modellierung praktisch undenkbar war.

So wie die Methodik jedoch derzeit vorgeschlagen wird, ist ihr Nutzen fraglich. Eine besonders unsichere Datenlage würde sich in der PSA durch eine besonders große Interquartilsregion (IQR) darstellen und damit einen großen Verhandlungsspielraum ermöglichen – sofern dieser die IQR berücksichtigt? Bei einer kleinen IQR hingegen wäre der Erstattungsbetrag praktisch analytisch vordeterminiert, was der bisherigen deutschen Diskussion zur Nicht-Anwendbarkeit von „Thresholds“ und statischen Regeln (z.B. im Sinne von Zahlungsbereitschaft) widerspricht. Weiterhin kann man bei der statistischen Verteilung der PSA-Ergebnisse nicht jedenfalls von einer Normalverteilung ausgehen. Entsprechend wird sich ggf. auch der Mittelwert vom 50%-IQR unterscheiden, was weiteres Konfliktpotential birgt. Ähnlich für die gemäß Methodenentwurf 4.2 mögliche Abweichung der Darstellung der IQR. So heißt es: „Grundsätzlich kann auch die Angabe von anderen Regionen mit anderen Maßen sinnvoll sein“.

Insgesamt halten wir den vorgeschlagenen Weg deshalb für methodisch interessant, aber für nur sehr bedingt hilfreich um eine wirklich rationalere Verhandlungssituation zu erreichen. Es ist insgesamt eine Verlagerung der strittigen Punkte zwischen den Verhandlungsparteien auf methodische Aspekte der PSA zu befürchten. Dabei wird die PSA auch heute noch bei weitem nicht in allen Fach-Publikationen von ökonomischen Kosten-Nutzen bzw. -Nutzwert Analysen

überhaupt durchgeführt und berichtet (2). Zudem besitzt gerade die PSA viele Einflussfaktoren wie beispielsweise die jeweils zugrunde gelegten Verteilungen. Entsprechend sind tiefe Modellierungskennntnisse zur Interpretation und Bewertung notwendig. Dass eine Verhandlung durch ein derart komplexes Hilfsmittel erleichtert wird, halten wir für fraglich und deshalb derzeit den Vorschlag des IQWiG im deutschen Kontext für nicht zielführend.

München, 05.08.2014

Prof. Dr. Aljoscha S. Neubauer
Director Market Access & Pharmacoeconomics

IfG Institut für Gesundheitsökonomik
Direktor: Univ.-Prof. Dr. rer. pol. G. Neubauer
Nixenweg 2b; D-81739 München
Tel. Sekretariat: (089) 605198
Fax: (089) 6061187
Email: ifg@ifg-muenchen.com
Homepage: <http://www.ifg-muenchen.com>

Referenzen:

- (1) Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: A report of the ISPOR-SMDM modeling good research practices task force-4. Value Health 2012;15:843-50.
- (2) Rochau U, Schwarzer R, Jahn B, Sroczynski G, Kluibenschaedl M, Wolf D, Radich J, Brixner D, Gastl G, Siebert U. Systematic assessment of decision-analytic models for chronic myeloid leukemia. Appl Health Econ Health Policy. 2014 Apr;12(2):103-15. doi: 10.1007/s40258-013-0071-8.

A.1.23 – MArS Market Access & Pricing Strategy GmbH und AHEAD GmbH

Stellungnahme der MArS Market Access & Pricing Strategy GmbH und der AHEAD GmbH zur inhaltlichen Aktualisierungen und Ergänzungen im Entwurf der Version 4.2: Allgemeine Methoden

3.8 Potenzialbewertung

S75: „Abweichend zum Vorgehen bei Nutzenbewertungen (siehe Abschnitt 3.1.2) werden bei der Potenzialbewertung auch Surrogatendpunkte berücksichtigt, für die noch keine hinreichende Validität gezeigt werden konnte. Surrogatendpunkte sollen jedoch etabliert und plausibel sein, um hierüber ein Potenzial begründen zu können.“

Kommentar: Wir begrüßen den Vorschlag des IQWiG auch Surrogatendpunkte zur Potentialbewertung zu berücksichtigen. In diesem Zusammenhang wäre es jedoch wünschenswert die Begriffe „etabliert“ und „plausibel“ näher zu definieren um möglichen Leistungserbringern, Antragstellern, und Dienstleistern einen tieferen Einblick in die Bewertungskriterien des IQWiG zu geben, und es somit zu ermöglichen, die Darstellung und Auswahl möglicher Surrogatendpunkte im Rahmen eines Antrages zur Potentialbewertung, den noch zu definierenden Maßstäben des IQWiG anzupassen.

3.8 Potenzialbewertung

S75: „Sofern das Potenzial diagnostischer Methoden zu bewerten ist, finden auch Daten zur Testgüte Berücksichtigung. Hierbei ist die Ergebnissicherheit der zugrunde liegenden Studien zu prüfen (siehe Abschnitte 3.5 und 8.3.11). Im zweiten Schritt erfolgt eine Plausibilitätsprüfung der diagnostischen Methode hinsichtlich der vom Antragsteller postulierten Effekte bezüglich patientenrelevanter Endpunkte, d. h. möglicher direkter Effekte der Methode sowie therapeutischer Konsequenzen, über die die diagnostische Methode patientenrelevante Zielgrößen beeinflussen könnte.“

Kommentar: Wir begrüßen den Vorschlag des IQWiG die Testgüte bei diagnostischen Methoden ebenfalls als Prüfkriterium hinzuzuziehen. Im zweiten Schritt, der Plausibilitätsprüfung, wäre es aus wissenschaftlicher Sicht wünschenswert, wenn das IQWiG die Prüfkriterien hinsichtlich der Vorhersage und damit der Korrelation der Testgüte auf mögliche patientenrelevante Endpunkte offenlegen würde. Wichtig erscheinen uns in diesem Zusammenhang vor allem mögliche Korrelationsstudien und die Akzeptanz von Surrogatendpunkten bei einer möglichen Studie zur Erprobungsregelung.

4.1.3 Zeithorizont

S77: „Der Zeithorizont der Kostenbestimmung muss mit dem der Nutzenbestimmung übereinstimmen.“

Kommentar: Der „Zeithorizont der Nutzenbestimmung“ sollte in diesem Zusammenhang klarer definiert werden. Die derzeitige Formulierung lässt offen ob es sich bei dem „Zeithorizont der Nutzenbestimmung“ um den Zeithorizont der KNB handelt oder um den Zeithorizont der zugrundeliegenden klinischen Studie(n).

4.1.4 Auswahl der Komparatoren

S77: „Substanzen z. B. können auch zu Substanzklassen zusammengefasst werden, wenn ausreichende Homogenität vorliegt.“

Kommentar: In diesem Zusammenhang sollte der Begriff „ausreichende Homogenität“ näher definiert werden. Ein möglicher Ansatz wäre dabei die Anwendung etablierter Verfahren zur Homogenitätsprüfungen bei Metanalysen.

4.1.6 Modell

S78: „Aus diesen Gründen ist die Modellierung der Kosten einer Intervention eine wesentliche Komponente der KNB. Ebenso kann in einer KNB der Nutzen modelliert werden, wenn einer KNB ein längerer als der studienbelegte Zeithorizont zugrunde gelegt werden soll.“

Kommentar: Hier sollte im Hinblick auf die Spezifizierung der „Modellierung der Kosten“ und der „Modellierung des Nutzens“ auf die Abschnitte „4.4 Kosten“ und „4.3 Nutzen“ verwiesen werden.

Interpretation der Ergebnisse (4.1.9) sowie Besonderheiten der Kosten-Nutzenbewertung nach § 35b SGB V (4.9.2-4)

Als Mitglieder der DGGÖ unterstützen wir die aktuelle Stellungnahme der Deutschen Gesellschaft für Gesundheitsökonomie die in Ihrer Stellungnahme die Limitationen des Ansatzes der Effizienzgrenze wiederholt aufzeigt.

4.3.3 Maß des Gesamtnutzens

S. 87 “Wenn der G-BA für eine KNB nach § 35b Absatz 1 Satz 2 SGB V das Maß des Gesamtnutzens festlegt (siehe Abschnitt 4.9), werden ein entsprechendes Instrument und ggf. die dafür festgelegten Erhebungsmethoden oder eine schon festgelegte Gewichtung von Endpunkten nach den Vorgaben im Auftrag eingesetzt. Die Ergebnisse sollten dem Entscheidungsträger zusammen mit der Endpunktgewichtung zur Verfügung gestellt werden. Hierdurch ergibt sich für den Entscheidungsträger die Option, einen aus mehreren Zusatznutzenbereinigten Erstattungspreisen gewichteten Erstattungsbetrag zu verhandeln.”

Kommentar: Nach einem kurzen Überblick zu den international bekannten Möglichkeiten das Maß des Gesamtnutzens zu bestimmen, zieht das IQWiG sich darauf zurück, dass der G-BA die „festgelegten Erhebungsmethoden oder eine schon festgelegte Gewichtung von Endpunkten nach den Vorgaben im Auftrag“ festlegt. Es wäre in diesem Zusammenhang wünschenswert wenn schon im Methodenpapier des IQWiG ein konkreteres methodisches Vorgehen inklusive einer möglichen praktischen Anwendung beschrieben werden würde. Gerade im Hinblick auf eine praktische Anwendung bleiben hier viele Fragen ungeklärt: z.B. wie bei einer KNB Nutzendaten aus unterschiedlichen Studien mit unterschiedlichen primären Endpunkten genutzt werden sollen, welcher Endpunkt in einem solchen Fall als patientenrelevant angesehen werden kann, wie eine valide Interpretation durchgeführt werden kann und wie eine transparente Bewertung erfolgen könnte?

GKV Versichertenperspektive

S95: „Grundsätzlich werden dem Prinzip der Effizienz folgend die günstigsten Vertreter eines Wirkstoffs bzw. einer Wirkstoffklasse ausgewählt.“

Kommentar: Die Auswahl des günstigsten Preises reflektiert nicht die tatsächlichen durchschnittlichen Preise die derzeit von den gesetzlichen Krankenkassen erstattet werden. In der Regel lässt sich ein Durchschnittspreis auf Basis der tatsächlichen Verordnungen verschiedener Präparate eines Wirkstoffs ableiten, was dann die tatsächliche Versorgungspraxis abbilden würde. Gerade im Hinblick auf die GKV Perspektive wäre eine derartige Abbildung der tatsächlichen Versorgungspraxis und damit der tatsächlichen Erstattungspreise erstrebenswert. Der Einfluss der im Methodenpapier als gängige Praxis postulierte „Mindestpreisannahme“ könnte dann in Form von Sensitivitätsanalysen untersucht werden.

A.1.24 – MSD Sharp & Dohme GmbH



MSD SHARP & DOHME GMBH · Postfach 12 02 · 85530 Haar
Institut für Qualität und Wirtschaftlichkeit
im Gesundheitswesen (IQWiG)
Herrn Prof. Dr. med. Jürgen Windeler
Im Mediapark 8 (KölnTurm)
50670 Köln

06. August 2014

**Stellungnahme der MSD SHARP & DOHME GmbH
Entwurf der Version 4.2 der Allgemeinen Methoden vom 18.6.2014**

Sehr geehrter Herr Professor Windeler,

wir begrüßen ausdrücklich die intensive Diskussionskultur um die Methoden und ihre kontinuierliche Weiterentwicklung.

Wir haben eine grundlegende sowie zwei kleinere Anmerkungen, welche wir im Folgenden ausführen.

**1. Keine Rechtsgrundlage für zwingend indikationsspezifische
Wirtschaftlichkeitskriterien**

Das IQWiG sucht für die von ihm vertretenen indikationsspezifischen Kriterien einer Kosten-Nutzen-Bewertung (KNB) von Arzneimitteln in seinem Entwurf der Version 4.2 der Allgemeinen Methoden eine Basis in § 12 SGB V. Zur Unterstützung zitiert das IQWiG das Bundesministerium für Gesundheit (BMG) aus dem Jahre 2008: „Maßgebend für die Verordnungsentscheidung des Arztes kann nur sein, ob das Arzneimittel zur Erreichung des Therapieziels, d. h. im jeweiligen Anwendungsgebiet, zweckmäßig und wirtschaftlich ist (§ 12 SGB V).“ Das IQWiG zieht daraus den Schluss, „daher muss die KNB einen Vergleich der Effizienz von Interventionen in einem bestimmten Indikationsgebiet ermöglichen“. Die KNB müsse also indikationsspezifischen Wirtschaftlichkeitskriterien unterworfen werden.

Dieser Schluss des IQWiG trifft nicht zu. Weder § 12 SGB V noch das Zitat des BMG stützen die These, indikationsspezifische Wirtschaftlichkeitskriterien seien rechtlich geboten:

- § 12 SGB V normiert das Wirtschaftlichkeitsgebot: Leistungen müssen ausreichend, zweckmäßig und wirtschaftlich sein; sie dürfen das Maß des Notwendigen nicht überschreiten. Leistungen, die nicht notwendig oder unwirtschaftlich sind, können



Versicherte nicht beanspruchen, dürfen die Leistungserbringer nicht bewirken und die Krankenkassen nicht bewilligen. Das Wirtschaftlichkeitsgebot ist damit zunächst grundlegend methodenoffen formuliert und auch intendiert.

- Das vom IQWiG herangezogene Zitat des BMG bezieht sich zunächst nur auf die Nutzenbewertung, nicht auf die Kosten-Nutzen-Bewertung. In einem vorhergehenden Absatz differenziert das BMG diese beiden Bewertungen durch das IQWiG auch ausdrücklich. Anders als im Methodenpapier verlinkt, hat das BMG das Zitat von seiner Homepage im Übrigen eliminiert.

Inhaltlich stellt das BMG in dem Zitat lediglich zutreffend fest, dass das Wirtschaftlichkeitsgebot für die konkrete Verordnungsentscheidung des Arztes auch bei Arzneimitteln maßgebend ist, und zwar mit Blick auf das konkrete Therapieziel, also im jeweiligen Anwendungsgebiet. Darin erschöpft sich die Aussage des BMG. Damit beschreibt das BMG den Kern von § 12 SGB V: das Gebot der Wirtschaftlichkeit jeder konkreten Leistung an einen Versicherten der Gesetzlichen Krankenversicherung. Über die Ausgestaltung genereller, über die konkrete Leistung hinausgehender Instrumente der Leistungssteuerung, enthält das BMG-Zitat keine Aussage. Die These, „daher“ müsse die KNB zwingend einen Vergleich der Effizienz von Interventionen in einem bestimmten Indikationsgebiet ermöglichen, suggeriert einen Zusammenhang, der nicht besteht.

Auch jenseits von § 12 stützt das SGB V in keiner Norm die These, eine Kosten-Nutzen-Bewertung müsse auf bestimmte indikationsspezifische Wirtschaftlichkeitskriterien begrenzt werden. Das wurde bereits in der Vergangenheit von der Rechtswissenschaft in der Diskussion der KNB-Methoden festgestellt. Demnach erkennt das Gesetz – gerade § 35b SGB V – die Pluralität gesundheitsökonomischer Standards an, die das IQWiG nicht „durch eine Methodik sui generis zu revidieren“ befugt sei (Gassner, Der Vorschlag des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen zur Methodik einer Kosten-Nutzen-Bewertung auf dem rechtlichen Prüfstand. Pharm. Ind. 70, Nr. 6, 701-707 (2008)). Auch andere Normen lassen sich nicht für indikationsspezifische Wirtschaftlichkeitskriterien in der Bewertung durch das IQWiG heranziehen. So „ist nach § 92 Abs. 2 Satz 2 SGB V zwar die Zusammenstellung der Arzneimittel in den Arzneimittel-Richtlinien nach Indikationsgebieten und Stoffgruppen zu gliedern. Wie die KNB vorzunehmen ist, regelt diese Vorschrift hingegen gerade nicht. § 92 Abs. 1 Satz 1 Halbs. 3 SGB V trägt zur Frage der indikationsspezifischen Methodenwahl ebenfalls nichts bei. Eine Aussage des Inhalts, dass die KNB für Arzneimittel auf indikationsspezifische Wirtschaftlichkeitskriterien angewiesen sei, lässt sich dieser Bestimmung schlechterdings nicht entnehmen. § 92 Abs. 1 Satz 1 Halbs. 3 SGB V konkretisiert vielmehr – abweichend von § 35b SGB V – den Wirtschaftlichkeitsgrundsatz der GKV in der Weise, dass bei zwei Behandlungsformen mit vergleichbarem Nutzen die

unwirtschaftlichere eingeschränkt bzw. ausgeschlossen werden kann. Die entscheidende Frage, wie die Wirtschaftlichkeit zu bemessen ist, lässt diese Vorschrift aber gerade offen (Gassner, a.a.O. Pharm. Ind. 70 (2008), S. 704).

Anders als § 92 Abs. 1 SGB V bezieht sich die gesetzliche Regelung der KNB in § 35b SGB V nicht einmal auf die Elemente des Wirtschaftlichkeitsgebotes. Der vom IQWiG vorgeschlagene Schritt, indikationsspezifische Wirtschaftlichkeitskriterien anzulegen, findet im § 35b und im gesamten SGB V demnach keine zwingende Stütze.

Vielmehr führt eine rein indikationsspezifische Bewertung, wie sie das Methodenpapier des IQWiG vorschlägt, zu massiven Gleichheitsproblemen (*Huster, Die Methodik der Kosten-Nutzen-Bewertung in der Gesetzlichen Krankenversicherung - Analyse der rechtlichen Vorgaben, Medizinrecht 2010: 234–240, S. 239*), weil diese Herangehensweise nicht sicherstellt, dass sich Kosten-Nutzen-Bewertungen und damit darauf basierende Erstattungsbeträge in unterschiedlichen Indikationsgebieten **an vergleichbaren Kriterien orientieren**. Eine indikationsspezifische Bewertung lässt es zu, dass in zwei Indikationen für den gleichen Zusatznutzen unterschiedliche Bewertungen erfolgen, da kein indikationsübergreifender Abgleich erfolgt. Rechtlich besteht die Gefahr, „dass Entscheidungen, die nicht auf indikationsübergreifenden Informationen beruhen, inkonsistent werden und zu Ungerechtigkeiten gegenüber betroffenen Patientengruppen führen sowie auch das Grundrecht der Hersteller auf faire Teilhabe am Wettbewerb beeinträchtigen“ (*Gassner, a.a.O. Pharm. Ind. 70 (2008), S. 707*).

Die KNB ist kein Selbstzweck, sondern soll eine Grundlage für einen angemessenen Erstattungsbetrag bereitstellen. Spätestens zu dessen Berechnung bzw. Vereinbarung braucht es indikationsübergreifender Maßstäbe, um zu gleichheitsgerechten Entscheidungen zu gelangen. Da liegt es nahe, dass das IQWiG im Sinne einer guten Grundlage entsprechende Informationen zur Verfügung stellt. Ansonsten besteht die Gefahr, dass die Kosten-Nutzen-Bewertungen des IQWiG in einem zentralen Punkt unvollständig und daher nur eingeschränkt in dem eigentlichen Entscheidungsprozess verwertbar sind (vgl. *Huster, MedR 2010 a.a.O., S. 240*).

Wie in der rechtswissenschaftlichen Diskussion der KNB-Methoden bereits festgestellt, liegt in der fehlenden Rechtsgrundlage für indikationsspezifische Wirtschaftlichkeitskriterien ein erhebliches Legitimitätsdefizit und ein Verstoß gegen den rechtsstaatlich unterfütterten Grundsatz vom Vorbehalt des Gesetzes. „Denn auch wenn das IQWiG keine rechtsverbindlichen Entscheidungen trifft, sondern nur eine „Empfehlung“ (vgl. § 35 b Abs. 2 S. 1 SGB V) abgibt, gegen die der Rechtsschutz nach § 35 b Abs. 4 SGB V ausdrücklich ausgeschlossen ist, wird von seinen Bewertungen in der Sache sehr viel abhängen“ (*Huster, MedR 2010 a.a.O., S. 237*). Die Ergebnisse determinieren wegen ihrer faktischen Bindungswirkung die

spätere Entscheidung des G-BA und sind daher als mittelbare Grundrechtseingriffe zu Lasten der betroffenen Akteure (vor allem Patienten und Arzneimittelhersteller) einzustufen.

Im Ergebnis kann das IQWiG für sein Modell indikationsspezifischer Wirtschaftlichkeitskriterien keine zwingende Logik des Gesetzes beanspruchen.

Keine normative Kraft des Flüchtigen

Das Modell der Effizienzgrenzen in der vom IQWiG vertretenen Form erhebt zudem die zum Vergleichszeitpunkt bestehende (faktische) Effizienz – genauer: die Steigung des letzten Segmentes der Effizienzgrenze – zur allgemein gültigen Norm in der betreffenden Indikation. Weder sieht jedoch das SGB V diese normative Kraft des Faktischen vor, noch erscheint es sinnvoll, die zu einem willkürlichen Zeitpunkt berechnete letzte Steigung der Effizienzgrenze zum Maßstab der nächsten geforderten zu machen: Schon eine Preisanpassung oder der Ablauf des Patentschutzes der Arzneimittel des letzten Segmentes der Effizienzgrenze verändert die Steigung des Segmentes und damit das Kriterium; aus derart Beliebigem kann kein Maßstab hervorgehen.

Auch keine indikationsspezifischen Zusatznutzenbewertungskriterien

Abschließend eine Parallele zur Nutzenbewertung: Auch die Nutzenbewertung ist von ihrer Durchführung her natürlich zunächst indikationsbezogen; es findet eine Bewertung bei bestimmten, durch die Zulassung definierten Patientengruppen statt. Die Bewertungskriterien der Nutzenbewertung (assessment) allerdings sind bereits indikationsübergreifend, in Form einheitlicher Obergrenzen (aus Anhang A der IQWiG-Bewertung von Ticagrelor) bzw. in Form einheitlicher Kriterien aus der AM-NutzenV (zum Beispiel eine "erhebliche Verlängerung der Überlebensdauer" oder eine "langfristige Freiheit von schwerwiegenden Symptomen"). Diese Bewertungskriterien sind indikationsunabhängig. Die Parallele zu dieser vom Gesetzgeber erkannten und vom IQWiG praktizierten Logik der Nutzenbewertung sind indikationsunabhängige Bewertungskriterien auch für die Kosten-Nutzen-Bewertung. In der Beschlussphase (appraisal) können dann bei beiden Fragestellungen, - der Nutzenbewertung wie auch der Kosten-Nutzen-Bewertung - bedarfsbezogene (nicht indikationsbezogene) Kriterien hinzukommen, z. B. die Lebensendsituation bei einer bestimmten Tumorerkrankung (Indikation), in An- oder Abwesenheit von Alternativen (Bedarf).

Wir fordern das IQWiG daher auf, in Version 4.2 der Allgemeinen Methoden indikationsübergreifende Wirtschaftlichkeitskriterien aus den internationalen Methoden der Gesundheitsökonomie zu berücksichtigen.

2. Zusatznutzenbasierte anstelle von Zusatznutzenbereinigter Erstattung

Zu ermitteln sind nicht Preise, aus welchen der Zusatznutzen "entfernt" wurde. Wie bereits in unserer Stellungnahme zum Vorbericht der Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin dargelegt, empfehlen wir daher den Begriff "zusatznutzenbasierter Erstattungspreis".

3. Suche im Studienregister PharmNet.Bund Pflicht

Auf Seite 166 wurden folgende Änderungen vorgeschlagen:

Die Nutzenbewertung des Instituts umfasst regelhaft die Suche in großen, allgemeinen Studienregistern sowie Metaregistern solcher Studienregister. Dazu gehören insbesondere das Studienregister Clinicaltrials.gov des US-amerikanischen National Institute of Health (NIH) sowie, das Metaregister International Clinical Trials Registry Platform Search Portal (ICTRP) der WHO sowie das EU Clinical Trials Register der Europäischen Zulassungsbehörde EMA. Bei der Nutzenbewertung von Arzneimitteln werden zusätzlich Studienregister der pharmazeutischen Industrie (Studienregister einzelner Firmen und Metaregister) durchsucht. In krankheitsspezifischen Studienregistern wird nur im Ausnahmefall recherchiert. Das bei der Europäischen Zulassungsbehörde EMA angesiedelte europäische Studienregister (clinicaltrialsregister.eu) stellt wegen fehlender Funktionalitäten [548] und aufgrund des unvollständigen Datenbestands in der derzeitigen Form keine Quelle dar, die regelhaft bei der Nutzenbewertung des IQWiG herangezogen wird.

Wir folgern aus dieser Ergänzung, dass das IQWiG für die Nutzenbewertung von Arzneimittel keine regelhaften Suchen im Studienregister PharmNet.Bund durchführt, da dieses nicht gelistet wird. Diese Suche ist jedoch laut Verfahrensordnung für pharmazeutische Unternehmen Pflicht; seit 18.04.2013 sind die Suchergebnisse auch als ris-Datei in Modul 5 abzulegen. Wir weisen darauf hin, dass diesem Studienregister - wie einst dem EU Clinical Trials Register - die Exportfunktionalität fehlt. Wir bitten, dies zu erörtern, ggfs. auch mit dem Gemeinsamen Bundesausschuss.



Mit freundlichen Grüßen

MSD SHARP & DOHME GMBH
Outcomes Research / HTA

A handwritten signature in blue ink, appearing to read "K. Krobot".

Dr. Dr. Karl J. Krobot, MPH (USA)
Direktor Outcomes Research / HTA

MSD SHARP & DOHME GMBH
Outcomes Research / HTA

A handwritten signature in blue ink, appearing to read "Monika Scheuringer".

Dr. Monika Scheuringer
Head Biostatistics & HTA Management

MSD SHARP & DOHME GMBH
Rechtsabteilung

A handwritten signature in blue ink, appearing to read "S. Schulz".

Dr. Stefan Schulz
Rechtsanwalt

A.1.25 – Novartis Pharma GmbH

**Stellungnahme der Novartis Pharma GmbH zum Entwurf des
IQWiG „Allgemeine Methoden Version 4.2“ vom 18.06.2014**

Autoren: Volker Claus, Eva Dütting, Sven Klebs, Christiane Rämisch

**Novartis Pharma GmbH
Roonstrasse 25, 90429 Nürnberg**

Stellungnahme der Novartis Pharma zum Entwurf des IQWiG:

„Allgemeine Methoden Version 4.2“ vom 18.06.2014

Die Stellungnahme der Novartis Pharma GmbH (im folgenden NVS) bezieht sich auf die im Vergleich zur Version 4.1 „Allgemeine Methoden“ des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (im folgenden IQWiG) vom 28.11.2013 vorgenommenen Änderungen, d.h. auf die in der Entwurfsfassung Version 4.2 „Allgemeine Methoden“ des IQWiG vom 18.06.2014 gelb markierten Abschnitte sowie auf einzelne grundsätzliche Aspekte der „Allgemeine Methoden“ des IQWiG.

A) Abschnitt 2.1.4 Kosten-Nutzen-Bewertung nach § 35b SGB V

In Abbildung 4 auf Seite 23 wird vom IQWiG beschrieben, dass die Einbindung externer Experten nur optional erfolgt. Mehrere Fachgesellschaften haben in der Vergangenheit die aus ihrer Sicht unzureichende Einbindung durch das IQWiG kritisiert (Deutsche Diabetes Gesellschaft 2013). Diese Einschätzung teilen wir. Unserer Auffassung nach sind eine engere Einbindung sowie die Implementierung weiterer Fachexpertise notwendig, damit insbesondere auch die konkrete Versorgungssituation der Patienten bessere Berücksichtigung bei der Bewertung findet. Die Einbindung sollte regelhaft erfolgen.

In der Fußnote der Abbildung 4 auf Seite 23 ist ferner dargestellt, dass eine mündliche wissenschaftliche Erörterung nur optional durchgeführt wird. Aufgrund Neuheit und Bedeutung des Verfahrens für betroffene pharmazeutische Unternehmer, ist unserer Ansicht nach eine mündliche Erörterung generell notwendig.

Auf Seite 24 ist beschrieben: „Die **medizinische Expertise** wird primär auf Basis eines Fragenkatalogs eingeholt, der zu Beginn der Bewertung an die externen Sachverständigen verschickt wird.“ Da weder der konkrete Fragenkatalog noch die Antworten der externen Experten veröffentlicht werden, fehlt es dem Verfahren an dieser Stelle an Transparenz. Zudem obliegt die Zusammenstellung der Fragen und damit auch die Antwortmöglichkeiten der Experten der Entscheidungshoheit des IQWiG. Die externen Experten haben keine vollumfängliche Möglichkeit, eine eigene Position darzustellen bzw. über den Fragenkatalog hinaus Position zu beziehen. Dies wurde von den Fachgesellschaften bereits in der

Vergangenheit bemängelt (Deutsche Diabetes Gesellschaft 2013). Der Kritik schließen wir uns an und fordern daher eine transparente und umfassendere Einbeziehung der externen Sachverständigen.

Die gleiche Problematik gilt analog für den Patientenfragebogen und die entsprechenden Rückmeldungen der Patientenorganisationen. Dies ist ebenfalls auf Seite 24 beschrieben: „Die **Patientenperspektive** wird auf Basis eines Fragebogens erhoben, der zu Beginn der Bewertung an Patientinnen und Patienten bzw. Patientenorganisationen verschickt wird.“ Da es nicht klar ist, in wie weit die Sichtweise der Patienten und die für sie relevanten Punkte in die Bewertung des IQWiG einfließen, können die pharmazeutischen Unternehmer lediglich unzureichend Stellung nehmen. Hier wäre eine höhere Transparenz notwendig.

Auf Seite 24 ist ferner beschrieben: „Für eine Frist von drei Wochen wird dann der Öffentlichkeit Gelegenheit zur Stellungnahme (Anhörung) gegeben (schriftliche Stellungnahmen).“ Indes ist eine dreiwöchige Frist zur Stellungnahme eindeutig zu kurz. Dies resultiert z.B. daraus, dass der Vorbericht zur „Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medikamentösen Behandlungen; G09-01; 1.0; 09.11.2012“ einen Umfang von über 500 Seiten aufwies, der im ersten Schritt zunächst durchzuarbeiten und auszuwerten ist, bevor eine Stellungnahme verfasst werden kann. Um angemessene Stellung nehmen zu können, bedarf es mindestens einer Frist von 6-8 Wochen.

B) Abschnitt 2.1.5 Potenzialbewertung

In der Abbildung 5 auf Seite 25 beschreibt das IQWiG, dass externe medizinische Expertise optional einbezogen werden kann. Mehrere Fachgesellschaften haben in der Vergangenheit die aus ihrer Sicht unzureichende Einbindung externer Expertise durch das IQWiG kritisiert (Deutsche Diabetes Gesellschaft 2013). Der Kritik schließen wir uns an. Eine engere Einbindung führt aus unserer Sicht neben der Implementierung der weiteren Fachexpertise insbesondere zur besseren Berücksichtigung der konkreten Versorgungssituation der Patienten. Daher ist für uns nicht nachvollziehbar, warum es ausreichend sein sollte, dass externe Experten lediglich optional einbezogen werden.

C) Abschnitt 2.1.6 Addendum

Das IQWiG führt in Allgemeine Methoden Version 4.2. zur Vorgehensweise zur Veröffentlichung eines Addendums in einer ergänzenden Formulierung aus: *„Das Verfahren zur Veröffentlichung eines Addendums orientiert sich an dem des ursprünglichen Institutsprodukts.“* Weiter wird dann präzisiert: *„So wird beispielsweise ein Addendum zu Berichten zunächst an den Auftraggeber sowie an Stiftungsrat und Vorstand versandt. I. d. R. eine Woche später wird es dem Kuratorium der Stiftung übermittelt und weitere drei Wochen später auf der Website des Instituts veröffentlicht.“* Diese skizzierte Vorgehensweise ist für Addendum-Produkte in Verfahren zur Dossierbewertung nach § 35a SGB V nicht ausreichend. In Allgemeine Methode Version 4.2 Abschnitt 2.1.3 ist zur Vorgehensweise bei der Dossierbewertung nach § 35a SGB V festgeschrieben, dass die Publikation der IQWiG Dossierbewertung zeitnah im Anschluss an die Fertigstellung und Weiterleitung an den Gemeinsamen Bundesausschuss als den Auftraggeber erfolgt. Dies entspricht der vom gesetzlichen Normengeber geforderten Verfahrenstransparenz. *„Die Dossierbewertung wird nach ihrer Fertigstellung dem G-BA zugeleitet. Zeitnah im Anschluss daran erfolgen die Weiterleitung an den Vorstand der Stiftung, den Stiftungsrat und an das Kuratorium der Stiftung sowie die Publikation der Dossierbewertung auf der Website des Instituts.“*

Wir schlagen daher vor, in Verfahren nach § 35a SGB V die Publikation eines Addendum-Produkts ebenso zeitnah vorzusehen, genau wie dies bei der Publikation einer Dossierbewertung erfolgt. Damit kann auch hier der vom Gesetzgeber geforderten Verfahrenstransparenz entsprochen werden. Basierend auf der in den bisherigen Verfahren nach § 35a SGB V von G-BA und IQWiG regelhaft angewendeten Praxis der Veröffentlichung der Dossierbewertungen auf den Websites von IQWiG und G-BA binnen 3 bis 4 Kalendertagen nach Fertigstellung einer Dossierbewertung durch das IQWiG, schlagen wir eine vergleichbar zeitnahe Veröffentlichung auch des Addendums in Verfahren nach § 35a SGB V auf der Website des IQWiG binnen 3 bis 4 Kalendertage nach dessen Fertigstellung durch das IQWiG als ebenfalls ausreichend vor.

D) Abschnitt 2.2.1 Auswahl externer Sachverständiger

Auf Seite 33 streicht das IQWiG ohne ersichtlichen Grund die Veröffentlichung der durch das Institut zu vergebenden Aufträge nach § 139b Abs. 3 SGB V. Es wird nicht begründet, woraus diese Entscheidung resultiert. Demnach ist die Entscheidung nicht nachvollziehbar und diesseits bestehen Bedenken an der notwendigen Transparenz. Daher schlagen wir vor, die ursprüngliche Version zu belassen.

E) Kapitel 3 zur Nutzenbewertung medizinischer Interventionen (Abschnitt 3.1.1 und 3.2.4)

In diesem Kapitel schlägt das IQWiG die Streichung der Zielgrößen „Patientenzufriedenheit“ bzw. „Behandlungszufriedenheit“ vor. Als patientenrelevant gelten bislang neben der *Mortalität*, *Morbidität* und *gesundheitsbezogenen Lebensqualität* ergänzend der interventions- und erkrankungsbezogene Aufwand sowie die Behandlungs- bzw. Patientenzufriedenheit. Der Entwurf Allgemeine Methoden Version 4.2 sieht an verschiedenen Stellen (u.a. Abschnitt 3.1.1 und 3.2.4) vor, die Endpunkte „*Behandlungszufriedenheit*“ und „*Patientenzufriedenheit*“ nicht weiter als Zielgröße bei der Nutzenbewertung von Interventionen mit zu berücksichtigen. Dieser Vorschlag wurde vom IQWiG nicht weiter begründet, weshalb die Entscheidung insofern in keiner Weise nachvollziehbar ist. Zur Erhebung der Patientenzufriedenheit existieren validierte Fragebögen für verschiedene Krankheitsgebiete (z.B. Atkinson 2004, Peyrot 2014, Cobos 2014, Bharmal 2009), die geeignet sind, diesen Parameter standardisiert und verlässlich zu erheben. Gerade bei unterschiedlichen Applikationsweisen kann bei vergleichbarer Wirksamkeit und Sicherheit unterschiedlicher Arzneimittel die Patientenzufriedenheit mit einem Arzneimittel einen Effekt auf dessen regelmäßige Anwendung haben.

Vor diesem Hintergrund schlagen wir vor, „*Patientenzufriedenheit*“ bzw. „*Behandlungszufriedenheit*“ als für die Nutzenbewertung von medizinischen Interventionen für die Praxis relevante und angewandte Zielgröße beizubehalten und die diesbezüglich im Entwurf Allgemeine Methoden Version 4.2 des IQWiG enthaltenen Vorschläge zur Löschung der entsprechenden Textstellen nicht umzusetzen.

F) Abschnitt 3.3.3 Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V

Das IQWiG sollte die aktuelle Überarbeitung seiner Allgemeine Methoden Version 4.2 nutzen, eine weitere Präzisierung auch des Abschnitts 3.3.3 Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V vorzunehmen. Die in den einzelnen bisherigen Nutzenbewertungsverfahren gemäß § 35a SGB V gemachten Erfahrungen sollten in einer Überarbeitung des Abschnitt 3.3.3 dahingehend aufgegriffen werden, dass die vom gesetzlichen Normengeber klar formulierte Forderung nach vorrangiger Berücksichtigung der Zulassungsstudien zur Bewertung von Arzneimitteln in den Verfahren nach § 35a SGB V in den Methoden des IQWiG deutlich abgebildet wird.

Die Arzneimittelnutzenbewertungsverordnung (AM-NutzenV) schreibt in § 5 Abs. 3 vor: „Für die *erstmalige Bewertung nach § 35a des Fünften Buches Sozialgesetzbuch zum Zeitpunkt der Markteinführung sind für die Bewertung des Arzneimittels mit neuen Wirkstoffen grundsätzlich die Zulassungsstudien zugrunde zu legen.*“ Entsprechend des besonderen Stellenwerts der Zulassungsstudien im Bewertungsverfahren nach § 35a SGB V ist in § 4 Abs. 6 AM-NutzenV die

vollständige Übermittlung aller Zulassungsstudien festgehalten: „Für das zu bewertende Arzneimittel mit neuem Wirkstoff legt der pharmazeutische Unternehmer im Dossier den Ergebnisbericht der Zulassungsstudien einschließlich der Studienprotokolle und des Bewertungsberichtes der Zulassungsbehörde vor, sowie alle Studien, die der Zulassungsbehörde übermittelt worden sind.“ Weitere Nachweise können vom Gemeinsamen Bundesausschuss verlangt werden, sofern die Zulassungsstudien nicht ausreichen. Geforderte Studien dürfen dabei nicht *unmöglich* oder *unangemessen* sein - in diesen Fällen sind best-verfügbare Nachweise zu übermitteln. Gemäß der aktuellen europäischen Schutzvorschriften durch Richtlinie 2001/20/EG, umgesetzt in deutschem Arzneimittelgesetz (AMG) und Verordnung zur Good Clinical Practice, muss für die Durchführung einer klinischen Prüfungsstudie eine vorherige zustimmende Bewertung der zuständigen Ethik-Kommission sowie die Genehmigung durch die zuständige Bundesoberbehörde vorliegen. Dies bestimmt Anforderungen, Machbarkeit und konsenterte Umsetzung eines Studiendesigns wesentlich mit, wie beispielsweise die Durchführung von Studien mit erlaubtem Behandlungswechsel in bestimmten Therapiegebieten.

Daher schlagen wir vor, den Abschnitt 3.3.3 *Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V* der Allgemeine Methoden Version 4.2. inhaltlich dahingehend zu ergänzen und die besondere Wertigkeit der Zulassungsstudien in Bewertungsverfahren nach § 35a SGB V explizit gemäß der Vorgaben durch die Arzneimittelnutzenbewertungsverordnung zu benennen - beispielsweise durch Integrierung des vorherigen Absatzes dieser Stellungnahme in den Abschnitts 3.3.3 Allgemeine Methoden Version 4.2 des IQWiG.

G) Abschnitt 3.8 Potenzialbewertung

Bei der Potentialbewertung führt das IQWiG auf den Seiten 74/75 drei neue Grade für die Ergebnissicherheit von Evidenz nicht-randomisierter Studien ein. Diese reichen von „gering“ über „sehr gering“ zu „minimal“.

Ein Potential würde dann aus der Kombination der Ergebnissicherheit und Effektgröße abgeleitet werden. Bei geringerer Ergebnissicherheit müssen nach Maßgabe des IQWiG die Studien unterschiedlich große Effekte aufweisen. Es werden Grenzen für das relative Risiko vorgeschlagen („Als ungefähre Grenzen zwischen kleinen, mittleren und großen Effekten können für das relative Risiko Werte von 0,8 und 0,5 dienen [147,419].“). Dies führt dazu, dass neben der Schwelle der statistischen Signifikanz (üblicherweise 5% Irrtumswahrscheinlichkeit) hierüber ein weiteres Entscheidungskriterium für die abschließende Potentialableitung etabliert wird. Eine der beiden vom IQWiG angeführten

Literaturstellen (Quelle 419) beschreibt, dass sehr große Behandlungseffekte vor allem aus kleineren Studien stammten. Ferner beziehen beide Quellen sich häufiger auf Laborparameter. Unabhängig vom einzelnen Endpunkt und der Größe der Studie scheint es daher nicht angemessen, generelle Grenzen für das relative Risiko von 0,8 bzw. 0,5 anzusetzen, ohne die Besonderheiten der Studie, der Endpunkte und der Indikation in umfassender Weise mit einzubeziehen. Die a priori Definition und vorgeschlagene einheitliche Anwendung dieser Schwellen zur Potentialableitung ist daher nicht zielführend.

H) Abschnitt 4.1.4 Auswahl der Komparatoren

Das IQWiG führt auf Seite 77 an: „Substanzen z. B. können auch zu Substanzklassen zusammengefasst werden, wenn ausreichende Homogenität vorliegt.“

Die konkrete Beschreibung des Vorgehens an dieser Stelle ist unzureichend. Es wird zudem nicht näher ausgeführt, wie „ausreichende Homogenität“ zu quantifizieren ist. Im Rahmen von Meta-Analysen beschreibt der Begriff der Homogenität bzw. Abwesenheit von Heterogenität lediglich die Vergleichbarkeit der betrachteten Studien bzw. der in den Studien beobachteten Effekte, wobei die Homogenität jedoch nicht für alle betrachteten Endpunkte gegeben sein muss. Ferner können spezifische Wechselwirkungen oder Gegenanzeigen in der Fachinformation zu anderen Einsatzschwerpunkten der Arzneimittel führen. Die Zusammenfassung zu Substanzklassen ist daher kritisch zu betrachten.

I) Kapitel 6 Evidenzbasierte Gesundheitsinformation für Bürgerinnen und Bürger

Im Kapitel 6 Evidenzbasierte Gesundheitsinformation für Bürgerinnen und Bürger erläutert das IQWiG in Abschnitt 6.2 seine Vorgehensweise zur Themenwahl, Informationsbeschaffung und wissenschaftliche Bewertung. Es wird ausgeführt, dass „zur Erstellung des Themenkataloges ... primär der Versorgungsreport des Wissenschaftlichen Instituts der AOK herangezogen“ wird. Damit plant das IQWiG zukünftig von seiner bisherigen Vorgehensweise gemäß Allgemeine Methoden Version 4.1. abzuweichen, wonach „...das Institut verschiedene Quellen u. a. zu Todesursachen, Diagnosen bei stationären Einweisungen, krankheitsbedingten Fehlzeiten am Arbeitsplatz, häufigen Diagnosen und Arzneimittelverordnungen in Deutschland sowie die Liste der im Rahmen des morbiditätsorientierten Risikostrukturausgleichs ausgewählten Krankheiten“ nutzt. Die geplante Änderung seiner Methoden und primäre Berücksichtigung nur einer Quelle bei die Themenwahl wird vom IQWiG nicht weiter erläutert und begründet.

Wir schlagen daher vor, anstatt eine Einschränkung der Informationsquellen bei der Themenidentifizierung und -auswahl zur Gesundheitsinformation für Bürgerinnen und Bürger vorzunehmen, weiterhin auf verschiedene öffentlich verfügbare Quellen zurückzugreifen und diese im Methodenpapier Version 4.2. explizit zu benennen. Beispielhaft sei als weitere relevante Quelle im deutschen Versorgungskontext die „Gesundheitsberichterstattung Bund“ genannt.

J) Kapitel 7 Informationsbeschaffung

Das IQWiG führt auf Seite 164 an: *„Ebenso werden unaufgefordert übermittelte Daten, d. h. außerhalb von Stellungnahmeverfahren oder außerhalb sonstiger bestehender Regelungen (z. B. Herstelleranfragen), inhaltlich nicht berücksichtigt. Die unaufgeforderte Übermittlung von Studiendaten birgt das Risiko, dass sie selektiv geschieht und es dadurch zu einer Verzerrung des Ergebnisses der Nutzenbewertung kommt.“* Aus unserer Sicht ist dieser kategorische Ausschluss problematisch, da fehlerhafte oder unvollständige Darstellungen/Veröffentlichungen des IQWiGs damit anlassbezogen nicht korrigiert werden können. Es wäre unverhältnismäßig, wenn der pharmazeutische Unternehmer in diesen Fällen gegebenenfalls Jahre auf ein erneutes Stellungnahmeverfahren warten müsste, um die entsprechende Evidenz vorzulegen. Da das IQWiG entsprechend der Punkte 7.1 und 7.2 ohnehin eigene Recherchen durchführt bzw. die Informationsbeschaffung überprüft, sollten die vom pharmazeutischen Unternehmer zusätzlich übermittelten Daten zu keiner kritischen Verzerrung führen, da sie die Informationen aus den Abschnitten 7.1 bzw. 7.2 ergänzen bzw. in den Kontext der aus diesen Quellen stammenden Recherchen gesetzt werden können.

Literatur:

Atkinson MJ et al. Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease. *Health Qual Life Outcomes*. 2004 Feb 26;2:12

Bharmal M et al. Validation of an abbreviated Treatment Satisfaction Questionnaire for Medication (TSQM-9) among patients on antihypertensive medications. *Health Qual Life Outcomes*. 2009 Apr 27;7:36. doi: 10.1186/1477-7525-7-36.

Cobos JP et al. Development and validation of the scale to assess satisfaction with medications for addiction treatment-methadone for heroin addiction (SASMAT-METHER). *Drug Alcohol Depend*. 2014 Jun 11. pii: S0376-8716(14)00907-7. doi: 10.1016/j.drugalcdep.2014.05.024. [Epub ahead of print]

Delestras S et al. Comparison between two generic questionnaires to assess satisfaction with medication in chronic diseases. *PLoS One*. 2013;8(2):e56247. doi: 10.1371/journal.pone.0056247. Epub 2013 Feb 20.

Deutsche Diabetes Gesellschaft et al. Gemeinsame Stellungnahme Deutsche Diabetes Gesellschaft (DDG), Deutsche Gesellschaft für Innere Medizin (DGIM), Deutsche Gesellschaft für Kardiologie (DGK), Deutsche Krebsgesellschaft (DKG), Deutsche Gesellschaft für Verdauungs- und Stoffwechselkrankheiten (DGVS) zum Entwurf „Aktualisierung einiger Abschnitte der Allgemeinen Methoden Version 4.0 sowie neue Abschnitte zur Erstellung der Allgemeinen Methoden Version 4.1“ des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) vom 18.04.2013. http://www.deutsche-diabetes-gesellschaft.de/fileadmin/Redakteur/Stellungnahmen/220513_Gemeinsame_Stellungnahme_zum_Methodenpapier_des_IQWiG.pdf; zuletzt abgerufen am 05.08.2014

IQWiG. Allgemeine Methoden. Version 4.1 vom 28.11.2013.

IQWiG. Allgemeine Methoden zur Bewertung von Verhältnissen zwischen Nutzen und Kosten. Version 1.0 vom 12.10.2009.

IQWiG. Vorbericht (vorläufige Kosten-Nutzen-Bewertung). Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medikamentösen Behandlungen. 09.11.2012. https://www.iqwig.de/download/G09-01_Vorbericht_Kosten-Nutzen-Bewertung-von-Venlafaxin-Duloxetin-Bupropion-und-Mirtazapin.pdf; zuletzt abgerufen am 05.08.2014

Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses zur Änderung der Arzneimittel-Richtlinie (AM-RL) Anlage XII zur Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V- Extrakt aus Cannabis Sativa vom 21. Juni 2012

Gesundheitsberichterstattung Bund. <http://www.gbe-bund.de/>

Peyrot M et al. Development and validation of the Diabetes Medication System Rating Questionnaire-Short Form. Diabet Med. 2014 Mar 26. doi: 10.1111/dme.12453. [Epub ahead of print]

A.1.26 – Sanofi-Aventis Deutschland GmbH

Stellungnahme der Sanofi-Aventis Deutschland GmbH zum IQWiG Methodenpapier Entwurf für Version 4.2

Zum Stellungnahmeverfahren allgemein:

Am 18.06.2014 wurde das schriftliche Stellungnahmeverfahren zur Entwurfsversion V 4.2 „Allgemeine Methoden“ eingeleitet.¹ Auf seiner Homepage beschreibt das IQWiG das weitere Procedere wie folgt: „Zu den neuen Inhalten des Entwurfs für Version 4.2 können bis zum 31. Juli 2014 schriftliche Stellungnahmen eingereicht werden. Das IQWiG wird diese sichten und würdigen. Danach werden die betreffenden Abschnitte entsprechend überarbeitet und das Methodenpapier auf den Internetseiten des Instituts in der Version 4.2 veröffentlicht.“² Wie bereits im Verfahren zur Überarbeitung des Dokuments „Allgemeine Methoden“ V.4.1. im Jahr 2013 geschehen, ist auch hier eine mündliche Anhörung im Nachgang zum schriftlichen Stellungnahmeverfahren unerlässlich. Neben der Verpflichtung des IQWiG gemäß § 139a SGB V, in regelmäßigen Abständen über Arbeitsprozesse und -ergebnisse einschließlich der Grundlagen für die Entscheidungsfindung öffentlich zu berichten, fördert solch ein Anhörungsverfahren den kontinuierlichen wissenschaftlichen Austausch.

Zu 2.1. „Produktspezifische Verfahrensabläufe“

Hinsichtlich Potenzialbewertungen wird hier festgelegt, dass Potenzialbewertungen im Auftrag des G-BA durchgeführt werden und entsprechende Anhörungsverfahren nicht beim IQWiG erfolgen. Erst im Falle einer Erprobung würde der G-BA ein Stellungnahmeverfahren zur Erprobungsrichtlinie durchführen. (s. S. 13 Allgemeine Methoden V 4.2.) Im Licht des Transparenzgrundsatzes ist es nicht nachvollziehbar, warum das IQWiG ein Anhörungsverfahren im Rahmen der Potenzialbewertung ausschließt, denn auch unter der gebotenen Vertraulichkeit sollte den betroffenen Parteien die Möglichkeit gegeben werden, zu offenen Fragen Stellung zu nehmen.

Zu 2.1.5. „Potenzialbewertung“

In der Abbildung 5 wird der Ablauf der Potenzialbewertung durch das IQWiG schematisch dargestellt, ohne dass jedoch der Prozess selbst inhaltlich transparent dargestellt wird. Auch eine Einbeziehung von externen Sachverständigen ist nur fakultativ vorgesehen. (s. S. 25 Allgemeine Methoden V 4.2) Da bei der Potenzialbewertung von Medizinprodukten regelhaft zu erwarten ist, dass entsprechende Evidenz nur begrenzt vorhanden ist, sollte es für das Institut verpflichtend sein, externe Sachverständige heranzuziehen. Insbesondere der interne Qualitätssicherungsprozess für die Potenzialbewertung sollte transparent im Methodenpapier dargelegt werden.

Zu 2.1.7. „Gesundheitsinformation“

¹ IQWiG 2014: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden. Entwurf für Version 4.2. vom 18.06.2014, Köln 2014.

² IQWiG 2014a: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, https://www.iqwig.de/de/presse/pressemitteilungen/pressemitteilungen/ueberarbeitung_methodenpapier_iq_wig_stellt_entwurf_fur_version_4_2_zur_diskussion.6134.html, entnommen am 14.07.2014

Es ist nicht nachvollziehbar, warum im Zuge der Gesundheitsinformationserstellung eine externe Stellungnahme nur auf einen kleinen Kreis beschränkt werden soll, nämlich Auftraggeber, Kuratorium und weitere Gremien des Instituts/ IQWiG. Im Kuratorium sind die Organisationen der Leistungserbringer und die Selbstverwaltungsorgane der Trägerorganisationen des Gemeinsamen Bundesausschusses, inklusive Patientenvertretern, repräsentiert. Was fehlt, ist die Möglichkeit, dass auch Personen und Institutionen in der Wissenschaft, Hersteller von pharmazeutischen Produkten und Medizinprodukten und sonstige Interessierte offiziell Stellung nehmen können. Es wird daher gefordert, die Möglichkeit der Stellungnahme auf einen größeren Kreis Interessierter und Betroffener zu erweitern. (s. S. 30 Allgemeine Methoden V 4.2)

Zu 2.2.1. „Auswahl externer Sachverständiger“, 2.2.2 Gewährleistung der fachlichen Unabhängigkeit und 2.2.3 Review der Produkte des Instituts

Laut eigener Aussage vergibt das IQWiG Aufträge an externe Sachverständige in einem transparenten und nichtdiskriminierenden Wettbewerb. Dies erfolgt über eine Registrierung Interessierter in eine Sachverständigen-Datenbank beim IQWiG, aus der das IQWiG je nach Auftrag Sachverständige nach gesetzten Kriterien auswählt. (s. S. 33 Allgemeine Methoden V 4.2.) Um dem Grundsatz der Transparenz nachzukommen, wäre es jedoch wünschenswert und notwendig, das Auswahlverfahren, die in Frage kommenden Sachverständigen und auch die letztlich angewendeten Kriterien, die zur Wahl eines bestimmten Sachverständigen geführt haben, zu veröffentlichen.

Im Sinne der Transparenz wäre es weiterhin wünschenswert, dass auch die externen Reviews veröffentlicht werden, was bis dato nicht der Fall ist. (s. S. 36 Allgemeine Methoden V 4.2)

Zu 3.1. „Patientenrelevanter medizinischer Nutzen und Schaden“ (Lebensqualität)

„Die Zielgrößen (Mortalität, Morbidität und gesundheitsbezogene Lebensqualität) sind auch im SGB V als vorrangig zu berücksichtigende Zielgrößen genannt, z. B. in § 35 Abs. 1b SGB V. Ergänzend kann der interventions- und erkrankungsbezogene Aufwand **und die Zufriedenheit der Patientinnen und Patienten** mit der Behandlung berücksichtigt werden.“

Die Streichung der „Zufriedenheit des Patienten“ neben den Erfolgskriterien Mortalität, Morbidität und Lebensqualität erfolgt ohne Begründung. Offen bleibt, warum die Zufriedenheit der Patienten in Zukunft keine Rolle mehr spielen soll.

In diesem Zusammenhang sei hier zunächst auf die umfassende Definition der WHO Quality of Life Gruppe verwiesen, die Lebensqualität als „die Wahrnehmung von Individuen bezüglich ihrer Position im Leben im Kontext der Kultur und der Wertesysteme, in denen sie leben, und in Bezug auf ihre Ziele, Erwartungen, Standards und Interessen“.³ Inzwischen hat sich eine operationale Definition durchgesetzt, nach der sich Lebensqualität als multidimensionales Konstrukt aus Wohlbefinden und Funktionsfähigkeit aus Sicht der Patienten zusammensetzt.⁴ Grundlegende Dimensionen betreffen die körperliche, die emotionale, die mentale, die soziale und die alltagsfunktionale Komponente.

³ WHO. The WHO quality of life assessment (WHOQOL): Position paper from the WHO. Soc Sci Med 1995; 41: 1403-09

⁴ LC Doward, SP McKenna. Defining Patient Reported Outcomes. Value in health 2004; 7/1: S4-S8

Darüberhinaus kann sich auch die Patientenzufriedenheit auf die gesundheitsbezogene Lebensqualität auswirken.

So zählen auch subjektive Wohlbefindensmodelle zu den zentralen konzeptuellen Ansätzen der Lebensqualitätsmessung.⁵ Als breit akzeptiertes Instrument sei hier beispielhaft der Diabetes Treatment Satisfaction Questionnaire (DTSQ) genannt. Der DTSQ (Status Version) besteht aus acht Positionen (Items). Sechst davon erfassen die Behandlungszufriedenheit und zwei weitere die wahrgenommene Frequenz von Hyper- und Hypoglykämien. Jedes Item wird auf einer Skala von 0 (sehr unzufrieden) bis 6 (sehr zufrieden) bewertet.⁶ Der DTSQ wurde von der Weltgesundheitsorganisation (WHO) sowie der International Diabetes Federation (IDF) als sehr wertvoll in der Erhebung von Outcomes der Diabetesbehandlung eingestuft.⁷ Mittlerweile liegen unzählige Publikationen vor, in denen der DTSQ zur Messung der Behandlungszufriedenheit eingesetzt wurde.⁸

Zu 3.1.1 „Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens“ und 3.2.4. „Patientenberichtete Endpunkte“

Das IQWiG sieht in seinem Methodenentwurf V.4.2. vor, die Termini „Patientenzufriedenheit“ und „Behandlungszufriedenheit“ als Aspekte der patientenberichteten Endpunkte (PROs) ersatzlos zu streichen. Im Gegensatz dazu thematisiert die Europäische Arzneimittelzulassungsbehörde EMA (European Medicines Agency) die Bedeutung der verschiedenen Aspekte patientenberichteter Endpunkte in ihrem kürzlich publizierten „Reflection Paper on the use of patient reported outcome (PRO) measures in oncology studies“. Dort wird dargestellt, dass PROs gewissermaßen als Oberbegriff für Symptome, gesundheitsbezogene Lebensqualität, Gesundheitszustand, Behandlungssadhärenz und Behandlungszufriedenheit, die direkt vom Patienten berichtet werden, zu sehen sind⁹. Keinem dieser Aspekte wird dort eine höhere Wertigkeit zugesprochen, umso weniger verständlich ist es, dass das IQWiG mit der Streichung der Behandlungszufriedenheit diese offenbar in ihrer Bedeutung abwertet.

Da der Patient und sein objektives und auch subjektives Wohlbefinden im Mittelpunkt stehen sollten, stellen Patientenzufriedenheit bzw. Behandlungszufriedenheit mit der gewählten Therapie wichtige Indikatoren für den Nutzen der Therapie dar. Erfassung und Darstellung der Patientenzufriedenheit und Behandlungszufriedenheit sollten daher auch weiterhin wichtige Bausteine in der Bewertung verschiedener Therapien bleiben.

⁵ M. Bullinger. Das Konzept der Lebensqualität in der Medizin – Entwicklung und heutiger Stellenwert. Z. Evid. Fortbild. Qual. Gesundh. wesen 2014; 108: 97-103

⁶ C. Bradley. The Diabetes treatment satisfaction questionnaire: DTSQ. In Bradley C, ed. Handbook of Psychology and Diabetes. A guide to psychological measurement in diabetes research and practice. Chur, Switzerland: Harwood Academic Publishers 1994; 11-133

⁷ C. Bradley, DS Gamsu. Guidelines for encouraging psychological well-being: report of a working group of the WHO regional office for Europe and IDF European region St. Vincent declaration action programme for diabetes. Diabet Med 1994; 11: 510 - 516

⁸ C. Bradley, CJB Gilbride. Improving treatment satisfaction and other patient reported outcomes in people with type 2 diabetes: the role of once-daily insulin glargine. Diabetes, Obesity and Metabolism 2008; 10/1: 50 - 65

⁹ European Medicines Agency. Reflection Paper on the use of patient reported outcome 4 (PRO) measures in oncology studies. 17.06.2014. EMA/CHMP/292464/2014

Zu 3.8. „Potenzialbewertung“

Das IQWiG führt in seinem Methodenpapier aus, dass „aufgrund der besonderen Zielsetzung [...] bei der Potenzialbewertung im Vergleich zur Nutzenbewertung deutlich niedrigere Anforderungen an die Evidenz zu stellen“ sind. Zugleich werden hier zusätzliche Grade der Ergebnissicherheit „gering“, „sehr gering“ und „minimal“ neu eingeführt. Für jeden dieser Grade werden zudem minimale Effektstärken definiert, für die wiederum im Entwurf des Methodenpapiers 4.2 nur „ungefähre Grenzen“ festgelegt werden. (s. S. 73-75 Allgemeine Methoden V 4.2.)

Diese neu eingeführten Effektstärken „gering“, „sehr gering“ und „minimal“ führen eher dazu, dass selbst die bestmöglich verfügbare Evidenz bei der Testung von Medizinprodukten sprachlich herabgewürdigt wird. Um der Tatsache Rechnung zu tragen, dass im Bereich von Medizinprodukten schon die bestmöglich generierbare Evidenz unterhalb dessen liegt, was bei Arzneimitteln im Rahmen von klinischen Studien möglich ist, sollte daher die Ergebnissicherheit bei Medizinprodukten deskriptiv dargestellt werden.

Das Design einer möglichen Erprobungsstudie, die sich aus einer Potenzialbewertung ergibt, sollte in jedem Fall im Dialog zwischen dem IQWiG, der Selbstverwaltung und dem Hersteller, ggf. unter Hinzunahme von externen Sachverständigen, erarbeitet werden. Wie bereits oben ausgeführt, könnte das z.B. im Rahmen eines Anhörungsverfahrens, die das Methodenpapier leider nicht vorsieht, diskutiert werden. (s. oben Kommentar zu 2.1. „Potenzialbewertung“)

Zu 4. „Kosten-Nutzen-Bewertung medizinischer Interventionen“

4.1 Einführung

„Gemäß SGB V bestimmt das Institut auftragsbezogen über die Methoden und Kriterien für die Erarbeitung von Kosten-Nutzen-Bewertungen (KNB) auf der Grundlage der in den jeweiligen Fachkreisen anerkannten internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie.“ „Alle Abweichungen von den hier vorgelegten Methoden sind im Einzelfall zu begründen.“

Diese Generalklausel eröffnet dem IQWiG einen unbegrenzten Spielraum für Einzelfallentscheidungen. Sie bedeutet gleichzeitig eine maximale Planungsunsicherheit für den pharmazeutischen Unternehmer. Insofern sollte jeder KNB ein Scoping-Verfahren vorangestellt werden, in dem sich beide „Parteien“ gemeinsam auf eine verbindliche Vorgehensweise einigen.

„Daher muss die KNB einen Vergleich der Effizienz von Interventionen in einem Indikationsgebiet ermöglichen, ohne die Frage einer indikationsübergreifenden Priorisierung im Gesundheitssystem zu beantworten.“

Die Beschränkung der KNB auf Interventionen in nur einem Indikationsgebiet wird einer gesamtwirtschaftlichen und gesundheitspolitischen Betrachtung nicht gerecht. Indikationsübergreifende Bewertungen dienen der Allokation von Ressourcen in diejenigen Bereiche, in denen sie den größten gesamtgesellschaftlichen Nutzen entfalten.

4.1.2 Perspektive

Der Unterschied zwischen der GKV-Perspektive sowie der Perspektive der GKV-Versicherten-gemeinschaft sollte präzisiert werden.

4.1.5 Versorgungspfad

Der Terminus „Versorgungspfad“ sollte definiert werden. Der Begriff „Versorgungskontext“ sollte generell durch Versorgungspfad ersetzt werden, sofern damit das Gleiche gemeint ist. Ist dies nicht der Fall, so muss auch der Begriff „Versorgungskontext“ definiert werden.

4.1.7 Datengrundlage

„Für das Maß des Gesamtnutzens werden einzelne Studien oder Datenerhebungen (siehe auch Abschnitt 4.3.3) verwendet.“

Das Maß des Gesamtnutzens stellt den Kern jeder KNB dar und bedarf deshalb einer präzisen Darstellung. Insbesondere müssen methodische Hinweise zur Identifikation, Gewichtung und Aggregation mehrerer Endpunkte aus Patientenperspektive ergänzt werden (z.B. in Anlehnung an die Methode der multikriteriellen Entscheidungsanalyse „Multiple Criteria Decision Analysis (MCDA)“. So liefern innovative Studienformen wie z.B. die präferenzbasierte Nutzenmessung mittels Discrete Choice Experiment (DCE) oder Analytic-Hierarchy-Process (AHP) Verfahren wertvolle Zusatzinformationen über die Bedeutung von Therapien aus Sicht der Betroffenen.^{10 11 12} Einschlägig bekannte Pilotprojekte dazu sind bereits seit längerer Zeit abgeschlossen und liegen dem IQWiG vor.

Im Übrigen gibt es aus gesundheitsökonomischer Sicht keine inhaltlichen oder methodischen Anhaltspunkte, welche den Einsatz unterschiedlicher Instrumentarien im Kontext der Nutzen-Schaden Abwägung (Kapitel 3) und der Kosten-Nutzen Analyse (Kapitel 4) sinnvoll erscheinen lassen. Eine gesundheitsökonomisch sinnvolle Kosten-Nutzen-Bewertung sollte in beiden Anwendungsfeldern auf die gleichen Methoden und Instrumente zurückgreifen.

4.2.3 Einflussdiagramm und Modellkonzept

Das Einflussdiagramm sollte durch ein Beispiel veranschaulicht werden.

4.3.2 Endpunkte

„Wenn mehrere patientenrelevante Endpunkte nebeneinander dargestellt werden, wird für jeden patientenrelevanten Endpunkt eine eigene Effizienzgrenze erstellt.“

Diese Vorgehensweise ist nicht mehr zeitgemäß. Es existiert ein ausreichend erprobtes methodisches Instrumentarium mit dem verschiedene Nutzenkomponenten zu einem einzigen Maß des Gesamtnutzens aggregiert werden können (siehe Ausführungen zu Pkt. 4.1.7).

Zu 7. „Informationsbeschaffung“

¹⁰ Rottenkolber D. Discrete-Choice-Experimente zur Messung der Zahlungsbereitschaft für Gesundheitsleistungen - ein anwendungsbezogener Literaturreview. *Gesundh ökon Qual manag* 2011; 16: 232-244

¹¹ Mühlbacher A.C., Bethge S., Tockhorn A. Präferenzmessung im Gesundheitswesen: Grundlagen von Discrete-Choice-Experimenten. *Gesundh ökon Qual manag* 2012; 17: 159–172

¹² Johnson F.R., Lanscar E, Marshall D. ISPOR Task Force Report: Constructing Experimental Designs for Discrete-Choice-Experiments. *Value in health* 2013; 16: 3-13

Im einleitenden Teil von Kapitel 7 (vormals Kapitel 6) wird – ergänzend zur bisherigen Fassung des Methodenpapiers – festgelegt, dass außerhalb bestehender Regelungen unaufgefordert übermittelte Studiendaten seitens des Institutes bei der Abfassung seiner Berichte inhaltlich nicht berücksichtigt werden. Begründet wird dies mit dem potentiellen Risiko einer selektiven Auswahl der übermittelten Daten, was wiederum zu einer Verzerrung der Berichtsergebnisse führen könne. (s. S. 164 Allgemeine Methoden V 4.2) Auffällig ist, dass in dieser Begründung nun ganz explizit die Ergebnisse von Nutzenbewertungen genannt werden, während der eingefügte Abschnitt ansonsten bezüglich anderer möglicherweise betroffener Produkte des Instituts unspezifisch ist.

Eine derart pauschale Einschränkung hinsichtlich der Erhebung der Datengrundlage für Berichte aller Art, wie sie hier vorgesehen ist, könnte dazu führen, dass relevante Informationen bzw. neueste und aktuellste Evidenz in der Berichterstellung keine Berücksichtigung fänden, was dann umgekehrt ebenfalls eine Verzerrung des Berichtsergebnisses nach sich zöge. Diese Vorgehensweise erscheint auch nicht konform mit der im zur Kommentierung stehenden Methodenpapier selbst in Abschnitt 1.2.4, Absatz 4 aufgeführten zentralen Strategie der evidenzbasierten Medizin, der sich das Institut in seiner Arbeitsweise offenbar verpflichtet sieht („Eine weitere zentrale Strategie der EbM besteht darin, alle nach der Qualität ihres Designs und ihrer Durchführung angemessenen Studien zu einer Frage zu identifizieren und so den Stand des zuverlässigen Wissens zusammenzufassen...“, s. S. 7 Allgemeine Methoden V 4.2) und damit letztendlich auch nicht im Sinne der von einer zu bewertenden Therapie potentiell betroffenen Patienten.

Das in den Raum gestellte Risiko des Selektionsbias in der Begründung der geplanten Einschränkung hinsichtlich der Akzeptanz unaufgefordert eingereicherter Daten erscheint auch deshalb nicht stichhaltig, da es gemäß Abschnitt 7.2 des vorliegenden Entwurfs des Methodenpapiers (Überprüfung der Informationsbeschaffung) ohnehin zur Arbeitsweise des Instituts gehört, eingereichte Daten zu überprüfen – zumindest im Zusammenhang mit einer Dossierbewertung. Aber auch außerhalb von Dossierbewertungen bleibt es dem Institut jederzeit unbenommen zu prüfen, ob eine zu den unaufgefordert eingereichten Daten möglicherweise gegenteilige Evidenz verfügbar ist und diese dann gegebenenfalls entsprechend in seine Bewertung mit einzubeziehen, bzw. die Daten wegen mangelnder Qualität unberücksichtigt zu lassen und insofern einen potentiellen Bias zu vermeiden.

Darüber hinaus erscheint dieses Risiko auch insofern eher artifiziell, als keinerlei Angaben dazu gemacht werden, inwieweit sich aus der bisherigen Arbeit des Instituts stichhaltige Belege dafür finden lassen, dass es in der Vergangenheit zu derartigen Verzerrungen, wie sie eingangs als Risiko beschrieben wurden, gekommen ist.

Zu 8.1.3. „Rangordnung verschiedener Studienarten/ Evidenzgrade“

Das Institut führt zunächst aus, dass es verschiedene Ansätze gibt, Studientypen gemäß Ihrer Evidenzgrade in eine Rangfolge zu bringen, dass aber keines dieser Systeme allgemein akzeptiert und universell einsetzbar sei. Das Institut orientiert sich weitgehend an der Evidenzklassifizierung in der aktuellen Verfahrensordnung des Gemeinsamen Bundesausschusses¹³, die in der Arzneimittel-

¹³ Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. Link: https://www.g-ba.de/downloads/62-492-873/VerfO_2014-03-20.pdf, entnommen am 09.07.2014

Nutzenbewertungsverordnung des Bundesministeriums für Gesundheit in § 5 (6) wie folgt festgelegt ist¹⁴:

1. I a systematische Übersichtsarbeiten von Studien der Evidenzstufe Ib
2. I b randomisierte klinische Studien (RCT)
3. II a systematische Übersichtsarbeiten der Evidenzstufe IIb
4. II b prospektiv vergleichende Kohortenstudien
5. III retrospektiv vergleichende Studien
6. IV Fallserien und andere nicht vergleichende Studien
7. V Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte, nicht mit Studien belegte Meinungen anerkannter Experten, Konsensuskonferenzen und Berichte von Expertenkomitees.

Dementsprechend kommt unter den verschiedenen Studientypen randomisiert kontrollierten Studien (RCT) der höchste Evidenzgrad zu. Das Institut betont, dass darüber hinaus RCT auch nach ihrer Qualität eingestuft werden.

Nicht-randomisierte Studien teilt das Institut grundsätzlich in vergleichende und nicht vergleichende Studien ein. Das Institut führt aus, dass bei nicht randomisierten vergleichenden Studien neben dem Studiendesign vor allem die Maßnahmen zur Kontrolle von Confoundern zu prüfen sind, dies jedoch die Bewertung des Verzerrungspotentials und nicht den Evidenzgrad berührt.

Wie bisherige Nutzenbewertungen gezeigt haben, wurden in internationalen Entwicklungsprogrammen neuer Arzneimittel häufig andere Komparatoren verwendet als die unter nationalen Gesichtspunkten vom Gemeinsamen Bundesausschuss gesetzten zweckmäßigen Vergleichstherapien. In der Arzneimittel-Nutzenbewertungsverordnung¹⁵ findet sich hierzu in § 5 (5) der Hinweis:

„Liegen keine direkten Vergleichsstudien für das neue Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vor oder lassen diese keine ausreichenden Aussagen über einen Zusatznutzen zu, können verfügbare klinische Studien für die zweckmäßige Vergleichstherapie herangezogen werden, die sich für einen indirekten Vergleich mit dem Arzneimittel mit neuen Wirkstoffen eignen.“

Der pharmazeutische Unternehmer hat also in diesen Situationen die Option, indirekte Vergleiche gegen die vom Gemeinsamen Bundesausschuss gesetzte zweckmäßige Vergleichstherapie durchzuführen und vorzulegen. Vor diesem Hintergrund erscheint es dringend geboten, dass das Institut in 8.1.3 ausführt, wie es indirekte Vergleiche bzw. Netzwerk-Metaanalysen in die Evidenzhierarchie eingruppiert.

¹⁴ Bundesministerium der Justiz und für Verbraucherschutz. Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung – AM-NutzenV). Link: <http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf>, entnommen 09.07.2014

¹⁵ Siehe Fußnote 5.

Zu 8.3.3. „Beurteilung klinischer Relevanz“

Die im Abschnitt 8.3.3. (vormals 7.3.3) vorgenommene Ergänzung hinsichtlich patientenrelevanter Endpunkte sagt aus, dass eine Voraussetzung zur Erhebung patientenrelevanter Endpunkte mittels (komplexer) Skalen die Verwendung von validierten bzw. etablierten Instrumenten ist. (s. S. 189 Allgemeine Methoden V 4.2.) Der Begriff etabliertes Instrument bedarf einer näheren Definition und evtl. einer konkreten Aufzählung von etablierten Instrumenten inklusive aller Indikationen, in denen diese Instrumente eingesetzt werden können. In der Vergangenheit hat das IQWiG z.B. die nach unserer Meinung etablierten und standardisierten Instrumente EQ-5D (EuroQol; präferenzbasiertes Verfahren zur Messung der gesundheitsbezogenen Lebensqualität) und den SF-36 (Short Form (36) Gesundheitsfragebogen/ Erhebung der gesundheitsbezogenen Lebensqualität) nicht akzeptiert. Weitere Konkretisierungen erfolgen ebenfalls nicht, so dass sich in diesem Zusammenhang folgende Fragen stellen: Was sind weitere Voraussetzungen für die Berücksichtigung derartiger patientenrelevanter Endpunkte, da es im ergänzten Text „eine Voraussetzung“ und nicht „die Voraussetzung“ heißt? Wie wird seitens des Institutes vorgegangen, um zu beurteilen, ob es sich bei einem bestimmten Instrument um ein „validiertes“ und „etabliertes“ handelt? Insbesondere letztere Formulierung eröffnet in der vorliegenden, unspezifischen Form einen breiten Interpretationsspielraum für die Bewertung seitens des Institutes. Dies bringt hinsichtlich der Einbeziehung entsprechender geeigneter Instrumente in die Planung und Gestaltung klinischer Prüfungen ein hohes Maß an Unsicherheit mit sich. Welche generischen, nicht krankheitsspezifischen Instrumente zur Beurteilung der Lebensqualität (z.B. EQ-5D, SF-36) werden unter welchen Voraussetzungen als patientenrelevante Endpunkte anerkannt(s. oben)? Welche konkreten Voraussetzungen gelten für krankheitsspezifische Instrumente? Um dem Anspruch eines Methodenpapiers gerecht zu werden, wären an dieser Stelle detailliertere Aussagen dringend erforderlich.

Literaturverzeichnis

1. IQWiG 2014: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden. Entwurf für Version 4.2. vom 18.06.2014, Köln 2014.
2. IQWiG 2014a: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, https://www.iqwig.de/de/presse/pressemitteilungen/pressemitteilungen/ueberarbeitung_methodenpapier_iqwig_stellt_entwurf_fur_version_4_2_zur_diskussion.6134.html , entnommen am 14.07.2014
3. WHO. The WHO quality of life assessment (WHOQOL): Position paper from the WHO. Soc Sci Med 1995; 41: 1403-09
4. LC Doward, SP McKenna. Defining Patient Reported Outcomes. Value in health 2004; 7/1: S4-S8
5. M. Bullinger. Das Konzept der Lebensqualität in der Medizin – Entwicklung und heutiger Stellenwert. Z. Evid. Fortbild. Qual. Gesundh. wesen 2014; 108: 97-103
6. C. Bradley. The Diabetes treatment satisfaction questionnaire: DTSQ. In Bradley C, ed. Handbook of Psychology and Diabetes. A guide to psychological measurement in diabetes research and practice. Chur, Switzerland: Harwood Academic Publishers 1994; 11-133
7. C. Bradley, DS Gamsu. Guidelines for encouraging psychological well-being: report of a working group of the WHO regional office for Europe and IDF European region St. Vincent declaration action programme for diabetes. Diabet Med 1994; 11: 510 - 516
8. C. Bradley, CJB Gilbride. Improving treatment satisfaction and other patient reported outcomes in people with type 2 diabetes: the role of once-daily insulin glargine. Diabetes, Obesity and Metabolism 2008; 10/1: 50 – 65
9. European Medicines Agency. Reflection Paper on the use of patient reported outcome 4 (PRO) measures in oncology studies. 17.06.2014. EMA/CHMP/292464/2014
10. Rottenkolber D. Discrete-Choice-Experimente zur Messung der Zahlungsbereitschaft für Gesundheitsleistungen - ein anwendungsbezogener Literaturreview. Gesundh ökon Qual manag 2011; 16: 232-244
11. Mühlbacher A.C., Bethge S., Tockhorn A. Präferenzmessung im Gesundheitswesen: Grundlagen von Discrete-Choice-Experimenten. Gesundh ökon Qual manag 2012; 17: 159–172
12. Johnson F.R., Lanscar E, Marshall D. ISPOR Task Force Report: Constructing Experimental Designs for Discrete-Choice-Experiments. Value in health 2013; 16: 3-13
13. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. Link: https://www.g-ba.de/downloads/62-492-873/VerfO_2014-03-20.pdf , entnommen am 09.07.2014
14. Bundesministerium der Justiz und für Verbraucherschutz. Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung – AM-NutzenV). Link: <http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf> , entnommen 09.07.2014
15. s. Quelle 5

A.1.27 – Verband Forschender Arzneimittelhersteller e. V. (vfa)

**Stellungnahme des vfa zum Entwurf des IQWiG
„Allgemeinen Methoden Version 4.2“ vom 18.06.2014**

05.08.2014

Dr. Andrej Rasch/ Tina Orben / Dr. Ch.-Markos Dintsios

vfa. Die forschenden Pharma-Unternehmen
Hausvogteiplatz 13, D-10117 Berlin



Hausvogteiplatz 13
10117 Berlin
Telefon 030 206 04-0
Telefax 030 206 04-222
www.vfa.de

Inhaltsverzeichnis:

A.	Einführender Abschnitt (i-xv)/Abkürzungsverzeichnis.....	3
B.	Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen	3
C.	Kosten-Nutzen-Bewertung nach § 35b SGB V (Abschnitt 2.1.4)	4
D.	Potenzialbewertung (Abschnitt 2.1.5)	6
E.	Gesundheitsinformationen (Abschnitt 2.1.7)	7
F.	Patientenrelevanter medizinischer Nutzen und Schaden (Abschnitt 3.1)	7
G.	Nutzenbewertung medizinischer Interventionen, insbesondere Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens (Abschnitt 3.1.1) .	9
H.	Zusammenfassende Bewertung (Abschnitt 3.1.5).....	10
I.	Diagnostische Verfahren (Abschnitt 3.5)	12
J.	Potenzialbewertung (Abschnitt 3.8)	13
K.	Kosten-Nutzen-Bewertung medizinischer Interventionen (Abschnitt 4).....	15
L.	Informationsbeschaffung (Abschnitt 7)	33

Die Stellungnahme des Verbandes forschender Arzneimittelhersteller e.V. (vfa) bezieht sich sowohl auf methodische als auch auf formale Aspekte sowie einzelne Punkte zur Umsetzung der im Entwurf vom IQWiG zur **Stellungnahme freigegebenen „Allgemeinen Methoden Version 4.2“ vom 18.06.2014**. Die Stellungnahme erfolgt entlang des vom IQWiG vorgelegten Entwurfs und bezieht sich ausschließlich auf die durch das IQWiG als Änderung gekennzeichneten Textteile.

Seite 3/34

A. Einführender Abschnitt (i-xv)/Abkürzungsverzeichnis

Auf Seite xv lautet die Abkürzung für Quality-Adjusted Life Year (qualitätsadjustiertes Lebensjahr) versehentlichlicherweise QUALY statt QALY.

B. Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Hier bezieht sich das Methodenpapier unter 1.1 Gesetzliche Aufgaben u. a. auf *die Methoden und Kriterien für die Erarbeitung von Bewertungen auf der Grundlage der in den jeweiligen Fachkreisen anerkannten internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie. In Abschnitt 1.2 werden der Begriff der evidenzbasierten Medizin, seine Entwicklung und das dahinterliegende Konzept näher erläutert*. Es fällt auf, dass kein weiterer Bezug zu den in den jeweiligen Fachkreisen anerkannten internationalen Standards der Gesundheitsökonomie genommen wird, während in Abschnitt 1.2 explizit die evidenzbasierte Medizin nähergehend erläutert wird. Da nun im Entwurf 4.2 der Allgemeinen Methoden des IQWiG die Methoden für Kosten-Nutzen-Bewertungen (KNB) begrüßenswerterweise als eigenständiges neues Kapitel 4 integriert wurden, stellt sich die Frage nach einer analog zu den Standards der evidenzbasierten Medizin näheren Erläuterung des Begriffes Gesundheitsökonomie, seiner Entwicklung und des dahinterliegenden Konzeptes in einem hierfür eigenständigen Abschnitt 1.3. Ansonsten entsteht beim unvoreingenommenen Leser der Entwurfsversion 4.2 Allgemeine Methoden des IQWiG der Eindruck, dass die in den Fachkreisen anerkannten internationalen Standards der Gesundheitsökonomie eine mindere Rolle im Vergleich zu denen zur evidenzbasierten Medizin beim IQWiG spielen (Primat der evidenzbasierten Medizin) bzw. die unter Abschnitt 1.1 bereits aufgegriffenen gesetzlichen Aufgaben des IQWiG miss- oder eigeninterpretiert werden. Nach Auffassung des vfa sind wie aus den weiteren Ausführungen im Rahmen dieser Stellungnahme ersichtlich wird, im Gesetz neben der evidenzbasierten Medizin auch die gesundheitsökonomischen Grundsätze verankert, womit diesen eine ebenbürtige Wichtigkeit in den Methoden eingeräumt werden sollte. Im Zuge des AMNOG sei an dieser Stelle nochmals explizit darauf hingewiesen, dass das IQWiG sowohl nach § 139a (4) SGB V als auch § 35a (1) SGB V die Bewertung des medizini-

schen Nutzens nach den international anerkannten Standards der evidenzbasierten Medizin und die ökonomische Bewertung nach den hierfür maßgeblichen international anerkannten Standards, insbesondere der Gesundheitsökonomie bewerten soll. Nach § 35a (1) SGB V ist nicht nur für die Bewertung des (Zusatz-) Nutzens von Arzneimitteln mit neuen Wirkstoffen (siehe dazu auch §7 AM-NutzenV), sondern auch beispielsweise für die Bestimmung der zweckmäßigen Vergleichstherapie, die Anwendung internationaler Standards der evidenzbasierten Medizin und der Gesundheitsökonomie unerlässlich und deswegen maßgeblich. In der entsprechenden Rechtsverordnung (AM-NutzenV) unter § 7 Nutzenbewertung (2) AM-NutzenV wird auf den allgemein anerkannten Stand der medizinischen Erkenntnisse abgehoben und als Grundlage für die Beurteilung auf die internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie verwiesen. Damit ergibt sich, dass die internationalen Standards der Gesundheitsökonomie nicht nur für die Kosten-Nutzen-Bewertung (KNB) (§ 139a SGB V und § 35b SGB V) eingesetzt werden, sondern bereits im Zusammenhang mit der frühen Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach §35a SGB V zum Tragen kommen.

Seite 4/34

vfa-Position zusammenfassend: Die international anerkannten Standards der evidenzbasierten Medizin und der Gesundheitsökonomie sind ebenbürtig und werden sowohl vom Gesetzesgeber als auch vom Verordnungsgeber in den betreffenden Paragrafen des Sozialgesetzbuches V und der AM-NutzenV nebeneinander genannt. Dies sollte auch im Abschnitt 1 des Methodenentwurfs 4.2 und seinen Unterabschnitten entsprechend so gewürdigt werden.

C. Kosten-Nutzen-Bewertung nach § 35b SGB V (Abschnitt 2.1.4)

Auf Seite 22 heißt es mit Verweis auf Abbildung 4: „*Der Ablauf der Kosten-Nutzen-Bewertung nach § 35b SGB V ist in Abbildung 4 schematisch dargestellt. Alle Arbeitsschritte werden in Verantwortung des Instituts getätigt. Dabei werden regelhaft externe Sachverständige beteiligt. Bei Bedarf wird auch der wissenschaftliche Beirat des Instituts einbezogen. Der interne Qualitätssicherungsprozess ist in diesem Flussdiagramm nicht dargestellt*“. Es werden hierbei wie auch an anderen Stellen zu anderen IQWiG Produkten keine Gründe oder Regeln für die Beteiligung oder Einbeziehung des wissenschaftlichen Beirats angeführt und abstrakt auf einen Bedarf hierfür verwiesen. Dieser lässt sich allerdings nach dieser Formulierung kaum fassen und ist somit auch nicht überprüfbar bzw. belegbar.

Auf Seite 23 wird in der Abbildung 4 auf den Ablauf der Kosten-Nutzen-Bewertung nach § 35b SGB V eingegangen und auf die Einbeziehung der Patientenperspektive über Patientenvertreter und Patienten bei der In-

formationsbeschaffung sowie der wissenschaftlichen Bewertung zur hingewiesen. Allerdings können aufgrund von gegebenen Interessenkonstellationen unterschiedliche Motivationslagen bei institutionalisierten Patientenvertretern und betroffenen Patienten existieren. Institutionalisierte Patientenvertreter haben zwar keine Stimmrechte im G-BA, können aber dennoch bei Beratungen der Entscheidungsgremien insbesondere auch solcher von Erstattungsrelevanz im Rahmen der Selbstverwaltung im Gesundheitswesen partizipieren. Es sollte hier klar festgelegt werden, dass zur Einbeziehung der Patientenperspektive nur betroffene Patienten eingebunden werden. Dies gilt auch für den auf der Folgeseite (S. 24) stehenden Verweis des IQWiG zum Vorbericht, wonach bei der inhaltlichen Bewertung regelhaft die medizinische Expertise (über externe Sachverständige) und die Patientenperspektive (über Patientinnen und Patienten bzw. Patientenorganisationen) einbezogen werden. Da die Einbeziehung medizinischer Expertise und der Patientenperspektive mittels Fragekataloge bzw. Fragebögen erfolgt, und die Antworten externer Personen der Meinungsbildung beim IQWiG dienen, können dadurch die Ergebnisse der Kosten-Nutzen-Bewertung durchaus beeinflusst werden, weil sie zwangsläufig ein subjektives Moment beinhalten. Aus diesem Grund und um das Transparenzgebot, das sich das IQWiG selbst auferlegt hat, einzuhalten, sollten sowohl die Fragen als auch die Antworten veröffentlicht werden, damit der pharmazeutische Unternehmer (pU) und weitere Interessierte auch im Rahmen der Vorberichts-veröffentlichung dazu Stellung nehmen können. Außerdem wäre es wünschenswert, wenn in zukünftigen Verfahren mehrere Sachverständige beteiligt werden (z.B. per Delphi-Panel), um das potenzielle Risiko einzudämmen, die Nutzenbewertung auf gegebenenfalls verzerrte Einzelmeinungen aufzubauen. Die Einbeziehung medizinischer Fachgesellschaften sowie der Berufsverbände der betroffenen Arztgruppen für den Nutzenteil und entsprechender Fachgesellschaften für den gesundheitsökonomischen Part, wie beispielsweise der Deutsche Gesellschaft Gesundheitsökonomie (DGGÖ), ist vor diesem Hintergrund zu fordern.

Des Weiteren ist der vfa der Auffassung, dass ein externes Review der Qualitätssicherung der IQWiG Produkte dient und dass dieser Schritt bei allen IQWiG Vorberichten, also auch für diejenigen, die sich auf die Kosten-Nutzen-Bewertung (KNB) nach § 35b SGB V beziehen, regelhaft und nicht nur optional (siehe Abbildung 4: Ablauf der Kosten-Nutzen-Bewertung nach § 35b SGB V) beibehalten werden sollte. Dies v. a. in Anbetracht der Tatsache, dass die Kosten-Nutzen-Bewertungen einen sehr hohen Grad an Komplexität erzielen können – der Auftrag G09-01 zur Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin [G09-01] im Vergleich zu weiteren verordnungsfähigen medikamentösen Behandlungen, an welchem 14 externe Sachverständige sowie zumindest 16 IQWiG Mitarbeiter aktiv beteiligt waren, stellt ein eindrucksvolles Beispiel hierfür dar – und somit auch fehleranfällig sind. Das externe Review in seiner regelhaften Umsetzung erscheint einer optiona-

len Umsetzung somit überlegen. Um allerdings auch ein selektives Review bzw. eine interessen geleitete Auswahl der Reviewer zu vermeiden, wäre – dem Gedanken des Transparenzgebotes Folge leistend – eine klare Institutionalisierung des Reviews bezüglich der Auswahlkriterien der geladenen Reviewer, der einzuhaltenden Fristen, der Offenlegung der Beratungsfragen und der Ergebnisse wünschenswert.

Seite 6/34

Ferner sollte der unter Auftragserteilung in Klammern gesetzte Begriff „Scoping“ definiert oder erläutert werden. Dieser Begriff findet sich auch in anderen Abschnitten des Methodenentwurfs, allerdings bleibt die inhaltliche Definition unklar.

Bezugnehmend auf die Stellungnahmefrist zum Vorbericht, die vom IQWiG hier auf drei Wochen festgesetzt wird, sei angemerkt, dass generell für Kosten-Nutzen-Bewertungen aufgrund ihres Umfangs und des technischen Aufwandes, der durch die Überprüfung von Modellierungen und Simulationen bei den Stellungnehmenden entsteht, eine zweimonatige Stellungnahmefrist gefordert wurde. Bereits im Rahmen der Stellungnahmen und Anhörung zum Vorbericht Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin [G09-01] wurde dies gefordert und letztlich auch vom IQWiG eingeräumt. Eine dreiwöchige Stellungnahmefrist für die Öffentlichkeit erscheint definitiv zu kurz und nicht umsetzbar. Der vfa plädiert dafür, eine Frist von mindestens sechs Wochen vorzusehen.

vfa-Position zusammenfassend: Die Einholung externer Expertise bei der IQWiG Produkterstellung muss klaren Regeln hinsichtlich ihrer Vorgehensweise folgen und dem Transparenzgebot genügen. Sie sollte regelhaft systematisch vorgesehen sein und mithilfe valider und etablierter Instrumente sowie repräsentativ umgesetzt werden. Die Aussagen von medizinischen oder anderen relevanten wissenschaftlichen Fachgesellschaften sollten hierbei berücksichtigt werden. Bei Einbeziehung der Patientenperspektive sollten primär die Erfahrungen Betroffener berücksichtigt werden. Externe Reviews der IQWiG Produkte sollten regelhaft erfolgen. Der Begriff „Scoping“ sollte nähergehend erläutert werden. Bei umfangreichen Berichten, wie beispielsweise gesundheitsökonomische Evaluationen, die neben einer Nutzenbewertung zusätzlich eine Kosten-Nutzen-Bewertung enthalten, sollten die Stellungnahmefristen entsprechend verlängert werden.

D. Potenzialbewertung (Abschnitt 2.1.5)

Auf Seite 25 in Abbildung 5 wird der Ablauf der Erstellung einer Potenzialbewertung beschrieben und die Einbeziehung der medizinischen Expertise über externe Sachverständige für die Potenzialbewertung als optional dargestellt. Nach Auffassung des vfa sollte diese regelhaft vorgesehen

sein, da davon auszugehen ist, dass Medizinprodukte bzw. Prozeduren zumindest hinsichtlich ihrer Anwendung oder Umsetzung eine Darlegung der Erfahrung von entsprechenden medizinischen Experten und Patienten erfordern. Nur so kann eine umfassende Einschätzung zur Technologie getroffen werden. Weiter (Seiten 25f.) konkretisiert das Institut Eckpunkte einer Erprobungsstudie, welche ein optionaler Antragsinhalt sind, wenn der Antragsteller hierzu keine Angaben macht. Nach Auffassung des vfa müsste hierzu eine Liste vorgegeben werden, die auch eine begleitende Rationale enthält und ein konsistentes Verfahren des Instituts in solchen Fällen gewährleistet.

vfa-Position zusammenfassend: Bei der Potenzialbewertung sollte die Einbeziehung externer medizinischer Expertise regelhaft vorgesehen sein. Zu den Eckpunkten sollte, falls der Antragsteller keine Angaben macht, eine Liste mit begleitender Rationale vorgesehen sein.

E. Gesundheitsinformationen (Abschnitt 2.1.7)

Das IQWiG führt aus: *„Diese Informationen werden der Öffentlichkeit primär über die Website www.gesundheitsinformation.de zur Verfügung gestellt (auf Englisch unter informedhealthonline.org). Den Kern der Website bilden gesundheits- oder krankheitsbezogene Themen. Je nach Breite und Tiefe können innerhalb eines Themas unterschiedliche Textformate kombiniert werden“*. Da dem vfa keine weiteren Kanäle für diesen Zweck außer der Webseite des IQWiG www.gesundheitsinformation.de (bzw. deren englisches Pendant) bekannt sind, stellt sich die Frage nach dem „primär“ hier.

F. Patientenrelevanter medizinischer Nutzen und Schaden (Abschnitt 3.1)

Die umfangreichen Änderungen im Abschnitt 3.1 „Patientenrelevanter medizinischer Nutzen und Schaden“ hinsichtlich der Nutzendefinition und der zusammenfassenden Bewertung werden durch das Institut nicht weiter kommentiert und erläutert.

Im neuen Methodenentwurf 4.2 entfällt beispielsweise gänzlich die im Abschnitt 3.2.4 der Version 1.0 der KNB-Methoden „Nutzenbegriff“ enthaltene Formulierung. Diese lautete: *„Der Begriff „Nutzen“ wird in der gesundheitsökonomischen Literatur weit und eng verwendet. In seiner engen Bedeutung lehnt er sich an die EbM an und spiegelt den reinen medizinischen Nutzen zur Beurteilung einer Maßnahme wider. In seiner weiten Bedeutung umfasst der Begriff „Nutzen“ nicht nur die Effekte einer Intervention, sondern auch das, was in der ökonomischen Literatur im Allgemeinen als Wert bezeichnet wird. Darunter wird die präferenzbasier-*

te Bewertung des Nutzens verstanden." Ebenso entfällt die im Glossar der Version 1.0 der KNB-Methoden enthaltene Nutzendefinition: „(...) *In seiner weiten Bedeutung umfasst der Begriff „Nutzen“ nicht nur die Gesundheitseffekte einer Intervention per se, sondern berücksichtigt auch den Wert, den der Patient diesem Effekt zuschreibt (sogenannte Nutzwerte / Utilities). Diese Nutzwerte können zur Gewichtung von Effekten genutzt werden und hierdurch die Bedeutung eines Effektes aus Sicht der Betroffenen erhöhen oder reduzieren.*“ Aus welchen Gründen jegliche Methoden der Präferenzmessung der Patienten zur Gewichtung von Effekten sowie die weitere Definition des Nutzens (inkl. Nutzwerte) in der Definition des Nutzen des neuen Entwurfs 4.2 nicht mehr enthalten sind, lässt sich nicht nachvollziehen. Vielmehr verortet das IQWiG jegliche Ansätze der multikriteriellen Entscheidungsfindung und der Präferenzhebung (z.B. AHP oder CA) sowie Ansätze zur Verwendung von Maßen des Gesamtnutzens als Bestandteile einer ausschließlichen KNB nach §35b SGB V. Im Allgemeinen gilt es, so sehr auch die Integration der KNB in das Methodenpapier zu begrüßen ist, der Frage nachzugehen, weshalb es zwischen den Abschnitten 3 und 4 zu methodischen Unterschieden in der Nutzenbewertung kommt. Alternative Instrumente und Theorien der Nutzenbewertung, wie beispielsweise das Konzept der qualitätsadjustierten Lebensjahre und die Instrumente der multikriteriellen Entscheidungsanalyse (inklusive der Verfahren der Präferenzmessung), die bei der Nutzenbewertung zur Wirtschaftlichkeitsanalyse Anwendung finden (Abschnitt 4), können auch im Kontext der Abwägung von Nutzen und Schaden berücksichtigt werden (Abschnitt 3). Analogien werden zwar angedeutet, aber eine eindeutige methodische Stellungnahme zur Identifikation, Gewichtung und Aggregation von patienten-relevanten Endpunkten fehlt. Dies ist umso verwunderlicher, da Pilotprojekte des IQWiG hierzu seit längerer Zeit abgeschlossen sind und auch in verschiedenen Verfahren der frühen Nutzenbewertung eine Gewichtung von Endpunkten abgeleitet wurde (u.a. zu Afatinib).

Seite 8/34

Daraus ergeben sich allerdings mehrere schwerlich nachzuvollziehende Anpassungen, zu welchen weiter unten im Rahmen der vorliegenden Stellungnahme direkt Bezug genommen wird.

vfa-Position zusammenfassend: Eine in den vorherigen Versionen der Methodenpapiere (inklusive des Methodenpapiers zur Kosten-Nutzen-Bewertung) enthaltene Definition der Begrifflichkeiten zum Nutzen sollte wieder Eingang im aktuellen Methodenpapier finden. Methodische Unterschiede bei der Nutzenbewertung zwischen den verschiedenen Abschnitten des Methodenentwurfs werden weder begründet noch erschienen sie gerechtfertigt.

**G. Nutzenbewertung medizinischer Interventionen, insbesondere
Definition des patientenrelevanten medizinischen Nutzens
bzw. Schadens (Abschnitt 3.1.1)**

Seite 9/34

Bereits in der vorherigen Fassung der IQWiG Methoden 4.1 zählten sowohl Mortalität als auch Morbidität sowie gesundheitsbezogenen Lebensqualität ergänzend um den interventions- und erkrankungsbezogenen Aufwand sowie Behandlungs- bzw. Patientenzufriedenheit als patientenrelevant. Letztere wurde aus dem kompletten Methodenentwurf Allgemeine Methoden Version 4.2 vollständig gestrichen (neben Abschnitt 3.1.1 auch aus Abschnitt 3.2.4 usw.), ohne eine Rationale hierfür zu benennen.

Diese Änderung kann nicht nachvollzogen werden, da die Erhebung der Behandlungszufriedenheit einen wichtigen ergänzenden Endpunkt in vielen Indikationsfeldern darstellen kann und eine generelle Ausblendung der verfügbaren patientenberichteten Ergebnisse als inadäquat erscheint. Insbesondere ist auf die Inkonsistenz einer Streichung der Behandlungszufriedenheit durch Patienten als Zielgröße der Nutzenbewertung hinzuweisen, da das IQWiG selbst ebendiesen Endpunkt als einen patientenberichteten Ergebnisparameter in den zurückliegenden Nutzenbewertungen (z.B. zu Insulinanaloga) in die Bewertung eingeschlossen hat. Da es sich hierbei um eine grundlegende Änderung des Methodenpapiers handelt, ist der vfa der Auffassung, dass eine angemessene Rationale diesen Schritt begleiten sollte, die auch eine entsprechende Nachvollziehbarkeit der Reduktion der Zielgrößen gewährleistet. Es ist an dieser Stelle zu betonen, dass Patientenzufriedenheit kein Konstrukt ist, das inhaltsleer ist bzw. keine Anwendung findet. Aus den Erfahrungen im Rahmen der frühen Nutzenbewertung nach § 35a SGB V lässt sich beispielsweise ableiten, dass durchaus die Patientenzufriedenheit als Zielgröße für einen Nachweis eines (Zusatz-)Nutzens operationalisiert werden konnte. So ist Patienten- / Behandlungszufriedenheit in 13 Fällen von bisher über 90 Bewertungen als Endpunkt von pharmazeutischen Herstellern eingeschlossen worden (Mikrobielle Collagenase aus Clostridium histolyticum: Patientenzufriedenheit (5-Punkte-Likert-Skala); Extrakt aus Cannabis Sativa: SGIC; Fampridin: SGI (Subject Global Impression); Aclidiniumbromid: Behandlungszufriedenheit (Bedienung, Bevorzugung, Bereitschaft zur Weiterverwendung); Ingenolmebutat: TSQM; Teriflunomid: TSQM; Linaclotid: Behandlungszufriedenheit (5-Punkt-Ordinalskala); Emtricitabin/Rilpivirin/Tenofoviridisoproxil (neues AWG): HIV Treatment Satisfaction Questionnaire (HIVTSQ); Turoctocog alfa: HEMO-SAT; Linagliptin: DTSQ; Dapagliflozin + Metformin: DTSQ; Dapagliflozin: DTSQ; Vildagliptin: DTSQ). Vor dem Hintergrund der zuletzt in den Verfahren der frühen Nutzenbewertung betonten Rolle der „patient reported outcomes“ sollte die Option eines Einschusses solcher Studienergebnisse in die Nutzenbewertung nach Auffassung des vfa nicht aus dem Methodenpapier gestrichen werden.

Auf Seite 39 desselben Abschnitts streicht das IQWiG zudem folgende **Textpassage, welche die Definition des Nutzens konkretisiert: „Gemäß § 35b des SGB V sollen beim Patientennutzen insbesondere die Verbesserung des Gesundheitszustandes, eine Verkürzung der Krankheitsdauer, eine Verlängerung der Lebensdauer, eine Verringerung der Nebenwirkungen sowie eine Verbesserung der Lebensqualität angemessen berücksichtigt werden. Diese Nutzendimensionen sind durch die oben aufgeführten Zielgrößen abgebildet. So sind die Verbesserung des Gesundheitszustandes und die Verkürzung der Krankheitsdauer Aspekte der unmittelbaren krankheitsbedingten Morbidität, die Verringerung von Nebenwirkungen ist ein Aspekt der therapiebedingten Morbidität.“**

Es erscheint verwunderlich, dass das grundsätzliche Verständnis des Patientennutzens von den unterschiedlichen Verortungen in den jeweiligen Paragraphen des SGB V abhängig gemacht wird. Zudem wird insbesondere auch im § 2 Abs. 3 AM-NutzenV der Nutzenbegriff wie auch im § 35b SGB V näher konkretisiert: **„Der Nutzen eines Arzneimittels im Sinne dieser Verordnung ist der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustands, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität.“** Diese konkretisierenden Aspekte des Nutzens sollten daher nach Auffassung des vfa weiterhin im Methodenpapier beibehalten werden.

vfa-Position zusammenfassend: Die Reduktion der Zielgrößen und die damit parallel einhergehende Streichung der Patienten- bzw. Behandlungszufriedenheit bedarf einer nachvollziehbaren Begründung. Eine unterschiedliche Interpretation des Nutzens in Abhängigkeit von der sozialrechtlichen Verortung für die Nutzen- und Kosten-Nutzen-Bewertung ist nicht nachvollziehbar. Die konkretisierenden Aspekte des Nutzens sollten weiterhin im Methodenpapier beibehalten werden.

H. Zusammenfassende Bewertung (Abschnitt 3.1.5)

Zur Her- bzw. Ableitung von „Nutzen“ und „Schaden“ und ihrer Abwägung schreibt das Institut Folgendes: *„Eine Möglichkeit der gleichzeitigen Würdigung von Nutzen und Schaden ist die Gegenüberstellung der endpunktbezogenen Nutzen- und Schadenaspekte. Dabei werden die Effekte auf alle Endpunkte (qualitativ oder semiquantitativ wie in der frühen Nutzenbewertung nach § 35a SGB V) gegeneinander abgewogen mit dem Ziel, zu einer endpunktübergreifenden Aussage zum Nutzen bzw. Zusatznutzen einer Intervention zu kommen. Eine weitere Möglichkeit der gleichzeitigen Würdigung besteht darin, die verschiedenen patientenrelevanten Endpunkte zu einem einzigen Maß zu aggregieren oder über ihre Gewichtung zu einer Gesamtaussage zu kommen. Die gleichzeitige Würdigung von Nutzen und Schaden wird themenspezifisch konkretisiert“* und verweist

hierbei auf Abschnitt 4.3.3 hierbei. Eine Rationale zur Aufgabe im Rahmen der zusammenfassenden Bewertung der vormals (in Version 4.1 noch enthaltenen) alternativen Möglichkeiten über eine qualitative oder semi-quantitative Vorgehensweise hinaus, wie sie beispielsweise in Form von Summenscores oder Indizes oder mithilfe von Nutzwerten oder Gewichtsungsverfahren [Analytic Hierarchy Process (AHP), Conjoint Analyse (CA)] genannt wurde, fehlt gänzlich. Dieser klare und eindeutige Kurswechsel des Instituts ist nach Auffassung des vfa nicht nachvollziehbar, da es eher einen methodischen Rückschritt darstellt, wenn nun zusammenfassende Bewertungen ausschließlich qualitativ oder semi-quantitativ erfolgen sollen, während in vergangenen Versionen des Methodenpapiers weitere, auch rein quantitative Ansätze hierfür zumindest als Alternativen genannt wurden. Es wird dem Leser bzw. Anwender überlassen, auf Basis einer abstrakten Formulierung im Methodenentwurf, ohne irgendeine Handlungsanleitung bzw. methodische Vorgehensweise hierfür zu präsentieren, eine Methode zu identifizieren, deren mögliche Akzeptanz vom IQWiG nur rein spekulativ angenommen werden kann. Eine weitergehende Konkretisierung zu Ansätzen für eine zusammenfassende Bewertung ist hier notwendig. Nach Auffassung des vfa sollte sich das Institut für die Weiterentwicklung bzw. Adaptation vorhandener Methoden weiterhin einsetzen. **Ferner wird zur gleichzeitigen Würdigung von „Nutzen“ und „Schaden“** nicht mehr Bezug genommen, ob diese – wenn möglich – im Berichtsplan prospektiv erfolgen sollte oder nicht. Aus Sicht des vfa stellt vielmehr der Verweis auf die frühe Nutzenbewertung nach § 35a SGB V zur qualitativen bzw. semiquantitativen Abwägung der Endpunkte einen Versuch dar, das methodische Vorgehen formal Sozialbuchparagrafen zuzuordnen und nicht aus der wissenschaftlichen Fragestellung heraus eine entsprechende Methodenfindung bzw. -entwicklung zu betreiben. Ein rein qualitatives bzw. semiquantitatives Vorgehen, das hier hervorgehoben wird, läuft Gefahr, willkürlichen Werturteilen ausgesetzt zu sein. Wie bereits unter Abschnitt B) der vorliegenden Stellungnahme kommentiert, verweisen wir hier erneut auf den Stellenwert der internationalen Standards der Gesundheitsökonomie neben denen der evidenzbasierten Medizin für die frühe Nutzenbewertung nach § 35a (1) SGB V und § 7 (2) AM-NutzenV, und somit auch auf die Nutzwerte als Teil dieser Standards, die nun keine Nennung mehr in diesem Abschnitt finden. Nach Auffassung des vfa sollten zumindest Teile der nun entfernten Passagen, die sich zur **gleichzeitigen Würdigung von „Nutzen“ und „Schaden“ bezogen, weiterhin beibehalten** werden, wenn keine nachvollziehbare Rationale durch das Institut für ihre Entfernung angeboten werden kann. Des Weiteren möchte der vfa in **diesem Zusammenhang auf das „Benefit-Risk Methodology“ Projekt der EMA und das anschließende Arbeitspaket der Working Group 2 der gleichen Behörde hinweisen**, die weit über semiquantitative Vorgehen, geschweige denn qualitative Methoden hinausgehend bereits Stated Preference Risk-Benefit Trade-Offs und Verfahren der multikriteriellen Entscheidungsanalyse als relevante Alternativen zur „Nutzen-Schaden-Abwägung“ identifiziert haben und umzusetzen versuchen.

Seite 11/34

vfa-Position zusammenfassend: Eine weitergehende Konkretisierung zu Ansätzen für eine zusammenfassende Bewertung von „Nutzen“ und „Schaden“ ist notwendig. Das Institut sollte sich für die Weiterentwicklung bzw. Adaptation vorhandener Methoden weiterhin einsetzen. Rein qualitative oder semiquantitative Ansätze wie sie zurzeit im Rahmen der frühen Nutzenbewertung von Arzneimitteln durch das IQWiG umgesetzt werden, stellen bei Vorhandensein alternativer quantifizierender Methoden ein reduktionistisches Vorgehen dar. Die zusammenfassende Bewertung von „Nutzen“ und „Schaden“ kann nicht für die Nutzen- und Kosten-Nutzen-Bewertung unterschiedlich verlaufen.

I. Diagnostische Verfahren (Abschnitt 3.5)

Auf Seite 64 lautet es mit Verweis auf eine Publikation in ZEFQ: *„Diagnostische Verfahren zeichnen sich dadurch aus, dass ihr gesundheitsbezogener Nutzen (oder Schaden) im Wesentlichen erst dadurch zustande kommt, dass sich ihnen therapeutische oder präventive Verfahren anschließen. Die alleinige Gewinnung diagnostischer Informationen (ohne therapeutische oder präventive Konsequenzen) hat regelhaft keinen sozialrechtlich relevanten Nutzen“*. Dies ist nach Auffassung des vfa zu hinterfragen, kann beispielsweise durch eine geeignete Diagnostik mit anschließender Diagnosestellung eine diagnostische Odyssee vermieden werden, womit auch die Ausgaben für weitere unnötige Leistungen reduziert werden, so dass sich ein direkter Zusammenhang zur vom SGB V eingeforderten Wirtschaftlichkeit ergibt.

Auf Seite 66 äußert sich das Institut mit zwei Literaturverweisen: *„Insgesamt entscheidend ist weniger, inwieweit eine diagnostische oder prognostische Information einen aktuellen oder zukünftigen Gesundheitszustand feststellen kann, sondern dass diese Information auch prädiktive Bedeutung hat, also den höheren (oder geringeren) Nutzen einer Folgebehandlung vorhersagen kann“*. Der vfa kann sich der dahinter stehenden Rationale nicht anschließen, da ein prognostischer Biomarker auch für Therapieentscheidungen relevant sein kann. Bei einer sehr guten Prognose, könnte je nach Fallkonstellation beispielsweise auf eine Behandlung (mit schwerwiegenden Nebenwirkungen) verzichtet werden. Ferner ist doch in Zeiten der Aufklärung eine jegliche Information mit einem Eigenwert versehen. Übertragen auf eine medizinisch-diagnostische Fragestellung bedeutet das, dass selbst bei fehlender prädiktiver Bedeutung im Sinne einer fehlenden Vorhersage zum Nutzen einer Folgebehandlung durchaus eine diagnostische oder prognostische Information wertvoll für den Betroffenen sein kann, da er hierdurch selbst bei einer nicht vorhandenen Interventionsoption seine Lebensplanung dementsprechend nur alleine aufgrund der erhaltenen Information anders gestalten kann. Insbesondere bei genetischen Erkrankungen mit späterer Manifestation ist

dies zu beachten. Weiter lautet es auf derselben Seite zu den Anreicherungsdesigns *„Daher lassen sich aus solchen Designs nur dann tragfähige Schlüsse ziehen, wenn aufgrund anderer Informationen ausgeschlossen werden kann, dass ein in der randomisierten Patientengruppe beobachteter Effekt nicht auch in der nicht randomisierten Gruppe vorhanden gewesen wäre“*. Nach Auffassung des vfa sollte hier vom IQWiG Bezug zur Art und Aussagekraft dieser Informationen sowie eine Quelle für diesen Vorschlag bzw. dieses Vorgehen hinterlegt werden. Auf Seite 67 führt das IQWiG aus: *„Die bisherigen Ausführungen beziehen sich primär auf diagnostische Verfahren, die durch eine Erhöhung der Testgüte (also Sensitivität, Spezifität oder beides) mehr Patienten einer bestimmten therapeutischen Konsequenz zuführen. In diesen Fällen ist es regelhaft notwendig, durch eine Erfassung der gesamten diagnostisch-therapeutischen Behandlungskette den Einfluss des diagnostischen Verfahrens auf patientenrelevante Endpunkte zu untersuchen“*. Der vfa merkt hierzu an, dass dies an die Grenzen der Machbarkeit stößt, vor Allem in solchen Fällen, in welchen zwischen diagnostischer Maßnahme und therapeutischem Eingriff ein langer Zeitraum liegt bzw. mehrere therapeutische Alternativen zur Verfügung stehen, die nicht ohne Weiteres gleichzeitig im Rahmen einer gesamten diagnostisch-therapeutischen Behandlungskette überhaupt untersucht werden können.

vfa-Position zusammenfassend: Durch geeignete diagnostische Maßnahmen können weitere diagnostische Interventionen vermieden werden, womit sich auch ein direkter Zusammenhang zur vom SGB V eingeforderten Wirtschaftlichkeit ergibt. Prognostische Informationen haben für Therapieentscheidungen Relevanz und sind somit per se wertvoll. Selbst bei unsicherer prädiktiver Bedeutung haben diagnostische und prognostische Aussagen ihren Eigenwert, da sie betroffenen Patienten eine autonome Lebensplanung gewährleisten können. Für Anreicherungsdesigns sollte das IQWiG Art und Aussagekraft der Informationen zum Ausschluss nicht auch in den nicht randomisierten Gruppen vorhandener Effekte konkretisieren. Die Erfassung der gesamten diagnostisch-therapeutischen Behandlungskette zur Untersuchung des Einflusses diagnostischer Verfahren auf patientenrelevante Endpunkte stößt an die Grenzen der Machbarkeit.

J. Potenzialbewertung (Abschnitt 3.8)

Auf Seite 74 wird vom Institut Folgendes festgehalten: *„(...) aufgrund der besonderen Zielsetzung sind bei der Potenzialbewertung im Vergleich zur Nutzenbewertung deutlich niedrigere Anforderungen an die Evidenz zu stellen. Letztendlich besteht erst das Ziel der Erprobung darin, eine adäquate Datengrundlage für eine zukünftige Nutzenbewertung herzustellen. Dementsprechend lässt sich ein Potenzial insbesondere auch auf der Basis nicht randomisierter Studien begründen“* und weiter *„Im Gegensatz zur Nutzenbewertung wird im Rahmen von Potenzialbewertungen aufgrund*

der niedrigeren Anforderungen an die Evidenz eine erweiterte Bewertung der Ergebnissicherheit nicht randomisierter Studien vorgenommen". Es folgt anschließend vom Institut die Ergänzung neben den in Abschnitt 3.1.4 für randomisierte Studien genannten Stufen (hohe oder mäßige Ergebnissicherheit) folgende Grade von Ergebnissicherheit: 1) geringe Ergebnissicherheit: Ergebnis einer höherwertigen, nicht randomisiert vergleichenden Studie mit adäquater Kontrolle für Confounder (beispielsweise quasirandomisierte kontrollierte Studien, nicht randomisierte kontrollierte Studien mit aktiver Zuteilung der Intervention nach vorab geplanter Regel, prospektive vergleichende Kohortenstudien mit passiver Zuteilung der Intervention); 2) sehr geringe Ergebnissicherheit: Ergebnis einer höherwertigen, nicht randomisiert vergleichenden Studie (siehe vorheriger Punkt), jedoch ohne adäquate Kontrolle für Confounder oder Ergebnis einer sonstigen nicht randomisiert vergleichenden Studie (beispielsweise retrospektive vergleichende Kohortenstudien, historisch kontrollierte Studien, Fall-Kontroll-Studien) und 3) minimale Ergebnissicherheit: Ergebnis einer nicht vergleichenden Studie (beispielsweise einarmige Kohortenstudien, Verlaufsbeobachtungen oder Fallserien, Querschnittsstudien oder sonstige nicht vergleichende Studien). Weiter lautet es: „Da auch hochwertige, nicht randomisierte Studien ein erhebliches Verzerrungspotenzial bergen, muss bei der Ableitung eines Potenzials aus solchen Studien geprüft werden, ob die vorliegenden Studien bezüglich der interessierenden Intervention Unterschiede in einer Größenordnung zeigen, die erwarten lassen, dass durch geeignete zukünftige Studien ein Nutzen nachgewiesen werden kann, und nicht mehr allein durch einen durchschnittlich erwartbaren Einfluss von Verzerrung (Bias) erklärt werden können". Das Institut macht im Nachgang das Potenzial abhängig von der Größe der Effekte bei geringer bis minimaler Ergebnissicherheit: „Daher ergibt sich ein Potenzial insbesondere dann, wenn Studien geringer Ergebnissicherheit mindestens kleine Effekte zeigen, wenn Studien sehr geringer Ergebnissicherheit mindestens mittlere Effekte zeigen oder wenn Studien minimaler Ergebnissicherheit mindestens große Effekte zeigen. Als ungefähre Grenzen zwischen kleinen, mittleren und großen Effekten können für das relative Risiko Werte von 0,8 und 0,5 dienen". In § 5 (6) AM-NutzenV werden folgende Evidenzstufen genannt: 1) Ia systematische Übersichtsarbeiten von Studien der Evidenzstufe Ib; 2) Ib randomisierte klinische Studien; 3) IIa systematische Übersichtsarbeiten der Evidenzstufe IIb; 4) IIb prospektiv vergleichende Kohortenstudien; 5) III retrospektiv vergleichende Studien; 6) IV Fallserien und andere nicht vergleichende Studien und 7) V Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte, nicht mit Studien belegte Meinungen anerkannter Experten, Konsensuskonferenzen und Berichte von Expertenkomitees. Hier stellt sich nun die Frage, wieso nicht zumindest bei Evidenzstufen ab IIa zur Bestimmung der Ergebnissicherheit und der daraus abzuleitenden Aussage zum (Zusatz-)Nutzen ähnlich verfahren wird, wie bei der Potenzialbewertung. Dies v. a. in Anbetracht der Tatsache, dass sowohl laut

§ 35a SGB V und nach AM-NutzenV bei der frühen Nutzenbewertung von neuen Arzneimitteln durchaus auch andere Studien als RCTs (z.B. einarmige Studien, historische Vergleiche) einbezogen werden können und dies auch in einigen Fällen so gehandhabt wurde. Auch bei Nicht-Orphan Drugs wurden schon einarmige Zulassungsstudien vorgelegt und zum Beschluss eines Zusatznutzens akzeptiert (siehe Sofosbuvir Beschluss 17.7.2014). Auch hier könnte die Aussage zum Vorliegen eines (Zusatz-)Nutzens auf Basis der Größe der Effekte neben der Ergebnissicherheit erfolgen und somit eine Handlungsanleitung für diese Konstellationen im Methodenpapier in Analogie zur Potenzialbewertung festgehalten werden.

Seite 15/34

vfa-Position zusammenfassend: Für nicht-randomisierte Studien im Rahmen der Nutzenbewertung von Arzneimitteln könnte die Aussage zum Vorliegen eines (Zusatz-)Nutzens auch auf Basis der Größe der Effekte neben der Ergebnissicherheit in Analogie zur Potenzialbewertung erfolgen.

K. Kosten-Nutzen-Bewertung medizinischer Interventionen (Abschnitt 4)

Hinsichtlich der Methodendiskussion um die Wertigkeit des Effizienzgrenzenansatzes als Basis für Kosten-Nutzen-Bewertungen medizinischer Interventionen durch das IQWiG und ohne diese Diskussion erneut aufrollen zu wollen, verweist der vfa auf die eingegangenen Stellungnahmen zu den Methoden der Kosten-Nutzen-Bewertung (Version 1.0) sowie zum Vorbericht Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin [G09-01].

4.1 Einführung

Im einführenden Abschnitt 4.1 wird vom Institut Folgendes angegeben: *„Gemäß SGB V bestimmt das Institut auftragsbezogen über die Methoden und Kriterien für die Erarbeitung von Kosten-Nutzen-Bewertungen (KNB) auf der Grundlage der in den jeweiligen Fachkreisen anerkannten internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie“*. Um etwaigen Missverständnissen vorzubeugen, weist der vfa darauf hin, dass auch die Methoden zur Nutzenbewertung nach § 35a SGB V mit Verweis auf AM-NutzenV, wie bereits in Abschnitt B) der vorliegenden Stellungnahme erläutert, auf anerkannten internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie basieren und somit auch keine Trennung der Standards nach unterschiedlichen Disziplinen zwischen § 35a SGB V, § 35b SGB V und § 139a SGB V erfolgen kann.

4.1.1 Rechtliche Grundlagen für die Kosten-Nutzen-Bewertung nach SGB V

Seite 16/34

Das IQWiG führt aus, dass das BMG im Jahr 2008 zum Stellenwert von KNB in Berufung auf § 12 SGB V, dem sog. Wirtschaftlichkeitsgebot, folgendermaßen Stellung nahm: *„Maßgebend für die Verordnungsentscheidung des Arztes kann nur sein, ob das Arzneimittel zur Erreichung des Therapieziels, d. h. im jeweiligen Anwendungsgebiet, zweckmäßig und wirtschaftlich ist (§ 12 SGB V [78]).“* Es schlussfolgert daraus, dass die KNB einen Vergleich der Effizienz von Interventionen in einem Indikationsgebiet ermöglichen muss, ohne die Frage einer indikationsübergreifenden Priorisierung im Gesundheitssystem zu beantworten. Unabhängig von der Lesart der Stellungnahme des BMG suggeriert hier das IQWiG, dass ein deutscher Sonderweg für die KNB aufgrund gesetzlicher Vorgaben (§ 12 SGB V) eingeschlagen werden muss, der sich von den internationalen Standards der Gesundheitsökonomie diesbezüglich unterscheidet. Diese Logik besteht nicht. Der vfa schlägt vor, in Version 4.2 der Allgemeinen Methoden indikationsübergreifende Wirtschaftlichkeitskriterien nach den internationalen Methoden der Gesundheitsökonomie zu berücksichtigen.

4.1.3 Zeithorizont

Da bei einer KNB nach § 35b SGB V Vorgaben durch den G-BA zur Durchführung von entsprechenden Versorgungsstudien gemacht werden können, sollte der Zeithorizont gegebenenfalls auch kürzer als die durchschnittliche Studiendauer gewählt werden können, wenn beispielsweise eine Interimsauswertung erwünscht ist, aus deren Ergebnissen die Weiterführung der Versorgungsstudie abhängig gemacht wird.

4.1.4 Auswahl der Komparatoren

Da in einer Wirkstoffklasse Arzneistoffe mit pharmakologischen Gemeinsamkeiten zusammengefasst werden und Wirkstoffklassen primär der Abgrenzung und Ordnung der in ihnen erfassten Stoffe dienen, schlagen wir die Abänderung des allgemeineren Begriffs Substanzklasse, der auch Nichtarzneimittel beinhaltet (z.B. Herbizide) in den terminologisch gängigeren Begriff Wirkstoffklasse für Arzneimittelbewertungen vor. Ferner ist **vom IQWiG nach Auffassung des vfa der Terminus samt Adjektiv „ausreichende Homogenität“ näher in den Methoden zu spezifizieren. Auf Basis von G09-01 zur Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetine, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medikamentösen Behandlungen zumindest, ist als Kriterium zur Bildung von Wirkstoffklassen wie folgt definiert worden: „In den vorgeschalteten Nutzenbewertungen wurden endpunktspezifisch Wirkstoffe zu Wirkstoffklassen als Komparatoren zusammengefasst, sofern keine durch einzelne Wirkstoffe bedingte bedeutsame Heterogenität in den paarweisen Meta-**

Analysen vorlag". Nach Auffassung des vfa sollte fehlende Heterogenität nicht mit vorliegender Homogenität gleichgesetzt werden bzw. auf die fehlende Heterogenität in diesem Abschnitt abgestellt werden.

Seite 17/34

4.1.5 Versorgungspfad

Das IQWiG führt folgendes aus: *„für jede KNB sollte(n) zunächst (ein) Versorgungspfad(e) für das Indikationsgebiet entwickelt werden. Dieser Versorgungspfad dient als Basis, um das Modell zu erstellen. Weiterhin basieren auf dem Versorgungspfad auch die Recherchen nach Daten zu Kosten und weiteren notwendigen Daten für das Modell“*. Das Modell wird allerdings hier nicht näher beschrieben bzw. abgegrenzt und der Leser erfährt erst unter 4.1.6 Modell, was darunter gemeint ist bzw. unter 4.2.1 **Grundlagen der Modellierung, wonach unter „Modell“ ein gesundheitsökonomisches Entscheidungsmodell subsumiert wird**. Somit müsste der Terminus *„chronologisch“* bereits unter 4.1.5 eingeführt werden bzw. die Gliederung der Inhalte einer anderen Reihenfolge unterstellt werden.

4.1.6 Modell

Die Aussage, dass Piggy-back-Studien sehr selten vorliegen, sollte mit einem entsprechenden Zitat untermauert werden. Wir würden uns gleichzeitig auch eine Überarbeitung der Formulierungen wünschen.

4.1.7 Datengrundlage

Laut IQWiG werden Grundlagen und Bewertung von Daten, die auf der Nutzenseite in eine KNB einfließen, in Abschnitt 3.3 erläutert und für das Maß des Gesamtnutzens werden einzelne Studien oder Datenerhebungen (siehe auch Abschnitt 4.3.3) verwendet. Abschnitt 3.3 bezieht sich explizit auf die Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V. Nach § 139a SGB V aber ist das IQWiG u. a. auch mit Erstellung von wissenschaftlichen Ausarbeitungen, Gutachten und Stellungnahmen zu Fragen der Qualität und Wirtschaftlichkeit der im Rahmen der gesetzlichen Krankenversicherung erbrachten Leistungen beauftragt und somit kann es auch eine KNB für nichtmedikamentöse Verfahren durchführen, womit dann der explizite Verweis auf Abschnitt 3.3 hinfällig wäre, es sei denn das Institut übernimmt die dort dargestellten Vorgaben auch im Rahmen der Bewertung von nichtmedikamentösen Verfahren, was allerdings seinen eigenen Methoden und seinem Auftrag widersprechen würde. Somit erachtet der vfa den Verweis auf Abschnitt 3.3 bzgl. der Nutzenseite der KNB als unangebracht an und schlägt vor, die Nutzenseite in Abhängigkeit von dem jeweiligen Erkenntnisobjekt entweder unterschiedlich zu operationalisieren oder die KNB um Ausführungen zur Nutzenseite bei nichtmedikamentösen Verfahren und für die Bewertung von Arzneimitteln nach § 139a SGB V (also nicht ausschließlich im Rahmen der frühen Nutzenbewertung nach § 35a SGB V) zu ergänzen. Hinsichtlich des Verweises zum

Maß des Gesamtnutzens (Abschnitt 4.3.3) stellt sich die Frage nach der spezifischen Lokalisierung dieses Vorgehens nur im Rahmen der KNB, da ein Gesamtmaß des Nutzens durchaus auch für die Nutzenbewertung von Arzneimitteln nach § 35a SGB V relevant sein könnte. Der Vorschlag des vfa lautet hierzu, Ansätze zur Herleitung eines Gesamtnutzens in den alten Abschnitt 3.1.5 Zusammenfassende Bewertung zurück zu überführen, wo sie auch hingehören, da diese Fragestellung nicht alleine für Kosten-Nutzen-Bewertungen von Relevanz ist und in der Vorgängerversion der Allgemeinen Methoden (Version 4.1) zurecht unter dem Abschnitt Zusammenfassende Bewertung firmierte.

Weiterhin lautet es in diesem Unterabschnitt: *„Auswertungen von Sekundärdaten sollten sich an den Leitlinien und Empfehlungen zur guten Praxis Sekundärdatenanalyse ausrichten [19]. Insbesondere sollen die Auswahl der Datenbasis, die Größe und relevante Merkmale der Stichprobe und der Studienpopulation (inkl. Ein- und Ausschlusskriterien), die statistischen Methoden und die Kontrolle von Störgrößen transparent beschrieben und begründet werden. Die Generalisierbarkeit und die Repräsentativität der Ergebnisse sollten erläutert werden. Die einzelnen Auswertungsschritte müssen nachvollziehbar sein; Plausibilitätskontrollen sind sicherzustellen. Werden Leitlinien verwendet, sollten diese aus dem deutschen Gesundheitssystem stammen und möglichst evidenzbasiert sein. Darunter werden Leitlinien verstanden, deren Empfehlungen auf einer systematischen Literaturrecherche beruhen, grundsätzlich mit einer Evidenz- und/oder Empfehlungseinstufung (Level of Evidence [LoE] und / oder Grade of Recommendation [GoR]) versehen sind und mit den Referenzen der ihnen zugrunde liegenden Primär- und / oder Sekundärliteratur verknüpft sind (modifiziert nach AGREE)“*. Bereits aus bisherigen Verfahren der frühen Nutzenbewertung ist bekannt, dass deutsche Leitlinien im Vergleich zu europäischen oder weiteren internationalen Leitlinien (ASCO, ESMO) je nach Indikation oftmals zeitlich hinterherhinken. Eine ausschließliche Orientierung an nationale Leitlinien birgt somit die Gefahr überholte Empfehlungen als Datengrundlage für den gesetzten Versorgungskontext geltend zu machen und aktuellere Entwicklungen, die zwar noch nicht in Leitlinien umgesetzt wurden, sich aber dennoch in der Versorgungsrealität wieder finden, zu ignorieren. Das IQWiG selbst nimmt mit Verweis auf entsprechende Literatur im Abschnitt 6.3.4 Bezug auf die Aktualität von Leitlinienempfehlungen. Der vfa schlägt aus diesem Grund vor, den Aktualitätsgrad der Leitlinien und die Versorgungsrealität bei der Bestimmung der zu verwendenden Datengrundlage mit zu berücksichtigen.

4.1.9 Interpretation der Ergebnisse

Bei der Definition der Effizienzgrenze bedient sich das IQWiG einer Negativdefinition: *„Die Effizienzgrenze zeigt ineffiziente Interventionen auf (z. B. solche, die sowohl teurer als auch in Bezug auf einen Endpunkt von geringerem Nutzen sind als andere Interventionen)“*. Genauso gut zeigt

die Effizienzgrenze aber auch effiziente neben ineffizienten Interventionen auf. Wir schlagen deshalb hier vor, eine Definition der Effizienzgrenze als zusammengesetzte Verbindungssegmente der effizienten Interventionen in einer grafischen Darstellung des Nutzens und der Kosten anhand eines indikationsspezifischen und endpunktbestimmten Koordinatensystems einzuführen und erst im Nachgang auf ihre Eigenschaften einzugehen. Weiter lautet es in diesem Abschnitt, dass aus der Effizienzgrenze selbst nicht direkt ableitbar ist, welche Kosten für diese Intervention angemessen sind. Auf der nächsten Seite (Seite 80) wird mit Verweis auf Abbildung 9 erläutert, dass bei einem gegebenen Nutzen einer zu beurteilenden Intervention solche Kosten-Nutzen-Verhältnisse als angemessen angesehen, die gemessen an der Effizienzgrenze nicht zu einer Verschlechterung der Effizienz in einem gegebenen Indikationsgebiet führen. Da Angemessenheit ein unbestimmter Rechtsbegriff ist, kann die Definition dessen, was auf Basis der Effizienzgrenze als angemessen angesehen wird, nur ein rein technisches Hilfskonstrukt bzw. ein Behelf sein, der aufgrund der Umsetzung des Ansatzes der Effizienzgrenzen nicht ohne Weiteres Bestand hat bzw. haben muss. Damit wird eine statische Betrachtung im Sinne einer technischen Effizienz als Angemessenheitskriterium eingeführt, ohne auf die alloкатive Effizienz abzustellen. Da aber mit Hilfe der Effizienzgrenze Erstattungsempfehlungen des IQWiG an die Selbstverwaltung erfolgen sollen, greift diese Betrachtungsweise zu kurz, da von einer rein statischen Zahlungsbereitschaft anhand der technischen Effizienz ausgegangen wird. Nach Auffassung des vfa ist außerdem die Operationalisierung des unbestimmten Rechtsbegriffes Angemessenheit nicht rein algorithmisch möglich. Weitere Faktoren, die die Angemessenheit triggern können sind beispielsweise das Indikationsgebiet, der Innovationsgrad der Interventionen, der ungedeckte Bedarf oder die entsprechende Krankheitslast. Selbst mit dem Ansatz der Effizienzgrenzen ist es nicht möglich, eine Angemessenheit mit einer simplen Extrapolation zu schlussfolgern. Denn die Effizienzgrenze stellt zum jeweiligen Betrachtungszeitpunkt über die Steigung ihrer Verbindungssegmente die entsprechenden indikationsbezogenen endpunktspezifischen Zahlungsbereitschaften anhand der Erstattungsbeträge für die untersuchten Interventionen dar. Ob sie als Handlungsempfehlung bzw. entscheidungsrelevanter Korridor oder Bereich die Extrapolation alleine geeignet ist, bleibt offen und unbeantwortet angesichts der weiter oben genannten Einflussfaktoren auf die Zahlungsbereitschaft des Gesundheitssystems. In Anbetracht dieser Ausführungen ist der vfa der Auffassung, dass die Effizienzgrenze durchaus Ineffizienzen für bereits im System befindliche Interventionen aufzeigen kann. Eine Handlungsempfehlung als Operationalisierung der Angemessenheit durch Extrapolation des letzten Segmentes für Interventionen, die im Nord-Ost-Quadranten der Kosten-Effektivitäts-Ebene liegen, ist nicht zielführend und nicht ohne Weiteres umsetzbar.

Seite 19/34

Weiter lautet es in der Entwurfsversion 4.2: *„Wenn ein Maß des Gesamtnutzens festgelegt wird (siehe Abschnitt 4.3.3), ist dieses als primäres*

Ergebnis anzusehen. Ist die Bestimmung mehrerer Effizienzgrenzen für die Bewertung einer Intervention erforderlich, so bleibt es dem Entscheidungsträger unter Beachtung der Relevanz der patientenrelevanten Endpunkte vorbehalten, eine Gewichtung vorzunehmen. Eine ähnliche Vorgehensweise kann auch bei Zulassung einer Intervention in mehreren Indikationsgebieten erwogen werden". Der vfa kann dieser Argumentation nicht widerspruchlos folgen. Zwar obliegt dem Entscheidungsträger die Verantwortung für die Nutzenbewertung und die Aussagen zur Wirtschaftlichkeit, aber eine Gewichtung der patientenrelevanten Endpunkte, ohne jegliche Referenzpunkte bzw. Rationale als Werturteil stellt auch für den Entscheidungsträger eine schier nicht zu erfüllende Aufgabe, will er sich nicht dem Vorwurf der Willkürlichkeit preisgeben. Schlimmer noch überlässt das IQWiG dem Entscheidungsträger mit einer ähnlichen normativen Vorgehensweise bei Zulassung in mehreren Indikationsgebieten der zu bewertenden Intervention deren Gewichtung vorzunehmen, was wiederum jeglicher Ratio entbehrt, da kein Modus dafür genannt wird, sondern auf die Relevanz der patientenrelevanten Endpunkte wiederum rekurriert wird, während sich hier objektivierbar über Zielpopulationsgrößen, Verordnungsvolumina oder -anteile durchaus quantifizierbare Gewichtungsmodi ausfindig machen lassen, die sinnvoll umsetzbar sind. Ferner wird im nachfolgenden Abschnitt auf die Zumutbarkeit Bezug genommen. Diese stellt wiederum einen unbestimmten Rechtsbegriff dar, der zusammen mit dem ebenfalls unbestimmten Rechtsbegriff der Angemessenheit in Abschnitt 1 nach Auffassung des vfa samt den entsprechenden legalen Eigenschaften im Methodenentwurf eingeführt werden sollte.

4.2.1 Grundlagen (Unterabschnitt zu Abschnitt 4 Modellierung)

Auf Seite 81 definiert das IQWiG die gesundheitsökonomische Modellierung wie folgt: *„Gesundheitsökonomische Modelle sind somit, wie mathematisch-formalisierte Modelle, auch eine vereinfachte Abbildung der Realität. Durch eine bewusste Reduktion der Komplexität auf die für das Entscheidungsproblem relevanten Entscheidungsfaktoren und -variablen wird zudem analytische Klarheit geschaffen“*. Nach Auffassung des vfa ist der letzte Satz apodiktisch. Unser Vorschlag lautet: [...] wird zudem der Versuch unternommen, analytische Klarheit zu schaffen.

4.2.5 Modelldokumentation und Modellvalidierung (B Allgemeine Dokumentation)

Unter dem fünften Bullet führt das IQWiG aus: *„Auflistung aller Annahmen hinsichtlich der Datenquellen und der Modellstruktur. Besonders wichtig ist eine detaillierte Darstellung jeglicher Annahme und verwendeten Technik zur Projektion über den Zeitraum hinaus, für welchen die Daten gelten.“* Hierzu merkt der vfa an, dass die Annahmen, die in das Modell einfließenden Daten und seine Struktur betreffen und weniger deren Quelle. Somit sollte auch darauf direkt Bezug genommen werden.

4.2.5 Modelldokumentation und Modellvalidierung (C Technische Dokumentation und elektronische Version des Modells)

Das Institut fordert, dass das KNB-Modell für der Öffentlichkeit vollumfänglich und voll interaktiv zur Benutzung zur Verfügung gestellt werden soll. Aus Sicht des vfa ist diese Information ein Teil von Modul K5, welches nach Anlage VIII.a des G-BA (Format und Gliederung des Dossiers, einzureichende Unterlagen, Vorgaben für technische Standards) vom 14.11.2013 vertrauliche Inhalte einschließt und somit nicht per se für die Öffentlichkeit verfügbar ist. Diese Forderung des IQWiG ist kritisch zu hinterfragen, da sie intellektuelle Eigentumsrechte von Modellentwicklern, die als Dienstleister für die pharmazeutische Industrie fungieren können, verletzt und den Vorgaben des G-BA zu Modul K5 nicht entspricht.

4.3 Nutzen

Wie bereits unter 4.1.7 Datengrundlage ausgeführt, führt der Verweis auf Abschnitt 3.3 hinsichtlich der Bestimmung des (Zusatz-)Nutzens in die Irre, wenn keine KNB nach § 35a (5a) bzw. § 35b SGB V, sondern nach § 139a SGB V erfolgt. Dies ist beispielsweise für nicht-medikamentöse Verfahren der Fall, für welche oft aufgrund der rechtlichen Voraussetzungen (CE Zertifizierung) keine Zulassungsstudien vorliegen und somit die Evidenz zum Nutzen nicht auf dem Evidenzlevel der Arzneimittel generiert werden kann.

4.3.2 Endpunkte

Hierzu führt das IQWiG aus, dass *„wenn mehrere patientenrelevante Endpunkte nebeneinander dargestellt werden, wird für jeden patientenrelevanten Endpunkt eine eigene Effizienzgrenze erstellt. Andernfalls wird der Nutzen aggregiert zu einem einzigen Maß des Gesamtnutzens, das anschließend in einer Effizienzgrenze abgetragen wird. Ein Maß des Gesamtnutzens ist in einer ganz allgemeinen Definition eine Aggregation der Bewertung von Nutzen und Schaden in einer Größe, wobei unterschiedliche patientenrelevante Endpunkte zu einem einzigen Maß zusammengefasst werden“*. Nach Auffassung des vfa wird hier nicht klar und deutlich herausgearbeitet, dass der Ansatz der Effizienzgrenze eigentlich erschöpfend und ganzheitlich angelegt ist, was die Einbeziehung der patientenrelevanten Endpunkte anbelangt. D. h., dass für die Aggregation zu einem einzigen Maß alle patientenrelevanten Endpunkte einbezogen werden müssen, weil sonst kein aussagekräftiges Maß zum Gesamtnutzen abzuleiten ist bzw. sich Verzerrungen ergeben können, die das Ergebnis der KNB nicht verwertbar gestalten.

4.3.3 Maß des Gesamtnutzens

Seite 22/34

Wie bereits weiter oben angeführt, erachtet der vfa die Vorgehensweise zur Herleitung eines Maßes für den Gesamtnutzen nicht nur für Kosten-Nutzen-Bewertungen, sondern auch im Rahmen von zusammenfassenden Bewertungen als relevant und schlägt deswegen eine entsprechende Abhandlung auch unter Abschnitt 3.1.5 vor. Laut IQWiG gilt folgendes Vorgehen: *„Wenn der G-BA für eine KNB nach § 35b Absatz 1 Satz 2 SGB V das Maß des Gesamtnutzens festlegt (siehe Abschnitt 4.9), werden ein entsprechendes Instrument und ggf. die dafür festgelegten Erhebungsmethoden oder eine schon festgelegte Gewichtung von Endpunkten nach den Vorgaben im Auftrag eingesetzt. Die Ergebnisse sollten dem Entscheidungsträger zusammen mit der Endpunktgewichtung zur Verfügung gestellt werden. Hierdurch ergibt sich für den Entscheidungsträger die Option, einen aus mehreren zusatznutzenbereinigten Erstattungspreisen gewichteten Erstattungsbetrag zu verhandeln“*. Allerdings stellt sich für den Leser erneut die Frage nach der normativen Wertentscheidung des Entscheidungsträgers, die einer wissenschaftlichen Grundlage entbehrt, insofern diese Gewichte, die beispielsweise mittels Präferenzenerhebung generiert wurden, evtl. auch aus Interessenkalkülen zur Disposition gestellt werden, um aus mehreren zusatznutzenbereinigten Erstattungspreisen einen gewichteten Erstattungsbetrag zu verhandeln. Nach Auffassung des vfa steht einer Verhandlungsoption nach § 130b SGB V für Vereinbarungen zwischen dem Spitzenverband Bund der Krankenkassen und pharmazeutischen Unternehmen über Erstattungsbeträge für Arzneimittel sowie so ein algorithmisches Gewichtungsvorgehen entgegen, auch wenn wissenschaftlich begründet und umgesetzt. Für andere Preisfindungen allerdings, welche keinen Verhandlungsweg sozialrechtlich vorsehen, wäre der vom IQWiG vorgeschlagene Modus nicht zielführend, da hiermit der **Selbstverwaltung die Möglichkeit nach „eigenem Ermessen“** Preise bzw. Erstattungsbeträge zu gestalten, Tür und Tor geöffnet würde. Ferner ist aus vfa Sicht eine schon festgelegte Gewichtung von Endpunkten nach den Auftragsvorgaben nicht nachvollziehbar, da diese Vorgaben wiederum auf einer Rationale fußen müssen, die aber zumindest nach dieser o.g. Formulierung nicht immer für eine Gewichtung erkennbar ist. Abschließend sei hier noch darauf hingewiesen, dass die theoretische Verortung bzw. die Debatte hierüber nicht dadurch abgetan werden kann, dass auf das Sozialbuch und seine Terminologie **„Gewichtung“** verwiesen werden kann, da durchaus die entsprechenden Begrifflichkeiten und der Stellenwert der Erhebungsmethoden von großer Relevanz im Rahmen der Kosten-Nutzen-Bewertung, aber auch einer Herleitung einer zusammenfassenden Aussage zur Nutzenbewertung in Abhängigkeit von der untersuchten Fragestellung sein können, und die Frage des Welfarismus bzw. Extra-Welfarismus als Ansatz für die allokativen Effizienz in Deutschland von Bedeutung ist.

4.3.3 Maß des Gesamtnutzens (A) Das QALY als Maß des Gesamtnutzens

Seite 23/34

Das IQWiG scheint den QALY-Ansatz für die Kosten-Nutzen-Bewertung zu „rehabilitieren“: *„Das Institut schließt nicht aus, in Kosten-Nutzen-Bewertungen auf QALYs als Maß für den Gesamtnutzen zurückzugreifen. QALYs sollten nur verwendet werden, wenn die einfließenden Werte zu den Gesundheitszuständen zunächst bei den Betroffenen erhoben worden sind. Wenn generische Indexinstrumente eingesetzt wurden, muss ein in Deutschland validierter Tarif bei der Ermittlung des Nutzwertes angewendet werden. Die Verwendung von QALYs sowie ihre Erhebung und Umrechnung in einen deutschen Tarif müssen in jedem Fall nachvollziehbar dargestellt und begründet werden.“*. Allerdings verweist es explizit auf § 35b SGB V und verkennt wie bereits unter Abschnitt B) Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen explizit dargelegt, Nutzwerte durchaus auch im Rahmen der frühen Nutzenbewertung nach § 35a SGB V und der entsprechenden Verordnung (AM-NutzenV) ihren Platz haben und als Nutzenparameter sogar belastbarer als klinische Outcomes interpretiert werden können, da sie auf der Erwartungsnutzentheorie fußen und letztlich aus einer aggregierten und auf einen vorgegebenen Wertebereich normierten Nutzenfunktion hergeleitet werden. Die theoretische Fundierung von Nutzwerten ist somit über die Kosten-Nutzen-Bewertung auch für die (Zusatz-)Nutzenbewertung wichtig, weshalb der vfa die Verortung der Nutzwerte als alternative Möglichkeit zur Bestimmung eines Maßes für den Gesamtnutzen in den Abschnitt 3.1.5 wir in den Vorversionen der Allgemeinen Methoden favorisiert. Abschließend sei darauf hingewiesen, dass ein einfacher Verweis hinsichtlich der wissenschaftlichen Debatten zu ethischen und methodischen Fragestellungen des QALY auf entsprechende Literatur das IQWiG nicht davon befreit, eine Stellung hierzu zu beziehen, da sonst je nach Fallkonstellation genau Argumente aus dieser Debatte vorgetragen werden können, um selektiv Nutzwerte sowohl im Rahmen einer Kosten-Nutzen-Bewertung als **auch als Mittel zur Abwägung von „Nutzen“ und „Schaden“ zu akzeptieren** oder zu negieren. Als Industrieverband ist der vfa verpflichtet für ein rechtssicheres und planbares Vorgehen zu plädieren, um stabile Rahmenbedingungen für seine Mitglieder zu gewährleisten und nicht über das Einfallstor einer latenten Methodendiskussion, Argumente für das Für und Wider von Nutzwerten situativ und interessengesteuert instituts- oder systemseitig einzusetzen. Deswegen wäre eine genaue Position des IQWiG zum Umgang von Nutzwerten in der KNB und in der frühen Nutzenbewertung gewünscht. Hierzu gehören beispielsweise auch handlungsanleitende Aussagen zur genaueren Auswahl der zur Verfügung stehenden Verfahren (direkte vs. indirekte Erhebung), zu den Zielpopulationen ihres Einsatzes (Betroffene – nicht Betroffene), zur Möglichkeit des Einsatzes von Mapping krankheitsspezifischer Instrumente (beispielsweise Chronic Liver Disease Questionnaire (CLDQ) oder Cushing-QoL Scores) auf generische Indexinstrumente (beispielsweise SF6D) zur Herleitung von Nutzwerten usw. Abschließend möchte der vfa hervorheben, dass hier

auch ein direkter Bezug zu den internationalen Standards der Gesundheitsökonomie geschaffen werden muss, da die meisten forschenden Arzneimittelhersteller global aufgestellt sind und auch entsprechende wissenschaftliche Erkenntnisse und Methodenentwicklungen im Bereich der evidenzbasierten Medizin und der Gesundheitsökonomie international erlangen. Die Industrie und Forschung wird somit nur bedingt nationale Eigenheiten berücksichtigen können, weswegen auch in den Methodenpapieren verschiedener HTA-Institutionen ein gewisser Grad an Konsens gefunden werden muss und ein pragmatisches Vorgehen favorisiert werden sollte.

Seite 24/34

4.3.3 Maß des Gesamtnutzens B) Erhebung von Präferenzen zur Erstellung eines Maßes des Gesamtnutzens

Einführend soll festgehalten werden, dass Patientenpräferenzen international im Rahmen der Zulassung (beispielsweise EMA & FDA) und der Bewertung von Gesundheitstechnologien sowie der anschließenden Erstattungsentscheidungsfindung neben den Methoden der evidenzbasierten Medizin sukzessiv eine immer bedeutendere Rolle spielen. Dieser Entwicklung folgt offensichtlich das neokorporatistisch geprägte deutsche Gesundheitssystem allerdings noch nicht. Es verharrt eher auf den tradierten Pfaden des Interessenausgleichs und inkorporiert rein formal Patienteninteressen durch die gesetzgeberisch bedingte Partizipation der institutionalisierten Patientenvertreter nach §§ 140f, 140g SGB V in den Beratungsprozessen der Selbstverwaltung, ohne diese mit Stimmrechten auszustatten. Wissenschaftlich fundierte Präferenzhebungen in Bewertungs- und Entscheidungsprozessen werden systeminhärent anscheinend noch immer ignoriert, wie sich am Beispiel der Nutzenbewertung von Arzneimitteln nach dem AMNOG eindeutig zeigen lässt. Ferner bedient sich das angewandte Methodeninstrumentarium des IQWiG hierbei statistischen Eigenkonstrukten, die explizit patientenseitige Präferenzen exkludieren. Die sich hierdurch abzeichnende Paradoxität hinsichtlich der Zielsetzung einer breit propagierten Patientenorientierung und der fehlenden Einbeziehung von wissenschaftlich fundierten Patientenpräferenzen bei (Erstattungs-)Entscheidungen der Selbstverwaltung wird nunmehr offensichtlich. In der Entwurfsversion 4.2 der Allgemeinen Methoden des IQWiG verbannt das Institut Methoden zur Präferenzhebung von Patienten aus dem Abschnitt 3.1.5 Zusammenfassende Bewertung, der auf die Nutzenbewertung medizinischer Interventionen abstellt, nun explizit in die KNB und negiert damit die multikriterielle Entscheidungsfindung als mögliche Alternative für die Herleitung eines Gesamtnutzens im Rahmen der Nutzenbewertung von medizinischen Interventionen. Das IQWiG hält fest, dass indikationsbezogen auf AHP und CA zur Generierung eines Maßes des Gesamtnutzens zurückgegriffen werden kann, allerdings gäbe es noch ungelöste methodische Probleme beim Einsatz dieser Verfahren, sodass gegenwärtig eine routinemäßige Anwendung dieser Methoden nicht vorgesehen sei. Es bezieht sich dabei weder auf die theoretische Verortung

bzw. Fundierung der genannten Verfahren noch auf die Häufigkeit ihrer Anwendung im Rahmen von gesundheitsrelevanten Fragestellungen. Nach Ansicht des vfa wird der Leser nicht auf die vom IQWiG angemahnten methodischen Probleme dieser Methoden tiefergehend hingewiesen, um deren Stellenwert einschätzen zu können. Es werden auch anders als zu den Methoden selbst keine expliziten Literaturverweise zu deren methodischen Problemen gebracht. In der nachfolgenden Kurzbeschreibung und Gegenüberstellung der beiden vom IQWiG genannten Methoden AHP und CA werden eher Hinweise zur Machbarkeit (Feasibility) bzw. zur Zielpopulation oder Fragestellung (geschlossene Gruppe versus Kompensationsleistung für entgangenen Nutzen) gegeben, die keiner klaren Rationales folgen. Eine Methodengegenüberstellung und -diskussion anhand klassischer Kriterien der Sozialempirie, wie Repräsentativität, Objektivität, Reliabilität und Validität, erfolgt nicht und die Institutsaussage, eine klare Zuordnung, wann welches Verfahren bevorzugt werden soll, ließe sich somit kaum ableiten, ist nach Auffassung des vfa nicht klar herausgearbeitet, weil genau die Auseinandersetzung mit beiden Methoden anhand der weiter oben genannten Kriterien in der Entwurfsversion 4.2 fehlt bzw. auch keine aussagekräftigen Literaturverweise gegeben werden. Die vom IQWiG geforderte Nachvollziehbarkeit bzw. Transparenz in Planung, Durchführung, Auswertung und Bewertung einer jeden Umsetzung mag zwar zentral sein, sie ersetzt allerdings nicht die weiter oben genannten Gütekriterien (insbesondere Objektivität, Reliabilität und Validität) und den fehlenden Methodendiskurs.

Es wird abschließend begrüßt, dass die gängigen Verfahren zur Aggregation eines Gesamtnutzens im Methodenentwurf 4.2 vertieft werden. Qualitätsadjustierte Lebensjahre und die multikriterielle Entscheidungsanalyse werden allerdings ohne konzeptionellen gesundheitsökonomischen Hintergrund als Alternativen vorgestellt. Es werden keine Aussagen darüber getroffen, wann Vor- und Nachteile der Alternativen zu einer spezifischen Auswahl im Rahmen der untersuchten Fragestellung führen können. Das Methodenpapier reflektiert nicht die gesundheitsökonomische Diskussion in beiden Theoriefeldern. Es diskutiert auch nicht, wie diese Verfahren in Relation zu den international anerkannten Standards der Gesundheitsökonomie eingesetzt werden sollen.

4.4.1 Perspektive und zu berücksichtigende Kosten

Auf Seite 91 in der zusammenfassenden Tabelle 6 zur Perspektive und den relevanten zu berücksichtigenden Kosten werden nach Auffassung des vfa unbegründet die erstattungsfähigen direkten nichtmedizinischen Kosten aus der Perspektive der Sozialversicherung ausgeschlossen („nein“ in der Tabelle). **Allerdings** gibt es solche Kosten analog zur reinen GKV-Perspektive (beispielsweise Fahrtkosten bei Rehabilitationsmaßnahmen in stationären Rehabilitationseinrichtungen, die von Rentenversicherungsträgern oder anderen Sozialversicherungsträgern übernommen wer-

den), sodass deren Ausschluss aus dieser Perspektive nicht nachvollzogen werden kann bzw. fälschlicherweise erfolgt ist.

Seite 26/34

4.4.2 Abgrenzung der Kosten B) Indirekte Kosten

Das IQWiG führt aus, dass es Produktivitätsausfälle auf der Kostenseite berücksichtigt, wie es auch überwiegend die Literatur empfiehlt. Der vfa schlägt vor, hier den Terminus „primär“ auf der Kostenseite einzuführen, da das IQWiG in alten Berichtsplänen durchaus schon Arbeitsunfähigkeitstage auf der Nutzenseite berücksichtigt hat (Arbeits- und Erwerbsfähigkeit beispielsweise unter den Zielgrößen im Berichtsplan zur Bewertung von SNRI).

4.4.3 Schritte zur Kostenbestimmung C) Bewertung der Ressourcen: Besonderheiten in weiteren Perspektiven

Das IQWiG merkt auf Seite 96 an, dass in gesundheitsökonomischen Evaluationen zumeist nur die zusätzliche Betrachtung der indirekten Kosten international üblich ist. Ein Quellenhinweis für diese richtige Aussage sollte ergänzt werden.

4.4.3 Schritte zur Kostenbestimmung C) Bewertung der Ressourcen: Bewertung von indirekten Kosten

In der Entwurfsversion 4.2 der Allgemeinen Methoden lautet es auf Seite 96: *„Für Produktivitätsverluste berücksichtigt das Institut im Basisfall den Friktionskostenansatz, da der Humankapitalansatz auf einigen unrealistischen Annahmen (insbesondere Vollbeschäftigung am Arbeitsmarkt) beruht. In Sensitivitätsanalysen kann diese Schätzung dem Humankapitalansatz gegenübergestellt werden“* und weiter auf Seite 97 *„Die Friktionskosten werden mit 80% der Lohnkosten angenommen (analog zu den Niederlanden). Die Friktionsperiode wird, sofern keine aktuellen Daten verfügbar sind, in Anlehnung an die durchschnittliche tatsächliche Besetzungsdauer in Deutschland für das Jahr 2012, mit 82 Tagen angesetzt“*. Hierzu merkt der vfa an, dass zwar theoretisch der Friktionskostenansatz für die Bewertung von indirekten Kosten dem Humankapitalansatz überlegen ist, seine Umsetzung aber aufgrund fehlender Daten zur genauen branchen- und arbeitsstellenspezifischen Friktionsperiode kaum adäquat gewährleistet werden kann. Wir plädieren dafür im Basisfall der Humankapitalansatz zu berücksichtigen und nur im Rahmen von Sensitivitätsanalysen den Friktionskostenansatz anzuwenden. Beispielhaft für die Schwierigkeiten des Friktionskostenansatzes sind die vom IQWiG angenommenen 80% der Lohnkosten als Friktionskosten, deren empirische Basis aus den Niederlanden stammt, wobei die Niederlande einen völlig unterschiedlichen Arbeitsmarkt im Vergleich zu Deutschland aufweisen. Auch die Verwendung der durchschnittlichen tatsächlichen Besetzungsdauer in Deutschland (82 Tage in 2012), die wiederum branchenspezifisch

heterogen ausfällt, zeugt von den zumindest aktuell geltenden Schwierigkeiten der Umsetzung des Friktionskostenansatzes für Deutschland.

Seite 27/34

4.4.4 Datengrundlage

Laut IQWiG sind auf Seite 97 für das Kostengeschehen Sekundärdaten in Form von ausgewerteten GKV-Routinedaten anhand einer repräsentativen **Stichprobe die Datenquelle erster Wahl**. *„Wenn aktuelle Analysen nicht aus der Literatur entnommen werden können, sind bevorzugt eigene Auswertungen durchzuführen.“* Hierbei übersieht das IQWiG die Zugangsmöglichkeiten der pharmazeutischen Industrie zu solchen Daten. Während das IQWiG nach § 303e SGB V (Datenverarbeitung und -nutzung, Verordnungsermächtigung) Nutzungs- und Verarbeitungsrechte der bei der Datenaufbereitungsstelle gespeicherten Daten neben anderen Institutionen (meist öffentlich rechtlichen Einrichtungen) hat, ist die pharmazeutische Industrie hiervon ausgeschlossen. Selbst projektbezogen kooperierende Institutionen mit der pharmazeutischen Industrie, müssen sich den Zugang und die Zurverfügungstellung solcher Daten explizit genehmigen lassen. Ein Zugang der pharmazeutischen Industrie zu GKV-Routinedaten ist somit nicht gewährleistet, weshalb er entweder gesetzgeberisch eingebracht werden müsste oder alternative Daten hierzu akzeptiert werden müssten.

In diesem Abschnitt merkt das IQWiG an, dass ergänzend Leitlinien oder Ergebnisse aus Expertenbefragungen hinzugezogen werden können, wenn Routinedaten nicht hinreichend die Versorgung in allen Zuständen des Modells darstellen. Bevorzugt sollten laut IQWiG evidenzbasierte Leitlinien aus dem deutschen Gesundheitssystem mit Verweis auf Abschnitt 4.1.7 verwendet werden. Siehe dazu in der vorliegenden Stellungnahme enthaltenen Kommentaren zu Abschnitt 4.1.7. Das IQWiG äußert sich weiter auf **Seite 98: „Eine Übertragbarkeit von Versorgungspfaden und Kostendaten aus anderen Gesundheitssystemen ist aufgrund von Systemunterschieden selten gegeben und nur unter sehr strengen Voraussetzungen möglich. Die Übertragbarkeit von Kostendaten aus folgenden Ländern wird nicht grundsätzlich ausgeschlossen, da sie ein dem deutschen ähnliches System von ambulanter und stationärer Versorgung haben: Österreich, Schweiz, Niederlande, Belgien, Frankreich. Eine Verwendung muss aber jeweils begründet und diskutiert werden. Kostendaten aus weiteren Ländern dürfen in einer KNB nicht verwendet werden“**. Indirekt bestätigt das IQWiG mit dieser Äußerung, dass die Übertragung der Friktionskosten analog zu den Niederlanden nicht ohne Weiteres aufgrund von Systemunterschieden (Wirtschaftssystem, Gesundheitssystem, Wirtschaftszyklus) möglich ist. **Ferner wird bei der Nennung der Länder mit „ähnlicher“ stationärer und ambulanter Versorgung verkannt, dass die Übertragbarkeit von Kosten wegen der unterschiedlicher Honorierung, Erstattung und Finanzierung der Systeme unabhängig von der „ähnlichen“ Leistungserbringung nicht unbedingt gegeben ist.** Die Übernahme von Kostendaten aus den oben

genannten Gesundheitssystemen kann nach Auffassung des vfa nur als Ultima Ratio Vorgehen gesehen werden und ist für solche Ausnahmefälle auch zu begrüßen.

Seite 28/34

4.5.1 Daten

Unter diesen Abschnitt führt das IQWiG aus, dass eine besondere Rolle für die Modellierung die Basiswahrscheinlichkeiten für Ereignisse spielen. In einem Modell seien für jeden Endpunkt Angaben zu den endpunktbezogenen Ereignishäufigkeiten bzw. -wahrscheinlichkeiten nötig, die als Ausgangswerte in das entscheidungsanalytische Modell eingehen. Allerdings fehlen hier nach Auffassung des vfa jegliche Aussagen, wie diese Basiswahrscheinlichkeiten ermittelt werden sollen, womit wiederum die Gefahr besteht, unterschiedliche Vorgehensweisen in den KNB zu ermöglichen und somit Inkonsistenzen von KNB zu KNB einzubringen. Der vfa schlägt hier vor, ein einheitliches Vorgehen zur Ermittlung der in das Modell einfließenden Basiswahrscheinlichkeiten im Methodenpapier aufzugreifen.

4.6.2 Verfahrensverlauf

Das IQWiG merkt bei der im Rahmen der Berücksichtigung der erweiterten Dominanz resultierenden stufenförmigen absoluten Effizienzgrenze auf Seite 102 an, dass zu beachten sei, dass die absolute Effizienzgrenze keine Steigung mehr im Sinne eines Kehrwerts der Zahlungsbereitschaft wiedergibt und somit keine Schwellenwerte zu erheben wären. Nach Auffassung des vfa kann bei einer nicht möglichen Teilbarkeit der erweitert dominierten Intervention bzw. bei einer Budgetrestriktion, die aus Gründen des Budget Impacts die Anwendung der zwar effektiveren, aber gleichzeitig zu teuren Intervention restringiert, durchaus die erweitert dominierte Intervention relevant sein, eben weil sie verordnet bzw. angewendet wird und somit auch Zahlungsbereitschaften aus dem Verlauf der absoluten Effizienzgrenze und ihrer Verbindungssegmente im Bereich erweiterter Dominanz ablesbar werden.

4.7.3 Darstellung von Unsicherheit mittels des Net Health Benefit

Laut IQWiG ist der Net Health Benefit (NHB) ein etabliertes Verfahren zur Darstellung von Ergebnissen aus probabilistischen Sensitivitätsanalysen. Durch die NHB-Berechnung wird dem Problem der Parameterunsicherheit Rechnung getragen, da der NHB eine Funktion sowohl von Zusatznutzen und -kosten als auch der Effizienzgrenze ist und die Position der zu bewertenden Intervention als Abstand zu der sich verlagernden Effizienzgrenze bzw. zu dem sich verlagernden letzten Segment der Effizienzgrenze abbildet. Daher sollten nach dem IQWiG sowohl die Basisfallanalysen als auch die deterministischen und die probabilistischen Sensitivitätsanalysen basierend auf dem Konzept der NHB-Berechnung durchgeführt werden. Hierzu merkt der vfa an, dass der neu in die Methoden aufgenom-

mene und bereits im Projekt G09-01 umgesetzte NHB Ansatz anwendbar sein kann. Es ist zu berücksichtigen, dass der NHB Ansatz wegen der spezifischen rechtsschiefen Kostenverteilungen, die charakteristisch für Kosten-Nutzen-Bewertungen aufgrund der Anwendung von Interquartilsabständen als Ansatz für die probabilistischen Sensitivitätsanalysen sind, eher glättend wirken. Um Unsicherheiten darzustellen sind nach Auffassung des vfa folgende Darstellungen sinnvoll: Scatter Plots, Vertrauensintervalle, Konturplots, Kosten-Effektivitäts-Akzeptanzkurven und Kovarianzanalysen (sofern Kosten und Effekte aus derselben Quelle stammen).

Seite 29/34

4.8 Ausgaben-Einfluss-Analyse (AEA bzw. Budget-Impact-Analyse)

Der folgende Abschnitt zur AEA in der Entwurfsversion der Allgemeinen Methoden **„Eine AEA prognostiziert insbesondere, wie eine Veränderung im Mix der Interventionen für eine bestimmte Krankheit die Ausgaben für ein Indikationsgebiet zukünftig beeinflusst“** sollte wie folgt umgewandelt werden in **„[...] zukünftig beeinflussen kann“**. Dieser Vorschlag erfolgt aufgrund der Tatsache, dass der Interventionsmix abhängig von der Substitution der alten Technologien durch neue Gesundheitstechnologien abhängig ist. Die Substitutionsrate kann nicht sicher vorhergesagt werden, da sich sogar in einem relativ kurzen Zeitverlauf beispielsweise die Evidenz zu Kontraindikationen, der Eintritt neuer Technologien in den Markt sowie eine Veränderung der epidemiologischen Kenngrößen für das Krankheitsbild ergeben können.

Insgesamt sollte der Abschnitt zur Sensitivitätsanalyse mit einer Grafik zu ihren Kostenkomponenten und der dahinterstehenden Epidemiologie versehen werden.

4.9.1 Gesetzliche Anforderungen und Verfahrensablauf

Das IQWiG bezieht sich in diesem Abschnitt u. a. auch auf die Festlegungen des G-BA nach § 35b SGB V. Hierunter werden zwei Punkte genannt (i. zweckmäßige Vergleichstherapie und andere Arzneimittel und Behandlungsformen, mit denen das zu bewertende Arzneimittel verglichen werden soll; ii. Maß des Gesamtnutzens), die einen sehr großen Einfluss auf die Ergebnisse der Effizienzgrenze haben. Die Effizienzgrenze erfordert, wie ursprünglich in den Methoden zur KNB des IQWiG angegeben, eigentlich eine erschöpfende Berücksichtigung aller Interventionen in einem Indikationsgebiet oder reduktionistischer zumindest alle relevanten im Sinne des Versorgungsgeschehens. Nur so kann der Verlauf der Kurven nicht manipulationsabhängig gestaltet werden. Des Weiteren ist das Maß des Gesamtnutzens als aggregierter Endpunkt über positive und negative Effekte der Intervention von Interesse im Vergleich zu ihren Komparatoren ausschlaggebend für die Ergebniserzielung. Werden beide Punkte durch Festsetzungen des G-BA umgesetzt, die nicht den methodischen

Erfordernissen des Effizienzgrenzansatzes entsprechen, sind die Ergebnisse der Effizienzgrenzanalyse verzerrungsanfällig und weder eindeutig interpretierbar noch verwertbar, da sie Grundannahmen dieser Methodik verletzen können.

Seite 30/34

4.9.4 Interquartilsregion als Maß der Streuung für Preisverhandlungen

Hier wird erneut auf die Kommentare unter Abschnitt 4.7.3 verwiesen. In seinen Ausführungen merkt das IQWiG an, dass neben der Interquartilsregion zwischen oberem und unterem Quartil auch grundsätzlich die Angabe von anderen Regionen mit anderen Maßen sinnvoll sein kann. Hierzu gilt es festzuhalten, dass in Abhängigkeit der Parameterverteilung dies wohl in den meisten Fällen auch so sein wird und, wie weiter oben bereits angemerkt, eine Interquartilsregion mit 50% der Simulationen in den probabilistischen Sensitivitätsanalysen Gefahr läuft, glättend zu wirken und eine fiktive Sicherheit um den nutzenbereinigten Erstattungspreis erzeugen. Eine Berücksichtigung der gesamten Parametersimulationen in der probabilistischen Sensitivitätsanalyse würde ein Bild befördern, dass den realen Parameterverteilungen eher entspricht und einen breiteren Korridor für nutzenbereinigte Erstattungspreise generiert.

vfa-Position zusammenfassend: Hinsichtlich des Methodendiskurses um die Effizienzgrenze verweist der vfa auf bereits zu den Methoden der Kosten-Nutzen-Bewertung (Version 1.0) und zum Vorbericht G09-01 Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin eingegangene Kommentare.

Die Methoden zur Nutzenbewertung nach § 35a SGB V basieren auf den anerkannten internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie und somit kann auch keine Trennung der Standards nach unterschiedlichen Disziplinen zwischen § 35a SGB V, § 35b SGB V und § 139a SGB V erfolgen.

Auf Basis der evaluativen Gesundheitsökonomie erfolgt international primär die vergleichende Kosten-Nutzen-Bewertungen der untersuchten medizinischen Interventionen innerhalb der betrachteten Indikation. Eine KNB im Rahmen von § 35b SGB V kann prozedural betrachtet mehrere Indikationsgebiete einschließen, wenn zulassungsbedingt mehrere Indikationen abgedeckt werden. Der vfa schlägt vor, indikationsübergreifende Wirtschaftlichkeitskriterien nach den internationalen Methoden der Gesundheitsökonomie zu berücksichtigen.

Der Zeithorizont sollte gegebenenfalls auch kürzer als die durchschnittliche Studiendauer gewählt werden können, wenn beispielsweise eine Intermittenzbewertung erwünscht ist, aus deren Ergebnissen die Weiterführung einer Versorgungsstudie abhängig gemacht wird.

Der Terminus Substanzklasse, der auch Nichtarzneimittel beinhaltet (z. B. Herbizide) sollte in den terminologisch gängigeren Begriff Wirkstoffklasse für Arzneimittelbewertungen umgeändert werden. Fehlende Heterogenität ist nicht mit dem Vorhandensein von Homogenität bei der Zusammenfassung in Wirkstoffklassen gleichzusetzen und es sollte in den Methoden spezifisch darauf abgestellt werden.

Die Nutzenseite sollte in Abhängigkeit von dem jeweiligen Erkenntnisobjekt der KNB entweder unterschiedlich operationalisiert werden oder die KNB um Ausführungen zur Nutzenseite bei nichtmedikamentösen Verfahren und für die Bewertung von Arzneimitteln nach § 139a SGB V (also nicht ausschließlich im Rahmen der frühen Nutzenbewertung nach § 35a SGB V) ergänzt werden.

Eine ausschließliche Orientierung an nationale Leitlinien birgt die Gefahr überholte Empfehlungen als Datengrundlage für den gesetzten Versorgungskontext geltend zu machen und aktuellere Entwicklungen, die zwar noch nicht in Leitlinien umgesetzt wurden, sich aber dennoch in der Versorgungsrealität wieder finden, zu ignorieren. Aktualitätsgrad der Leitlinien und Versorgungsrealität sollten bei der Bestimmung der zu verwendenden Datengrundlage mitberücksichtigt werden.

Die Definition der Effizienzgrenze als zusammengesetzte Verbindungssegmente der effizienten Interventionen in einer grafischen Darstellung des Nutzens und der Kosten sollte anhand eines indikationsspezifischen und endpunktbestimmten Koordinatensystems eingeführt werden und erst im Nachgang auf ihre Eigenschaften eingegangen werden.

Die Effizienzgrenze kann potenzielle Ineffizienzen für bereits im System befindliche Interventionen aufzeigen. Eine Handlungsempfehlung als Operationalisierung der Angemessenheit durch Extrapolation des letzten Segmentes für Interventionen, die im Nord-Ost-Quadranten der Kosten-Effektivitäts-Ebene liegen, ist allerdings nicht zielführend und umsetzbar. Eine Gewichtung der patientenrelevanten Endpunkte, ohne jegliche Referenzpunkte bzw. Rationale als Werturteil, stellt auch für Entscheidungsträger eine schier nicht zu erfüllende Aufgabe. Ähnliches gilt auch bei Zulassung in mehreren Indikationsgebieten der zu bewertenden Intervention und deren Gewichtung.

KNB-Modelle können der Öffentlichkeit nicht vollumfänglich und voll interaktiv zur Benutzung zur Verfügung gestellt werden, da sie aus Gründen des intellektuellen Eigentums ihrer Entwickler geschützt werden können und somit nur als vertrauliche Informationen dem IQWiG vorgelegt werden sollten.

Der Ansatz der Effizienzgrenze ist erschöpfend und ganzheitlich angelegt, was die Einbeziehung der patientenrelevanten Endpunkte angeht.

Einer Verhandlungsoption nach § 130b SGB V für Vereinbarungen zwischen dem Spitzenverband Bund der Krankenkassen (GKV-SV) und pharmazeutischen Unternehmen über Erstattungsbeträge für Arzneimittel steht ein algorithmisches Gewichtungsvorgehen für Endpunkte entgegen. Für andere Preisfindungen, welche keinen Verhandlungsweg sozialrechtlich vorsehen, wäre der vom IQWiG vorgeschlagene Modus nicht zielführend, da hiermit der Selbstverwaltung die Möglichkeit nach freiem Ermessen Preise bzw. Erstattungsbeträge zu gestalten, Tür und Tor geöffnet würde. Eine schon festgelegte Gewichtung von Endpunkten nach den Auftragsvorgaben ist nicht nachvollziehbar, da diese Vorgaben wiederum auf einer Rationale fußen müssen, die aber zumindest nach dieser o.g. Formulierung nicht immer für eine Gewichtung erkennbar ist.

Die theoretische Fundierung von Nutzwerten ist über die KNB auch für die (Zusatz-)Nutzenbewertung wichtig, weshalb die Verortung der Nutzwerte als alternative Möglichkeit zur Bestimmung eines Maßes für den Gesamtnutzen in den entsprechenden Abschnitt zur Aggregation des Nutzens Eingang finden sollte.

Handlungsanleitende Aussagen zur genaueren Auswahl der zur Verfügung stehenden Verfahren für Nutzwerte (direkte vs. indirekte Erhebung), zu den Zielpopulationen ihres Einsatzes (Betroffene – nicht Betroffene), zur Möglichkeit des Einsatzes von Mapping krankheitsspezifischer Instrumente (beispielsweise Chronic Liver Disease Questionnaire (CLDQ) oder Cushing-QoL Scores) auf generische Indexinstrumente (beispielsweise SF6D) zur Herleitung von Nutzwerten usw. fehlen. Ein direkter Bezug zu den internationalen Standards der Gesundheitsökonomie muss geschaffen werden, da die meisten forschenden Arzneimittelhersteller global aufgestellt sind und auch entsprechende wissenschaftliche Erkenntnisse und Methodenentwicklungen im Bereich der evidenzbasierten Medizin und der Gesundheitsökonomie international erfolgen. Die Industrie und Forschung wird somit nur bedingt nationale Eigenheiten berücksichtigen können.

Bei der Erhebung von Patientenpräferenzen erfolgt die Methodengegenüberstellung und -diskussion nicht anhand klassischer Kriterien der Sozialempirie wie Repräsentativität, Objektivität, Reliabilität und Validität und die Aussage, eine klare Zuordnung, wann welches Verfahren bevorzugt werden soll, ließe sich somit kaum ableiten, ist nicht klar herausgearbeitet.

Erstattungsfähige direkte nichtmedizinische Kosten sollten auch aus der Perspektive der Sozialversicherung eingeschlossen werden.

Im Basisfall sollte für die Bestimmung der indirekten Kosten der Humankapitalansatz berücksichtigt werden und nur im Rahmen von Sensitivitätsanalysen der Friktionskostenansatz angewendet werden.

Ein Zugang der pharmazeutischen Industrie zu GKV-Routinedaten ist nicht gewährleistet, weshalb er entweder gesetzgeberisch eingebracht werden müsste oder alternative Daten hierzu im Rahmen der KNB akzeptiert werden müssten.

Ein einheitliches Vorgehen zur Ermittlung der in das Modell einfließenden Basiswahrscheinlichkeiten sollte im Methodenpapier aufgegriffen werden.

Bei einer nicht möglichen Teilbarkeit der erweitert dominierten Intervention bzw. bei einer Budgetrestriktion, die aus Gründen des Budget Impacts die Anwendung der zwar effektiveren, aber gleichzeitig zu teuren Intervention restringiert, kann durchaus die erweitert dominierte Intervention relevant sein und somit auch Zahlungsbereitschaften aus dem Verlauf der absoluten Effizienzgrenze und ihrer Verbindungssegmente im Bereich erweiterter Dominanz ablesbar werden.

Der Net-Health-Benefit (NHB) Ansatz kann wegen der spezifischen rechtschiefen Kostenverteilungen, die charakteristisch für Kosten-Nutzen-Bewertungen aufgrund der Anwendung von Interquartilsabständen als Ansatz für die probabilistischen Sensitivitätsanalysen sind, glättend wirken. Eine Berücksichtigung der gesamten Parametersimulationen in der probabilistischen Sensitivitätsanalyse würde ein Bild befördern, das den realen Parameterverteilungen eher entspricht und einen breiteren Korridor für nutzenbereinigte Erstattungspreise generiert. Um Unsicherheiten darzustellen, sind Scatter Plots, Vertrauensintervalle, Konturplots, Kosten-Effektivitäts-Akzeptanzkurven und Kovarianzanalysen (sofern Kosten und Effekte aus derselben Quelle stammen) sinnvoll.

Die Effizienzgrenze erfordert eine erschöpfende Berücksichtigung aller Interventionen in einem Indikationsgebiet oder reduktionistischer zumindest alle relevanten im Sinne des Versorgungsgeschehens. Nur so kann der Verlauf der Kurven nicht manipulationsabhängig gestaltet werden.

L. Informationsbeschaffung (Abschnitt 7)

Das IQWiG nimmt zu unaufgefordert übermittelten Daten wie folgt Stellung: *„Ebenso werden unaufgefordert übermittelte Daten, d. h. außerhalb von Stellungnahmeverfahren oder außerhalb sonstiger bestehender Regelungen (z. B. Herstelleranfragen), inhaltlich nicht berücksichtigt. Die unaufgeforderte Übermittlung von Studiendaten birgt das Risiko, dass sie selektiv geschieht und es dadurch zu einer Verzerrung des Ergebnisses der Nutzenbewertung kommt“*. Nach Auffassung des vfa sollte dieses Vorgehen mit unaufgefordert übermittelten Daten überdacht werden. Im Rahmen der evidenzbasierten Medizin sollte jegliche relevante Information berücksichtigt werden, unabhängig davon, ob sie aufgefordert oder unaufgefordert eingegangen ist. Ferner können unaufgefordert einge-

reichte Daten auch zu einer Aktualisierung oder Fehlerkorrektur der Institutsprodukte beitragen. Vielmehr sollten Kriterien für die Verwertbarkeit solcher Daten ihre Relevanz und ihr Verzerrungspotenzial sein, deren Überprüfung eigentlich eine originäre Aufgabe des Instituts ist.

Seite 34/34

vfa-Position zusammenfassend: Jegliche relevante und somit verwertbare Information sollte vom IQWiG berücksichtigt werden, unabhängig davon, ob sie aufgefordert oder unaufgefordert eingegangen ist.

A.1.28 – YES Pharmaceutical Development Services GmbH

Von: [Johannes Hofmann YES-GmbH](#)
An: [Methoden](#)
Thema: Stellungnahme Allgemeine Methoden: Entwurf für Version 4.2
Datum: Donnerstag, 24. Juli 2014 15:50:18
Anlagen: Formblatt YES Ph Dev Serv GmbH.pdf

Sehr geehrte Damen und Herren,

In Tabelle 16 / Seite 214 des Methoden-Entwurfs Version 4.2 wird in der Fußnote a) darauf verwiesen, dass die Risiken für die Lebensqualität als "Non-Response" zu definieren sind.

Da die Definition der Risiken als Response oder Non-Response potentiell einen Einfluss auf die Effektgröße, und damit den Zusatznutzen nehmen kann, wird zu diesem Punkt um Klärung gebeten:

In der Regel werden in der frühen Nutzenbewertung die Risiken zur Ableitung der relativen Effektmaße in dieser Zielgrößenkategorie (schwerwiegende Symptome und Nebenwirkungen) als "Response" (klinische Verbesserung) und nicht als "Non-Response" definiert. Es erscheint daher inkonsistent, die Risiken für die Responderanalysen in der Lebensqualität (die sich ja in derselben Zielgrößenkategorie befindet wie die schwerwiegenden Symptome und Nebenwirkungen) als "Non-Response" zu definieren.

Mit freundlichen Grüßen,

Johannes Hofmann

YES Pharmaceutical Development Services GmbH
Dr. Johannes Hofmann
Team Leader Medical Affairs & Toxicology
Bahnstr. 42-46
61381 Friedrichsdorf
GERMANY


web: www.yes-services.eu

Sitz: Friedrichsdorf, Handelsregister-Nr.: HRB 5042 Amtsgericht Bad Homburg,
Geschäftsführer: Dr. Thomas Dobmeyer, Dr. Markus Zimmer

CONFIDENTIAL

This email transmission and any documents attached to it may contain information that is confidential or legally privileged. If you are not the intended recipient or a person responsible for delivering this transmission to the intended recipient, you are hereby notified that any disclosure, copying, distribution or use of this transmission is not permitted. If you have received this transmission in error, please notify the sender immediately and destroy the original transmission and its attachments and destroy any hardcopies.

A.2 – Stellungnahmen von Privatpersonen

A.2.1 – Kohlmann, Thomas

Stellungnahme T. Kohlmann

Im Entwurf für die Version 4.2 zum Methodenpapier „Allgemeine Methoden“ werden in den Abschnitten 3.1.1 und 3.2.4 die in der aktuellen Version enthaltenen Zielgrößen „Zufriedenheit der Patientinnen und Patienten“, „Patientenzufriedenheit“ und „Behandlungszufriedenheit“ gestrichen.

Es handelt sich bei den genannten Aspekten der Patientenzufriedenheit jedoch um Zielgrößen, die wichtige Dimensionen des patientenrelevanten Nutzens abbilden können, die in den vorrangig zu berücksichtigenden Zielgrößen (Mortalität, Morbidität, gesundheitsbezogene Lebensqualität) nicht hinreichend erfasst werden. So können beispielsweise bestimmte Merkmale von Arzneimitteln (vereinfachte oder flexiblere Einnahme oder Anwendung, verbesserte Darreichungsform, längere Intervalle zwischen der Anwendung, seltenere Arztbesuche oder Laborkontrollen) zu einer wesentlichen Erhöhung des patientenrelevanten Nutzens führen. Aus diesen Gründen wird die Patientenzufriedenheit in den meisten Konzepten zur Definition und Strukturierung gesundheitsbezogener Outcomes als eigenständiges Bewertungskriterium geführt (z.B. Asadi-Lari M, Tamburini M, Gray D. Patients' needs, satisfaction, and health related quality of life: towards a comprehensive model. *Health and Quality of Life Outcomes* 2004; 2:32).

Anhand der genannten Beispiele wird deutlich, dass Aspekte der Patientenzufriedenheit einen unmittelbaren patientenrelevanten Nutzen ansprechen können. In der Überarbeitung des Methodenpapiers sollten deshalb diese Zielgrößen nicht gestrichen werden.

A.2.2 – Mühlbacher, Axel

Stellungnahme zum Entwurf des Methodenpapiers 4.2
(Version vom 18.06.2014)
des Instituts für
Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

07. August 2014

Prof. Dr. Axel Mühlbacher
Hochschule Neubrandenburg

Hochschule Neubrandenburg
Brodaer Straße 2
17033 Neubrandenburg



Gliederung

Einleitung und Gegenstand	4
Begründung der Fokussierung der Stellungnahme	6
Stellungnahme Kapitel 4: Nutzen	10
Maß des Gesamtnutzens als relevante Entscheidungsgröße	10
Alternative Vorgehensweise für die Bestimmung des Gesamtnutzens	10
Skalenniveau des transformierten Nutzens	12
Stellungnahme Kapitel 3: Nutzenbewertung medizinischer Interventionen	14
Evidenzgrad über die Werturteile in der frühen Nutzenbewertung	14
Patientenzufriedenheit und Nutzendimensionen	14
Begriff und Definition des Nutzen	15
Fakultative Würdigung von Schaden und Nutzen	17
Würdigung der Nutzen- und Schadensaspekte	17
Beliebigkeit der Abwägung von Nutzen und Schaden	19
Stellungnahme Kapitel 2: Produkte des Instituts	20
Informationsbeschaffung und wissenschaftliche Bewertung	20
Potenzialbewertung: Zielkriterien und Werturteile	21
Potenzialbewertung: Indikationsübergreifende Vergleichsgrößen	22
Stellungnahme Kapitel 1: Das Institut	23
Darstellung der wissenschaftlichen Grundlagen	23
Die Bereitstellung von Informationen über Prioritäten und Werturteile	24
Anforderungen an die Evidenz der Werturteile und Präferenzen	25
Lösungsvorschlag: Messen, Bewerten & Entscheiden	26
Phasen der Nutzenbewertung	27
Empfehlungen von Arbeitsschritten bei der Nutzenbewertung	31
Anforderungen an das Skalenniveau	32
Zusammenfassung	33
Literatur / Quellen	36

Abkürzungsverzeichnis

AHP	Analytic Hierarchie Process
CA	Conjoint Analyse
EbM	Evidenzbasierte Medizin
EG	Effizienzgrenze
FNB	Frühe Nutzenbewertung
GBA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KNB	Kosten-Nutzen Bewertung
MdG	Maß des Gesamtnutzens
SGB	Sozialgesetzbuch
ZN	Zusatznutzen

Einleitung und Gegenstand

Seit der Vorstellung des ersten Entwurfs des Methodenpapiers im Jahr 2009 wurden zahlreiche Stellungnahmen geschrieben und die methodischen Grundlagen der Nutzenbewertung in Deutschland diskutiert. Führt 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. die Bestimmung des Ausmaßes des Zusatznutzens weitere methodische Fragen auf. Die am 28. November 2013 publizierte Fassung "Allgemeine Methoden 4.1" zeigte Änderungen im Kapitel 3 und dort die Operationalisierung des Ausmaßes eines Zusatznutzens (Kapitel 3.3.3), unter anderem durch die Methoden der Präferenzmessung und Ansätze der multikriteriellen Entscheidungsfindung. Im Anhang wurde eine umfassende Begründung zur Rationale der zugrundeliegenden Methodik der Operationalisierung angefügt. In dem jetzt vorliegenden Entwurf wurden die Methoden zur Kosten-Nutzen Bewertung integriert, wesentliche Inhalte des Kapitels 3 (frühe Nutzenbewertung) gestrichen und in einem neuen Kontext für das Kapitel 4 (Kosten-Nutzen Bewertung) aufbereitet.

Evidenzbasierte Medizin als Grundlage: Heute – fünf Jahre nach dem ersten Entwurf – sind die Frage- und Problemstellungen des Messens von Zielkriterien und Nutzendimensionen im Zusammenhang mit der evidenzbasierten Medizin ausführlich 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. 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.

Werturteile und Präferenzen: 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? Welche spezifischen Anforderungen werden an das Skalenniveau bei der frühen Nutzenbewertung bzw.

der Kosten-Nutzen Bewertung gestellt? 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, kann von einem erheblichen Handlungsbedarf ausgegangen werden.

Gegenstand der Stellungnahme: Das Institut weist zu Beginn darauf hin, dass „Nur die als Änderungen gekennzeichneten Textteile (...) Gegenstand des Stellungnahmeverfahrens“ sind (S. ii). Inhaltliche Änderungen in einigen Abschnitten mögen begründeten Änderungs- oder Klärungsbedarf in anderen Abschnitten nach sich ziehen (auch wenn diese nicht im neuen Entwurf geändert wurden). Im neuen Entwurf werden konkrete Verweise auf andere Abschnitte in den Änderungen gemacht. So wird im ersten Satz des Abschnitts 4.3 direkt auf den Abschnitt 3.3 verwiesen und festgestellt, dass in diesem Abschnitt die Methoden „mit denen der Nutzen im Rahmen der Nutzenbewertung bestimmt wird“ beschrieben werden (S. 86). Trotz der zahlreichen Änderungen in den Kapiteln 3 und 4 wurden in diesem Abschnitt keine Änderungen oder Anpassungen vorgenommen. In Abschnitt 3.1.5 wird wiederum für die themenspezifische Konkretisierung und gleichzeitige Würdigung von Nutzen und Schaden auf Abschnitt 4.3.3 verwiesen (S. 49). In Abschnitt 4.3.3 sollen die Methoden für die dem Kapitel 3 zugrundeliegenden Entscheidungen erläutert werden. Die Einschränkung der Stellungnahme auf die gelb markierten Bereiche ist nicht zielführend, da dem Stellungnehmenden allgemein auch die Möglichkeit eingeräumt werden sollte, auf Kritikpunkte der vergangenen Jahre einzugehen, wenn diese im Methodenpapier keine Berücksichtigung gefunden haben. Im Sinne der Verfasser sollte es auch sein, dass die Stellungnehmenden ebenso auf Verbesserungspotential oder Korrekturbedarf hinweisen, auch wenn die Textpassagen nicht Gegenstand der Änderungen waren. Neue wissenschaftliche Erkenntnisse können die Beurteilung des Methodenpapiers über die Zeit beeinflussen, gerade dann, wenn diese Abschnitte in der Fortschreibung der Methoden nicht berücksichtigt wurden. Eine Einschränkung der Stellungnahmen auf die tatsächlich geänderten Textstellen entspricht nicht der Idee die (Fach-)Öffentlichkeit an der Weiterentwicklung der Methoden zu beteiligen. Zur Konkretisierung der adressierten Frage- und Problemstellungen des aktuellen Entwurfes des Methodenpapieres in der Version 4.2 ist es notwendig zunächst einzelne Aspekte aus dem neuen Kapitel 4 herauszuarbeiten. Fortfolgend werden dann erst die resultierenden Fragen in den vorangehenden Kapiteln 3, 2, 1 diskutiert.

Begründung der Fokussierung der Stellungnahme

Für die Begründung weiterführender entscheidungsunterstützender Informationen als Grundlage für die Nutzenbewertung des IQWiG im Kontext der Entscheidungen der Selbstverwaltung, soll hier die Analogie „von zwei Läufern in einem fairen Rennen“ (S.7) im Methodenbericht aufgegriffen werden.

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 Rangordnung genutzt werden. Die Entscheidungsregeln haben 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.

Komplexes Entscheidungsproblem: 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 Disziplinen 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 wie das Kugelstoßen und das Kugelstoßen höher wie 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. Die „Schiedsrichter“ verfügen 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 „fairen 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: 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 Entbehren 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 GBA liegt. Konkret: Das IQWiG bewertet die Messung, also die

Begründung kausaler Effekte, und der GBA 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 GBA die Informationen über die Wertmaßstäbe der Bevölkerung und Präferenzen der Patienten benötigt. Diese Informationen sollten demnach Gegenstand des Dossiers und Gegenstand des Berichtes des IQWiG an den GBA sein.

Stellungnahme Kapitel 4: Nutzen

Maß des Gesamtnutzens als relevante Entscheidungsgröße

In Abschnitt 4.3.2 (Endpunkte) definiert das Methodenpapier das „Maß des Gesamtnutzens“, als „eine Aggregation der Bewertung von Nutzen und Schaden in einer Größe, wobei unterschiedliche patientenrelevante Endpunkte zu einem einzigen Maß zusammengefasst werden.“ (S. 87).

Position und Fragen: Unklar bleibt, ob die Aggregation eine notwendige Bedingung für die Entscheidungsfindung darstellt? Oder, ob eine Kosten-Nutzenbewertung auch sinnvoll ist, wenn mehrere Endpunkte „nebeneinander dargestellt werden“ (S. 87) und für jeden „patientenrelevanten Endpunkt eine eigene Effizienzgrenze erstellt“ (S. 87) wird?

Alternative Vorgehensweise für die Bestimmung des Gesamtnutzens

Abschnitt 4.3.3 (Maß des Gesamtnutzens) nennt das qualitätsadjustierte Lebensjahr (QALY), das behinderungsbereinigte Lebensjahr (DALY) oder das Healthy Years Equivalent (HYE) als unterschiedliche Maße, um den Gesamtnutzen einer Intervention auszudrücken.

Position und Fragen: Keinen Sinn macht die Feststellung, dass „je nach methodischem Ansatz oder ökonomischer Theorie in der wissenschaftlichen Literatur die Begriffe Präferenzen (Preferences), Nutzwerte bzw. Teilnutzenwerte (Utilities) oder Werte (Values) verwendet“ (S. 87) werden (Auch fraglich ist ein Verweis an dieser Stelle auf das Buch Drummond et al. 2005 mit annähernd 400 Seiten). Diese Begrifflichkeiten werden weder als Synonyme verwendet, noch haben sie in unterschiedlichen methodischen Ansätzen eine ähnliche Bedeutung. (Dies wird auf den annähernd 400 Seiten von Drummond et al. 2005 auch ausgeführt). Auf eine Diskussion der methodischen Grundlagen und Begrifflichkeiten wird im Methodenpapier gänzlich verzichtet. Diese Diskussion wäre hilfreich, da diese Begrifflichkeiten unterschiedliche Theorieansätze betreffen (z. B. bei der Definition des Nutzenbegriffes) und in unterschiedlichen Kontexten verwendet werden (so bedarf es der Gewichtung von klinischen oder nicht-klinischen Endpunkten durch Präferenzgewichte, damit ein Nutzwert aggregiert werden kann). Insofern sind in diesem Absatz die Ausführungen irreführend, da nicht alle der oben genannten Begrifflichkeiten unter dem Begriff „Gewicht“ abgehandelt werden können (siehe Aussage auf S. 87).

Position und Fragen: Es wäre sinnvoll zu erfahren, wie diese Gewichte "mit deren Hilfe einzelne patientenrelevante Endpunkte in ein Maß des Gesamtnutzens überführt werden können" (S. 87), wissenschaftlich hergeleitet werden? Es ist sicherlich falsch zu behaupten, dass der QALY ein derartiges Gewicht darstellen kann. Vielmehr werden wahlbasierte Verfahren (z. B. Standard Gamble, Time Trade off, Discrete Choice Experimente) eingesetzt, um Gewichte zu ermitteln, damit erhobene Gesundheitszustände in Nutzenwerte überführt werden können. Methoden der Präferenzmessung können sowohl bei der Gewichtung von nicht-klinischen Endpunkten (z. B. bei Gesundheitszuständen bzw. Lebensqualität) als auch klinischen Endpunkten (z. B. Zielkriterien oder Nutzendimensionen) eingesetzt werden.

Das Methodenpapier unterscheidet zwei Ansätze zur Dokumentation eines Maßes des Gesamtnutzens: A) den „QALY als Maß des Gesamtnutzens“ und B) die „Erhebung von Präferenzen zur Erstellung eines Maßes des Gesamtnutzens“. Im Kern geht es bei dieser Unterscheidung darum, dass bei der ersten Alternative patientenberichtete subjektive Angaben zur Lebensqualität bzw. dem Gesundheitszustand (nicht-klinische Endpunkte) mit Hilfe der Präferenzmessung gewichtet werden, um dann einen Nutzenwert abzuleiten. Das zweite Maß basiert auf einem breiteren Ansatz (hier bezeichnet mit „Multikriterielle Entscheidungsfindung“) und gewichtet klinische und nicht-klinische Endpunkte auf Grundlage wissenschaftlich abgeleiteter Präferenzen. Beiden Verfahren ist also gemein, dass nur mit Hilfe der Gewichtung ein Nutzenwert als Maß des Gesamtnutzens abgeleitet werden kann. Unterschiede ergeben sich durch die subjektive Datenerhebung (patient reported outcomes in A) im Gegensatz zur objektiven Datenerhebung (clinician reported outcomes in B). Für die Gewichtung des QALYs werden grundsätzlich nur theoriebasierte Verfahren der Präferenzmessung eingesetzt. Traditionell die Lotterieverfahren wie der Standard Gamble und Time Trade Off. In jüngster Zeit werden zunehmend fortgeschrittene Stated Preference Methoden, wie das Discrete Choice Experiment bzw. die wahlbasierte Conjoint Analyse genutzt. Für die Variante B) diskutiert das IQWiG die Verfahren des Analytic Hierachy Process sowie die (wahlbasierte) Conjoint Analyse. Das Methodenpapier verweist darauf, dass es „noch ungelöste methodische Probleme beim Einsatz dieser Verfahren“ gibt, „sodass gegenwärtig eine routinemäßige Anwendung dieser Methoden nicht vorgesehen ist“. Eine weitere Erläuterung dieser Aussage wird jedoch nicht vorgenommen. Eine wissenschaftliche Argumentation hinsichtlich dieser Aussagen bzw. weiterführende Zitate sind nicht zu finden. In diesem Kontext werden weitere Behauptungen aufgestellt, die weder durch Literatur untermauert sondern lediglich argumentativ gestützt werden.

Position und Fragen: Geht das IQWiG davon aus, dass beide genannten Ansätze zur Operationalisierung des Gesamtnutzens systematisch und konsistent zu gleichen Ergebnissen kommen? Auf welchem Weg können die Ergebnisse beider Verfahren vergleichbar gemacht werden? Vor dem Hintergrund, dass hier zum einen subjektive Daten (patient reported outcomes) als auch objektive Daten (clinician reported outcomes) bei der Transformation der Effekte in einen Nutzenwert zugrunde gelegt werden, stellt sich die Frage nach der Ausschließlichkeit: Spielen bei der Ableitung des Maßes für den Gesamtnutzen über QALYs die klinisch gemessenen Endpunkte keine Rolle mehr?

Folgende Aussagen sind falsch: „Aus der Wahl der Szenarien wird dann über ein Regressionsmodell ein Gewichtungsfaktor für jedes Attribut errechnet“ (S. 89). Das Regressionsmodell schätzt lediglich die Koeffizienten für die Ausprägungen (Level) der Attribute. „Für die CA gilt, dass maximal 6 bis 7 Attribute einbezogen werden können.“ (S. 89) Die einschlägige Literatur zeigt, dass diese Aussage falsch ist, es können und werden mehr Attribute verwendet. „Beim AHP-Verfahren gibt es diese Grenze nicht“ (S. 89). Auch der AHP ist im selben Ausmaß begrenzt (was ausdrücklich durch den Begründer Saaty festgehalten wurde (Saaty 1990, Seite 20), denn es können nicht unendlich viele Paarvergleichsfragen durchgeführt werden. „Weiterhin scheint das AHP-Verfahren die Befragten vor geringere kognitive Anforderungen zu stellen“ (S. 89). Diese Aussage ist falsch, da die Paarvergleiche vom Teilnehmer einer Befragung verlangen, dass er den Nutzen einer Eigenschaft auf einer 9-stufigen Likert Skala hinsichtlich ihres Zusatznutzens bewertet. Ferner müssen beim AHP stets zwei Entscheidungen getroffen werden: Zum einen, welches Kriterium hat die größere Bedeutung und zum anderen, um wie viel größer die Bedeutung des jeweiligen Attributes ist. Dagegen verlangt die wahlbasierte Conjoint Analyse nur die Entscheidung zwischen zwei Alternativen. Es ist in der einschlägigen psychologischen und ökonomischen Literatur bekannt, dass eine numerische Bewertung des Nutzens auf einer Kardinalskala für den einzelnen Teilnehmer mit sehr großen Problemen verbunden ist.

Skalenniveau des transformierten Nutzens

In Abschnitt 4.3.1 wird die Skala definiert, auf der ein transformierter Nutzen abgetragen werden soll. Über den „Bereich, der für die Definition des patientenrelevanten Zusatznutzens von Bedeutung ist“ (S. 86) soll der Nutzenwert auf einer kardinalen Nutzenskala abgetragen werden. Im Methodenbericht wird diese Skala mit „approximativ kardinalskaliert“ bezeichnet.

Position und Fragen: Diese Bezeichnung ist irreführend, da im relevanten Bereich eine Kardinalskala eingefordert wird. Nachdem die Kapitel 3 und 4 Eingangs in ihren methodischen Grundlagen verknüpft

wurden, gelten diese Anforderungen hinsichtlich des Skalenniveaus des Gesamtnutzens nur für die Kosten-Nutzen Bewertung oder auch (entsprechend der Verweise aus den Abschnitten in Kapitel 3) für die frühe Nutzenbewertung?

Fazit Kapitel 4: Die einführende Argumentation dieser Stellungnahme zeigt die Notwendigkeit der Berücksichtigung von Werturteilen und Präferenzen. Folgt man diesen Ausführungen, gilt es in diesem Abschnitt und in den Kapiteln 3 und 2 die Trennung zwischen „Messen“ und „Bewerten“ herauszustellen und deutlich die Anforderungen an die Methoden und Instrumente der Gewichtung von klinischen und nicht-klinischen Zielkriterien zu formulieren. Gleich zu Beginn des Abschnitts 4.3 wird festgestellt, dass die „Methoden, mit denen der Nutzen im Rahmen der Nutzenbewertung bestimmt wird, (...) in Abschnitt 3.3 beschrieben“ (S. 86) wurden. Diese Aussage ist unzureichend, da im Abschnitt 3.3 keine Aussagen zur Bewertung des Nutzens, sondern nur Aussagen hinsichtlich der richtigen Messung der Effekte gemacht werden (siehe die Hinweise im folgenden Abschnitt). Auf Basis dieses Verweises wird das Kapitel 3 auch umfassend in die Stellungnahme einbezogen.

Stellungnahme Kapitel 3: Nutzenbewertung medizinischer Interventionen

Evidenzgrad über die Werturteile in der frühen Nutzenbewertung

In Abschnitt 3.1.1 deutet das Methodenpapier rudimentär an, dass realisierte positive und negative Effekte sich gegenseitig kompensieren (S. 38 und 39). Konkreter, dass eine Verlängerung der Lebensdauer durch eine stärkere Nebenwirkung kompensiert werden kann. Zudem können „Nutzen- und Schadensaspekte (...) eine unterschiedliche Wichtigkeit für die Betroffenen haben“ (S. 40). In diesem Zusammenhang weißt das Methodenpapier lapidar auf die Quelle der zugrundeliegenden Information hin, „die sich ggf. durch die qualitative Erhebung oder bereits bei der Beratung durch Betroffene, Patientenvertretungs- und / oder Verbraucherorganisationen (...) abzeichnet“ (S. 40).

Position und Fragen: Welcher Evidenzgrad wird bei der Berücksichtigung dieser Werturteile toleriert? Reicht in diesem Kontext die anekdotische Evidenz der qualitativen Befragung von nicht weiter spezifizierten Personengruppen? In welcher Form wird diese Evidenz über die Werte und Prioritäten der Patienten dokumentiert? In welcher Form wird das durch qualitative Erhebungen gewonnene Wissen den Entscheidungsträgern zur Verfügung gestellt? In welcher Form werden diese Entscheidungsgrundlagen im Bericht an den GBA aufbereitet? Mit welcher wissenschaftlichen Begründung ist die qualitative Erhebung ausreichend – würden diese Anforderungen auch für den pharmazeutischen Unternehmer im Rahmen der Berichterstattung des Dossiers gelten?

Patientenzufriedenheit und Nutzendimensionen

Vorrangig zu berücksichtigende Zielgrößen sind nach § 35 Abs. 1b SGB V die Mortalität, die Morbidität und die Lebensqualität, da sie eine Feststellung krankheits- und behandlungsbedingter Veränderungen erlauben (S. 39). Die „Zufriedenheit der Patientinnen und Patienten“ wurde gestrichen.

Position und Fragen: Welche wissenschaftlichen Erkenntnisse haben dazu geführt, die Patientenzufriedenheit nach fünf Jahren zu streichen und zukünftig nicht mehr als ergänzende Zielgröße anzuerkennen?

Wie die Patientenzufriedenheit, werden ohne Kommentar, die in der Gesetzgebung konkretisierenden Nutzendimensionen des § 35b SGB V (Verbesserung des Gesundheitszustandes, Verkürzung der

Krankheitsdauer, Verlängerung der Lebensdauer, Verringerung der Nebenwirkung, Verbesserung der Lebensqualität) gestrichen.

Position und Fragen: Welche wissenschaftlichen Erkenntnisse haben dazu geführt, diese konkretisierenden Nutzendimensionen im Kapitel 3 zu löschen? Werden diese Nutzendimensionen nicht mehr anerkannt?

Begriff und Definition des Nutzen

Grundsätzlich bleibt unklar, welche Zielkriterien aus den Zielgrößen oder Nutzendimensionen abgeleitet werden können. Die Ableitung von Zielkriterien bzw. patientenrelevanten Endpunkten ist abhängig von der dem Methodenpapier zugrundeliegenden Definition des Nutzens. Im Methodenbericht wird der Begriff „Nutzen“ als „kausal begründete positive Effekte“ einer medizinischen Intervention auf patientenrelevante Endpunkte definiert (S. 38). Hier scheint eine Definition gewählt, die sich ausschließlich an den klinischen Effekten orientiert, also ein „medizinischer Nutzenbegriff“ (soweit es diesen als eigenständige Definition geben kann). Vermutlich umfasst diese Definition nicht den ökonomischen Nutzen, der beim Patienten entsteht oder das Konzept des Erwartungsnutzens. Im Internet Glossar definiert das IQWiG einen „engen“ und einen „weiten“ Nutzenbegriff. „In seiner engen Bedeutung lehnt er sich an die EbM an und spiegelt den reinen medizinischen Nutzen zur Beurteilung einer Maßnahme wider (= Gesundheitseffekte / Outcomes)“ (IQWiG Glossar im Internet). Auf diesen Nutzenbegriff wird zu Beginn des Kapitels 3 zurückgegriffen. Erst in seiner weiten Bedeutung bezieht sich der Nutzenbegriff auf den beim Patienten entstandenen Nutzen oder den Erwartungsnutzen (unter Unsicherheit). Der Nutzen wird dann nicht nur als ein Gesundheitseffekt einer Intervention definiert, „sondern berücksichtigt auch den Wert, den der Patient diesem Effekt zuschreibt (sogenannte Nutzwerte / Utilities)“ (IQWiG Glossar im Internet). Dieser Begriff lehnt sich scheinbar an die neoklassische Theorie an, wonach der Nutzen nicht mehr über die Befriedigung von Bedürfnissen definiert ist, sondern sich durch die Wahlentscheidung zwischen zwei Handlungsalternativen definiert, d.h. über den Wert den der Konsument einer Intervention, Handlung oder Alternative beimisst. Wird eine Behandlung A einer Behandlung B vorgezogen, so hat A aus der Perspektive des Entscheiders (z.B. Arzt oder Patient) einen höheren Nutzen. Die Präferenz für A resultiert aus einem höheren Nutzen von A im Vergleich mit B (aus der Perspektive des Entscheiders). Der durch den Experten oder Konsumenten prognostizierte Nutzen muss nicht immer dem tatsächlich eintretenden Nutzen entsprechen. Man spricht vom Erwartungsnutzen, wenn zum Zeitpunkt der Entscheidung aufgrund von Unsicherheit der Nutzen nicht genau quantifiziert werden kann.

Position und Fragen: Es wäre sinnvoll zu prüfen, ob die dem Methodenpapier zugrundeliegende Definition des Nutzens neben der EbM auch den gesundheitsökonomischen Inhalten gerecht werden könnte. Der „reine medizinische Nutzen“ scheint hinsichtlich seines Wertes für den Patienten unbestimmt. Die Verbesserung des Patientennutzens ist jedoch eine relevante Handlungsmaxime. Der „enge“ Nutzenbegriff orientiert sich an den gemessenen klinischen Effekten und nicht an dem Wert, den ein Entscheider oder Konsument der Intervention zuweist (insofern ist dieser Nutzenbegriff nicht in der Ökonomie verankert). Man kann nicht davon ausgehen, dass dieser medizinische Nutzenbegriff mit dem Patientennutzen gleichzusetzen ist (gemessene patientenrelevante Endpunkte spiegeln genauso wenig den Patientennutzen wieder, wie Motorleistung und Benzinverbrauch den Nutzen eines Autos charakterisiert. Es kommt auf die Bewertung der Effekte aus der Perspektive des Konsumenten an).

Zudem setzt das Glossar im Internet den „weiten“ Nutzenbegriff gleich mit dem Nutzwert. Der neoklassische und damit vorherrschende Nutzenbegriff leitet sich aus der Wahlhandlung bzw. über die Präferenzen ab. Der Nutzwert orientiert sich am klassischen Begriff, der durch das Ausmaß der Eignung eines Gutes zur Befriedigung der Bedürfnisse bestimmt wird.

Position und Fragen: Offen bleibt dann, welcher ökonomische Nutzenbegriff handlungsleitend für die weiteren Ausführungen des Methodenpapiers ist. Weiterhin führt das Institut im Internet Glossar aus, dass diese „Nutzwerte (...) zur Gewichtung von Effekten genutzt werden und hierdurch die Bedeutung eines Effektes aus Sicht der Betroffenen erhöhen oder reduzieren“ können. Diese Feststellung ist falsch, da in die Berechnung der Nutzwerte bereits die gewichteten Effekte der Intervention eingegangen sind.

Haben die Ausführungen über den Nutzen oberflächlich gesehen wenig mit den geänderten Textzeilen im Entwurf des Methodenpapiers 4.2 zu tun, zeigt sich im folgenden, dass Missverständnisse aus der unterschiedlichen Definition des Nutzens gerade in den Kapiteln 3 und 4 nicht aufgelöst werden können. Zudem stellt sich für den Leser die Frage, ob beide Kapitel mit dem gleichen Nutzenbegriff arbeiten oder ob Kapitel 3 (zur FNB) und Kapitel 4 (zur KNB) einen unterschiedlichen Nutzenbegriff zu Grunde legen. Nach der Lektüre könnte man zu dem Ergebnis kommen, dass bei der frühen Nutzenbewertung der „enge“ Nutzenbegriff handlungsleitend ist, wobei bei der Kosten-Nutzen Bewertung der „weite“ Nutzenbegriff umgesetzt werden soll. Eine Begründung für die unterschiedliche Betrachtung ist in den Ausführungen nicht zu finden.

Fakultative Würdigung von Schaden und Nutzen

Der Begriff der „Nutzenbewertung“ wird definiert als der gesamte „Prozess der Evaluation medizinischer Interventionen hinsichtlich ihrer kausal begründeten positiven und negativen Effekte im Vergleich mit einer klar definierten anderen Therapie, einem Placebo (oder einer andersartigen Scheinbehandlung) oder keiner Behandlung. Dabei werden Nutzen- und Schadensaspekte zunächst endpunktbezogen evaluiert und dargestellt. Darüber hinaus ist eine gemeinsame Würdigung der endpunktbezogenen Nutzen- und Schadensaspekte möglich (...)“ (S. 40)

Position und Fragen: Was heißt in diesem Kontext „gemeinsame Würdigung“? Warum ist eine gemeinsame Würdigung von Nutzen- und Schadensaspekten „möglich“ – ist eine Entscheidung ohne die Abwägung von Nutzen und Schaden für das IQWiG denkbar?

Würdigung der Nutzen- und Schadensaspekte

„Schaden“ wird definiert als kausal begründete, negative Effekte einer medizinischen Intervention auf patientenrelevante Endpunkte. Abschnitt 3.1.3 begründet die Relevanz der unerwünschten Wirkungen, unter anderem durch die Tatsache, dass der Nutzen teilweise oder ganz durch den Schaden aufgewogen werden könnte. Von einer Relevanz der negativen Effekte ist dann auszugehen, wenn unter anderem die Effekte „von den Patientinnen und Patienten als besonders wichtig angesehen werden“ (S. 43). Die Auswahl der potenziell relevanten unerwünschten Wirkungen erfolgt unter anderem auf Basis dieses Kriteriums (S. 44).

Position und Fragen: Es bleibt offen, welche wissenschaftlichen Methoden und Instrumente zur Dokumentation der Wichtigkeit einer unerwünschten Wirkung herangezogen werden können? Bei einem unklaren Nutzenbegriff bleibt es offen, welche Nutzendimension und Zielkriterien bei der Abwägung von Nutzen und Schaden zur Bestimmung des Zusatznutzens gegenüber einer Alternative eingehen. Die gemeinsame Würdigung von Nutzen und Schaden verlangt einen transparenten Prozess, der die Wertentscheidungen, d.h. der die Gewichtung der positiven und negativen Effekte, offen legt. Wertentscheidungen der Abwägung sollten nicht Gegenstand „subjektiv geprägte(r) und damit häufig verzerrte(r) Bewertung(en)“ sein (dies ist auf S.4 in Abschnitt 1.2 als wesentliche Forderung der EbM vorgegeben). Die Gewichtung der Entscheidungskriterien als Grundlage des Abwägungsprozesses, ist die Voraussetzung für die gemeinsame Würdigung der endpunktbezogenen Nutzen- und Schadensaspekte.

Bei der Definition des Begriffes der Nutzenbewertung wird für die Details der „gemeinsamen Würdigung der endpunktbezogenen Nutzen- und Schadensaspekte“ auf Abschnitt 3.1.4 verwiesen (S. 40). In Abschnitt 3.1.4 wird die Abwägung zwischen Nutzen und Schaden anhand der Stärke der Ergebnisunsicherheit vorgenommen. Dabei wird der (Zusatz-)Nutzen und Schaden zusammenfassend anhand der Beleglage in vier Abstufungen bezüglich der Aussagesicherheit getroffen (Beleg, ein Hinweis, ein Anhaltspunkt oder keine der drei Situationen). Anhand der Gewichtung der Ergebnissicherheit sollen die gemessenen Effekte zusammenfassend beschrieben werden. Diese Operationalisierung der Ergebnissicherheit, als Gewichtungsparemeter der Studienergebnisse, dient der Beurteilung der Qualität des Messens.

Position und Fragen: Kapitel 3.1.4 geht nicht darauf ein, wie Nutzen und Schaden aufgrund von Werturteilen und Prioritäten gegeneinander abgewogen werden können. Es gibt keinen Hinweis darauf, wie eine Würdigung der endpunktbezogenen Nutzen- und Schadensaspekte hinsichtlich der Wichtigkeit der gemessenen Effekte für den Konsumenten Eingang in die Bewertung findet. Die Relevanz der einzelnen Zielkriterien aus Sicht der Patienten oder Experten wird nicht systematisch berücksichtigt. Anstatt der systematischen Berücksichtigung von Expertenurteilen und Patientenpräferenzen, erfolgt eine implizite Gewichtung durch die Beurteilung der Ergebnissicherheit des Messens. Eine transparente Bewertung der Nutzen- und Schadensaspekte anhand expliziter Informationen, welche Endpunkte von den Patienten als besonders wichtig angesehen werden, ist nicht Gegenstand der Abwägung (siehe Forderung in 3.1.3). Eine Transparenz hinsichtlich der Gewichtung der patientenrelevanten Endpunkte durch die Experten ist in Abschnitt 3.1.4 auch nicht vorgesehen.

Diese Bewertung deckt sich auch mit der im Anhang dargestellten „Rationale der Methodik zur Feststellung des Ausmaßes des Zusatznutzens“. Hier wird die „Hierarchisierung der Zielgrößen“ (S. 206) und die Strukturierung entsprechend der „Bedeutung“ (S. 208) auf Basis der Messung der Effektstärken vorgenommen (keine systematische und strukturierte Gewichtung der Studienergebnisse). Die Ergebnisskala differenziert zwischen „erheblich“, „beträchtlich“ und „gering“. Eine Zuweisung der Effekte der Gesamtmorbidität, der schwerwiegenden Symptome und Nebenwirkungen, der gesundheitsbezogenen Lebensqualität und der nicht schwerwiegenden Symptome und Nebenwirkungen erfolgt über unbestimmte Begriffe (S. 209 und 210). Irrenführend ist auch der Verweis auf S. 208, dass das schrittweise Vorgehen einer Gesamtabwägung, d.h. die Zusammenfassung zu einer globalen Aussage zum Ausmaß des Zusatznutzens, in Abschnitt 3.3.3 beschrieben wird. Ein Hinweis zur systematischen Aggregation auf Basis wissenschaftlicher

Erkenntnisse zur Gewichtung der Endpunkte hinsichtlich der Relevanz oder Wichtigkeit ist in diesem Abschnitt nicht zu finden. In Abschnitt 3.3.3 wird der Prozess der Nutzenbewertung erläutert, ohne jedoch die Methoden oder Arbeitsschritte zur Herleitung des (Zusatz-)Nutzens zu erläutern. So hoch die Anforderungen an den Evidenzgrad und die Dokumentation der Unsicherheit sind, so gering sind die Anforderungen an die Evidenz hinsichtlich der Werte und Prioritäten der Bürger. Evidenz auf Basis der Einzelmeinungen von Patientenvertretern, Gutachten von Patientenorganisation und Surveys unter den zugelassenen Organisationen nähren den Verdacht, dass man den einzelnen Patienten und deren Vertretern mehr traut als den klinischen Experten. (Oder wie ist es zu erklären, dass unterschiedliche Anforderungen hinsichtlich der medizinischen Evidenz und der Evidenz über die Patientenpräferenzen gestellt wird?; vgl. hierzu Abschnitt 1.2.7).

Beliebigkeit der Abwägung von Nutzen und Schaden

In Abschnitt 3.1.5 wird wiederum die Nutzen-Schaden Abwägung als bewertendes Fazit zur zusammenfassenden Bewertung der Beleglage für jeden patientenrelevanten Endpunkt aufgegriffen. Basierend auf den Zielgrößen (siehe Abschnitt 3.1.1) stellt das Institut den Nutzen, den Schaden und ggf. die Nutzen-Schaden-Abwägung dar „soweit dies aufgrund der vorliegenden Daten möglich ist“ (S. 49). Ohne wissenschaftlich belastbare Aussagen zu Wertmaßstäben und Präferenzen für jeden patientenrelevanten Endpunkt ist eine gleichzeitige Bewertung von Nutzen und Schaden nicht möglich. Unklar bleibt wieder, wie die Gegenüberstellung der endpunktbezogenen Nutzen- und Schadensaspekte „qualitativ oder semiquantitativ wie in der frühen Nutzenbewertung nach § 35a SGB V“ erfolgt. Die wissenschaftlichen Methoden und Verfahren sind im Methodenbericht nicht konkretisiert.

Position und Fragen: Im Methodenbericht wird weiter ausgeführt, dass eine weitere Möglichkeit der gleichzeitigen Würdigung darin besteht, „die verschiedenen patientenrelevanten Endpunkte zu einem einzigen Maß zu aggregieren oder über ihre Gewichtung zu einer Gesamtaussage zu kommen.“ (S. 49) Diese Formulierung ist schwer vor dem Hintergrund der gängigen Methoden zu interpretieren, da die Aggregation von Effekten nur über die Gewichtung der Zielerreichungsgrade jedes Endpunktes möglich ist. Die Zusammenhänge der Gewichtung und Aggregation werden hier missverständlich dargestellt. Für die themenspezifische Konkretisierung und gleichzeitige Würdigung von Nutzen und Schaden wird weiter auf Abschnitt 4.3.3 verwiesen. Ist diesem Verweis zu entnehmen, dass die in Abschnitt 4.3.3 dargestellten Methoden auch der Kapitel 3 zugrundeliegenden Entscheidungen zur frühen Nutzenbewertung als Basis dienen? Demnach gelten für Kapitel 3 die gleichen Frage- und Problemstellungen wie für Kapitel 4.

Stellungnahme Kapitel 2: Produkte des Instituts

Informationsbeschaffung und wissenschaftliche Bewertung

Nach der Übermittlung des Dossiers bezieht sich die inhaltliche Bewertung des Dossiers neben der medizinischen Expertise auch auf die Einbeziehung der Patientenperspektive. In Abschnitt 2.1.4 zeigt das Ablaufdiagramm (Abb. 4) zur Kosten-Nutzen Bewertung nach § 35b SGB V die Phase der „Informationsbeschaffung und wissenschaftlichen Bewertung“. Es kann sich in diesem Kontext nicht mehr um die Messung der klinischen Effektparameter handeln, da diese bereits in klinischen Studien gemessen und im Dossier des pharmazeutischen Unternehmers dokumentiert wurden. Die wissenschaftliche Bewertung kann sich zum einen auf die Bewertung der Qualität (Unsicherheit) der Messung beziehen, zum anderen könnte hier auch die Bewertung der Relevanz der Zielkriterien bzw. der Endpunkte (Gewichtung) für die Aggregation des Gesamtnutzens gemeint sein. Beide Aktivitäten sind anzunehmen, da das Ablaufdiagramm vorsieht, dass sowohl die medizinische Expertise über „externe Sachverständige“ als auch die „Patientenperspektive über Patienten / Patientenvertreter“ einbezogen werden sollen.

Position und Fragen: Da beide o.g. Aktivitäten eine wesentliche methodische Voraussetzung für die Dokumentation des Patientennutzens darstellen, stellt sich hier die Frage, ob nicht beide Aktivitäten getrennt im Ablaufdiagramm gewürdigt werden sollten. Diese Maßnahme würde nicht nur die Interpretation des Ablaufdiagramms auf Seite 23 erleichtern, sondern auch die Ablaufdiagramme auf den Seiten 16 und 19. Klärungsbedarf besteht dann auch in Abgrenzung zur Abbildung 3 (Ablauf der Erstellung einer Dossierbewertung). Dort wird neben den externen Sachverständigen auch die Patientenperspektive durch Patienten und Patientenvertreter vertreten. Dabei nennt sich die Ablaufphase jetzt „Bewertung des Arzneimittels auf Basis der Dossierinhalte“. Im Text fehlt die Konkretisierung der Begrifflichkeit „wissenschaftliche Bewertung“. Das Weglassen der Begrifflichkeit „wissenschaftlich“ deutet darauf hin, dass eine Bewertung hinsichtlich der Relevanz der Zielkriterien (oder Endpunkte) nicht mehr auf wissenschaftlichen Kriterien beruhen sollte. Warum werden wissenschaftliche Informationen zur Patientenperspektive nicht explizit im Dossier ausgewiesen?

Entgegen der unterschiedlichen Darstellung in den Ablaufdiagrammen, beschreiben die Abschnitte 2.1.3 und 2.1.4 dann die „Einbeziehung der Patientenperspektive“ deckungsgleich. Auf Basis eines Fragebogens werden Patienten und Patientenorganisationen, welche über die in § 140f SGB V

benannten maßgeblichen Organisationen identifiziert wurden, zu relevanten Endpunkten und wichtigen Subgruppen befragt. Gelten hier andere Anforderungen an die Daten als im Abschnitt 1.2.7 (Nutzen im Einzelfall)? Auf Basis welcher wissenschaftlichen Methoden wird hier die Evidenz über Werturteile und Präferenzen auf Basis der Antworten von Patienten und Patientenorganisationen generiert? Grenzen sich die Änderungen in 2.1.4 gegenüber der Dossierbewertung bei der Berücksichtigung der Patientenperspektive im Ablaufschema ab? Gelten für die Abwägung des Nutzens und Schadens für die frühe Nutzenbewertung (Ausmaß des Zusatznutzens) die gleichen methodischen Voraussetzungen, wie bei der Abwägung von Nutzen und Schaden in der Kosten-Nutzen Bewertung (Maß des Gesamtnutzens)?

Potenzialbewertung: Zielkriterien und Werturteile

Neue Inhalte, wie die Methoden zur Potenzialbewertung in Abschnitt 2.1.5 (S. 25), offenbaren alte und neue Frage- und Problemstellungen. Damit Innovationen schneller in der Regelversorgung eingesetzt werden, können auf Antrag Erprobungsrichtlinien zu neuen Untersuchungs- und Behandlungsmethoden beschlossen werden (§ 137e SGB V). Neben dem Erlaubnisvorbehalt (§ 135 SGB V) in der vertragsärztlichen Versorgung und dem Verbotsvorbehalt (§ 137c SGB V) in der stationären Versorgung können so neue Untersuchungs- und Behandlungsmethoden in den Leistungskatalog der gesetzlichen Krankenversicherung aufgenommen und zeitlich befristet vergütet werden (insofern entscheidet das Potenzial über die Erstattungsfähigkeit). Der Vorteil der Potenzialbewertung besteht darin, dass ein Nutzen nicht hinreichend belegt sein muss. Ausreichend ist das Potenzial einer Behandlungsalternative, welche durch das IQWiG bewertet und durch den GBA bei der Entscheidung berücksichtigt wird.

Position und Fragen: Das Methodenpapier macht deutlich, dass bei der Dokumentation des Potenzials nicht die gleichen Anforderungen an die Evidenz gestellt werden, wie bei der Nutzenbewertung von Arzneimitteln. Es wird darauf eingegangen, wie das Potenzial einer Methode gemessen werden kann, es werden keine Aussagen darüber getroffen, wie die gemessenen Effekte in die Bewertung des Potenzials letztendlich eingehen (wie die Werturteile über die Effektparameter in die Bewertung eingehen). Es bleibt unklar, welche Zielkriterien für die Bewertung des Potenzials herangezogen werden, ob diese Zielkriterien bei der Potenzialbewertung gewichtet werden und wie ein eindimensionales Maß für das Potenzial bestimmt werden kann.

Potenzialbewertung: Indikationsübergreifende Vergleichsgrößen

Mit den Beiträgen der Versicherten soll über einen zeitlich befristeten Zeitraum und mit einem abgestimmten Studiendesign Evidenz generiert werden. Die begründete Erwartung einer effektiven Behandlung wird anhand der Morbidität, Mortalität und Lebensqualität gemessen. Eine positive Beurteilung des Potenzials nach Antragstellung begründet keinen Anspruch auf Erprobung; alle eingereichten Anträge werden auf der jährlichen GBA Haushaltsberatung bewertet.

Position und Fragen: Auf Basis eines Vergleichs der Potenziale der Methoden und der Wahrscheinlichkeit einer erfolgreichen Erprobung werden die Methoden zur Erprobung ausgewählt. Damit genügt nicht alleine eine Überschreitung eines zuvor definierten Schwellenwertes für die positive Bewertung. Die innerhalb eines Jahres eingereichten neuen Untersuchungs- und Behandlungsmethoden werden entsprechend des Potenzials verglichen. Bei der Potenzialbewertung werden indikationsübergreifende Vergleiche angestrebt. Demnach muss die Methode der Potenzialbewertung einen indikationsübergreifenden Vergleich von Untersuchungs- und Behandlungsmethoden ermöglichen. Die Methode muss im Ergebnis mindestens die Rangreihung anhand des gemessenen und bewerteten Potenzials über alle Entscheidungsalternativen garantieren. Diese Anforderung für die Methoden ist neu und in den Kapitel 3 und 4 in dieser Form nicht umgesetzt. Indikationsübergreifende Vergleiche wurden bis heute als nicht sinnvoll erachtet und im Sinne des SGB V bislang abgelehnt. Es bleibt bei der Potenzialbewertung unklar, wie diese multiplen Zielkriterien zu einer indikationsübergreifenden Kennzahl für das „Potenzial“ aggregiert werden, um dann auf Basis dieser Information eine Entscheidung über die Durchführung einer Erprobungsstudie zu treffen. Es bleibt unklar in welchem Umfang sich die Methode der Potenzialbewertung auf die methodischen Erläuterungen in Kapitel 3 und Kapitel 4 bezieht. Welche methodischen Grundlagen werden der Potenzialbewertung zugrunde gelegt?

Stellungnahme Kapitel 1: Das Institut

Darstellung der wissenschaftlichen Grundlagen

Bei der Nutzenbewertung geht es darum, die notwendigen Informationen für eine evidenzbasierte Entscheidung zu generieren. In Abschnitt 1.2.1 werden aus der evidenzbasierten Medizin die Aufgaben des Instituts abgeleitet. Das Institut führt eine objektive Bewertung durch und dokumentiert, „mit welcher Sicherheit der Nutzen medizinischer Maßnahmen nachgewiesen“ (S.5) worden ist.

Position und Fragen: Die Kommentierungen in den Kapitel 4 und 3 legen offen, dass es hierbei auch um die Beurteilung der Informationen über die folgenden Fragestellungen gehen müsste: Welche Endpunkte sind aus der Perspektive der Betroffenen und Experten relevant? Wie wurde die Relevanz der Entscheidungskriterien im Dossier belegt? Mit welcher Bedeutung gehen diese Endpunkte in einer Gesamtbetrachtung ein? Wie kann ein Gesamtnutzen oder Zusatznutzen auf Basis der Werturteile über die Zielkriterien abgeleitet werden? Darüber hinaus ist darzulegen, unter welchen Annahmen der Nutzen auf eine Dimension des Gesamtnutzens reduziert worden ist und wie die Effekte der Zielkriterien aggregiert wurden?

Es geht in diesem Kontext auch um gesundheitsökonomische und entscheidungstheoretische Fragestellungen, die neben der evidenzbasierten Medizin (auch fünf Jahre nach dem ersten Entwurf des Methodenpapiers) in den einführenden Darlegungen der wissenschaftlichen Grundlagen bislang keinen Platz finden. Es geht bei der Nutzenbewertung nicht nur um den kausalen Nachweis eines Effektes bzw. die Messung von Effektgrößen, sondern auch darum, unter welchen Prämissen Nutzen und Schaden abgewogen werden können, um mit dieser Information eine rationale Entscheidung zu ermöglichen.

Damit die Methode der Nutzenbewertung ausreichend zuverlässige Belege zur Verfügung stellt, dass „Therapie A für Patientinnen und Patienten mit einer bestimmten Krankheit besser ist als Alternative B“, müssen neben den Methoden der EbM auch die Methoden der Gesundheitsökonomie erläutert werden. Sowohl der Gesetzgeber, als auch das Methodenpapier bezeichnet neben der evidenzbasierten Medizin, die Gesundheitsökonomie als wissenschaftliche Basis der Nutzenbewertung. In Kapitel 1 des Methodenpapiers werden jedoch nur die Grundlagen der evidenzbasierten Medizin herangezogen um die methodischen Grundlagen des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) näher zu erläutern.

Würde man die wissenschaftlichen Grundlagen der Gesundheitsökonomie berücksichtigen, würde dies eine wesentliche Strategie der EbM unterstreichen. Ein grundlegendes Prinzip der EbM ist die Berücksichtigung der Präferenzen der betroffenen Patienten. Die Notwendigkeit der Abwägung von Nutzen und Schaden auf Basis der Patientenpräferenzen ist bereits in den Ausführungen zur evidenzbasierten Medizin explizit genannt (siehe hier die Erläuterungen in Abschnitt 1.2.1). Der Patient gilt im Sinne der evidenzbasierten Medizin als Betroffener, dessen Präferenzen strukturiert und systematisch zu berücksichtigen sind. Leider werden die Präferenzen als theoretische Schnittstelle der evidenzbasierten Medizin und der Gesundheitsökonomie in Kapitel 1 nicht diskutiert. In Abschnitt 1.2.4 (die Strategien der EbM) müssten dann weitere charakteristische Standardelemente aufgenommen werden. Die Punkte (1) bis (4) befassen sich nur mit dem Messen, jedoch nicht mit dem Bewerten der klinischen Evidenz. Wenn die EbM letztendlich darauf verweist, dass die Präferenzen der Betroffenen im Kontext der Behandlungsentscheidungen eine Rolle spielen müssen, dann könnte die Bewertung der Zielkriterien auch im Abschnitt 1.2.4 in einem gesonderten Punkt als zentrale Strategie aufgegriffen werden.

Die charakteristischen Standardelemente in Abschnitt 1.2.3 sind somit auch um die gesundheitsökonomischen Problem- und Fragestellungen zu erweitern. So müsste die Vergleichbarkeit von unterschiedlichen Messergebnissen, die Gewichtung von unterschiedlich gemessenen Zielerreichungsgraden und letztendlich die Definition einer Aggregationsregel zur Analyse eines Nutzenmaßes als eindimensionale Kennzahl auf einer ordinalen oder kardinalen Skala, diskutiert werden.

Die Bereitstellung von Informationen über Prioritäten und Werturteile

In Abschnitt 1.2.5 stellt das Methodenpapier die Bedeutung der Ergebnissicherheit für die Nutzenbewertung heraus. Im Kontext der Anforderungen an die Präzision und Verlässlichkeit kommt dem GBA die Aufgabe zu, die Unsicherheit bei der Entscheidungsfindung zu beachten. Hier konkretisiert das Methodenpapier: „Neben der wissenschaftlichen Sachlage bezieht er auch andere Aspekte in seine Entscheidung ein, wozu neben der Wirtschaftlichkeit auch die Bedürfnisse und Werte der Menschen gehören. In einer wissenschaftlich unsicheren Situation gewinnen diese Aspekte an Gewicht“ (S. 8).

Position und Fragen: In diesem Kontext müsste klar herausgestellt werden, inwiefern diese Informationen über die Bedürfnisse und Werturteile durch den Antragsteller bereitgestellt werden bzw. wie diese Informationen für die Entscheidungsträger bewertet werden. Werturteile bzw. die Bedürfnisse der Menschen dienen der Bewertung der klinischen Evidenz. Sie sollten systematisch für die Bestimmung des Ausmaßes des Zusatznutzens und des Maßes für den Gesamtnutzen im Dossier bzw. der Kosten-Nutzen Bewertung bereitgestellt werden. Die Notwendigkeit derartiger Studien für die Dokumentation des Ausmaßes des Zusatznutzens und des Maßes des Gesamtnutzens, wird in den folgenden Kapiteln nur rudimentär berücksichtigt. Nach der Verschiebung der wesentlichen Textstellen, könnte aus den Ausführungen geschlossen werden, dass unterschiedliche Anforderungen an die frühe Nutzenbewertung gestellt werden sollen. Dieser verkürzten Berücksichtigung der Bedürfnisse der Menschen sollte bereits in Kapitel 1 klar entgegengetreten werden.

Anforderungen an die Evidenz der Werturteile und Präferenzen

Im Abschnitt 1.2.7 (Nutzen im Einzelfall) stellt das Methodenpapier fest, dass auf „Einzelfällen basierende Erfahrungen (...) für eine Nutzenbewertung ungeeignet“ sind, „da es nicht möglich ist, Ergebnisse im Einzelfall (d.h. ohne Vergleich) einer Intervention zuzuschreiben.“ Treffender Weise hält der Methodenbericht fest, dass Erfahrungen, welche auf Einzelfällen basieren nicht für die Nutzenbewertung geeignet sind.

Position und Fragen: Leider versäumen die Autoren in diesem Kontext darauf hinzuweisen, dass im Status Quo die Prioritäten und Werte der Patienten auf Einzelmeinungen und Einzelfällen beruhen (Vgl. hiermit auch die Anforderungen an die Dokumentation der Wichtigkeit in 3.1.4, weiter unten im Text). Das Methodenpapier sollte zukünftig darauf verweisen, dass es wissenschaftlicher Studien bedarf, um die Präferenzen und Werturteile systematisch wissenschaftlich zu erfassen und zu dokumentieren. Inwieweit wissenschaftliche Methoden zur Dokumentation der Patientenpräferenzen als Informationen über die Prioritäten und Werte bei der frühen Nutzenbewertung herangezogen werden, bleibt offen.

Lösungsvorschlag: Messen, Bewerten & Entscheiden

In Abschnitt 3.3.3 wird das eigentliche Selbstverständnis des Methodenpapiers deutlich: „Die Nutzenbewertung erfolgt auf Basis der im vorliegenden Methodenpapier beschriebenen Standards der evidenzbasierten Medizin, die Bewertung der Kosten auf Basis der Standards der Gesundheitsökonomie.“ (S. 59) Hier geht das Institut von einer klaren Trennung zwischen den Disziplinen aus. Diese Trennung erscheint vor dem Hintergrund der wissenschaftlichen Erkenntnisse zur Nutzenbewertung falsch. Wenn die wissenschaftlichen Instrumente und Methoden zur Erstellung und Diskussion des Studiendesigns und der Messung von klinischen und nicht-klinischen Endpunkten in der evidenzbasierten Medizin große Fortschritte gemacht haben, so ist es doch die Ökonomie die in Theorien, Methoden und Instrumenten die Bewertung des Konsumentennutzens zum Gegenstand hat. Der Nutzenbegriff ist in den Medizinwissenschaften nicht einheitlich belegt, wogegen der Nutzenbegriff in den Wirtschaftswissenschaften Gegenstand eines langanhaltenden wissenschaftlichen Diskurses ist.

Diese strikte Trennung zwischen den beiden Disziplinen resultiert in Lücken und offenen Fragen zur Operationalisierung des Nutzens, unabhängig davon, ob er bei der frühen Nutzenbewertung oder der Kosten-Nutzen Bewertung eingesetzt wird. Diese Lücken hatten sich bereits bei der Stellungnahme zum Methodenpapier 4.1 offenbart: Bereits in den vorangegangenen Stellungnahmen wird gewünscht, dass das Vorgehen bei der Ableitung einer Gesamtaussage zum Ausmaß des Zusatznutzens genauer beschrieben wird. Das IQWiG gesteht ein, dass eine „Operationalisierung für diesen Schritt der Gesamtschau (...) zurzeit nicht (...) möglich“ ist (Würdigung der Stellungnahmen des Methodenpapiers 4.1, S. 8). Die Begründung des IQWiG: „Für die Gesamtabwägung existiert keine allgemein akzeptierte Methodik.“ (Würdigung der Stellungnahmen des Methodenpapiers 4.1, S. 8) Was genau unter einer allgemein akzeptierten Methodik verstanden wird bleibt unklar. Festzustellen ist, dass kaum eine der im Methodenpapier vorgestellten Methoden „allgemein“ akzeptiert wird. Gerade die Operationalisierung zur Feststellung des Ausmaßes des Zusatznutzens wird in den Stellungnahmen als zu starr kritisiert. Die vorgeschlagene indikationsübergreifende Gewichtung widerspricht jedem ethischen Empfinden, da hier deutlich wird, dass diese Gewichtung unabhängig von der Indikation und unabhängig von den zur Bewertung herangezogenen Zielkriterien ist. Jedes Entscheidungsproblem erfährt die gleiche „Gewichtung“. Eine schwere lebensbedrohliche Krankheit wird ebenso gewichtet wie eine chronische Krankheit, sogar im gleichen Indikationsgebiet werden austherapierte Patienten (Lebensqualität steht im Vordergrund) gleichgestellt mit Patienten deren Therapie auf die Heilung zielt. Dieses Vorgehen deckt sich weder mit dem allgemeinen Gerechtigkeitsempfinden noch mit den Wertmaßstäben der Bevölkerung, noch mit den Präferenzen der Patienten. (Es wird im Internet aber durch das IQWiG

darauf hingewiesen, dass die Präferenzen bei der Nutzenbewertung eine Rolle spielen und bei der Nutzenbewertung berücksichtigt werden sollten;

https://www.iqwig.de/de/methoden/grundsätze/hintergrund_was_sind_patientenpraferenzen.3757.html)

Im Hinblick auf diese Kritik antwortet das IQWiG in der Würdigung der Stellungnahmen: „Die im Entwurf dargestellte Operationalisierung ist in der Tat abstrakt und unabhängig von der Indikation.“ Das IQWiG muss eine „Feststellung zum Ausmaß treffen“, wobei dies mit „möglichst nachvollziehbaren Verfahren erfolgen sollte“. Ebenso wird gefordert, dass „Verfahren möglichst wenige Werteentscheidungen (etwa zur unterschiedlichen „Wertigkeit“ von Indikationen) trifft, die dem G-BA aufgrund seiner Legitimation vorbehalten bleiben sollten.“ (Würdigung der Stellungnahmen des Methodenpapiers 4.1, S. 11) Die Trennung zwischen der Informationsaufbereitung und der tatsächlichen Entscheidung, also der Trennung der Aufgaben des IQWiG und des GBA macht Sinn. Jedoch muss der GBA in seiner Funktion auch in der Lage sein, auf den durch das IQWiG berichteten Daten zu entscheiden. Keinen Sinn macht die Tatsache, dass der GBA durch den Bericht des IQWiG, keinerlei belastbare Informationen über die Werturteile der Versicherten bzw. Präferenzen der Patienten bekommt. In mehreren Antworten verweist das IQWiG darauf, dass keine konkreten „Verbesserungs- oder Alternativvorschläge (...) dem Institut (...) bekannt“ sind, oder durch die Stellungnehmenden im Stellungnahmeverfahren gemacht wurden (Würdigung der Stellungnahmen des Methodenpapiers 4.1, S. 11). Bei der Nachfrage hinsichtlich der Pilotprojekte AHP und CA antwortete Herr Andreas Gerber: „Wir haben uns bewusst dazu entschieden, darauf diesmal nicht einzugehen, weil, wie Sie alle wissen, ein Papier zu AHP ist erschienen, zur Conjoint-Analyse wird es kommen. Und wir denken, dass es Sinn macht. (...) Wir werden im Grunde genommen das dann in einer neuen Version unterbringen. (...) Die inhaltliche Diskussion sollte man vor dem Hintergrund der beiden Arbeitspapiere führen und dann, wenn es in die Methoden Einzug findet.“ (Dokumentation und Würdigung der Stellungnahmen zu den IQWiG Methoden 4.1, S.69 und 70) Die Pilotprojekte wurden bereits jeweils publiziert, eine Anpassung des Methodenberichtes an die Methoden der Präferenzmessung stehen noch aus.

Phasen der Nutzenbewertung

Zielsetzung einer Intervention ist die positive Beeinflussung des Patientennutzens. Zielsetzung der Nutzenbewertung ist die Operationalisierung des Patientennutzens, um eine rationale Wahlentscheidung über Interventionen zu ermöglichen. Das Verfahren der Nutzenbewertung sollte eine nachvollziehbare Bewertung durch eine transparente Arbeitsweise garantieren, die dem Verfahren zugrundeliegenden wissenschaftlichen Grundlagen sollten dokumentiert werden und darüber hinaus sollte sich das Verfahren an den gesetzlichen Grundlagen orientieren.

Grundsätzlich kann die Nutzenbewertung in drei Phasen unterteilt werden:

- (1) Das **Messen** von kausalen Effekten einer Intervention, wobei klinische und nicht-klinische Zielgrößen beim Patienten, einer Patientenpopulation oder den Bürgern mit einem dafür geeigneten Studiendesign gemessen werden. Der Nutzenbegriff des IQWiG in der engen Fassung, bezieht sich auf die kausal begründete Wirksamkeit einer Intervention. Zielsetzung ist die Darlegung eines kausalen Zusammenhangs zwischen Intervention und positivem oder negativem Effekt. Dieser Wirksamkeitsnachweis ist Gegenstand der evidenzbasierten Medizin. Der Nachweis kausaler Effekte sollte die Prinzipien der Strukturgleichheit (Randomisierung), Behandlungsgleichheit (Verblindung) und der Beobachtungsgleichheit (Doppel-Verblindung) berücksichtigen. Gesundheit ist ein multidimensionales Konstrukt und wird in klinischen und nicht-klinischen Effektparametern gemessen. Zielgrößen sind die Mortalität, die Morbidität und die Lebensqualität. Unter Nutzendimensionen versteht man die Verbesserung des Gesundheitszustandes, Verkürzung der Krankheitsdauer, Verlängerung der Lebensdauer, Verringerung der Nebenwirkung und die Verbesserung der Lebensqualität. Endpunkte sollten zuverlässig und korrekt Änderungen des Gesundheitszustandes abbilden. Unter anderem sind die folgenden Fragestellungen relevant: Definition der Zielkriterien, Festlegung der Messgrößen, Zuordnung der Effekte im Experiment, Bewertung der Unsicherheit und Interpretation der Ergebnisse bzw. Ergebnissicherheit.
- (2) Die **Bewertung** der gemessenen Effekte erfolgt im nächsten Schritt. Kausal begründete Effekte können einen Nutzen oder Schaden für den Patienten darstellen. Nutzen im ökonomischen Sinne resultiert aus der Bewertung der klinischen Effekte und erfolgt in der Regel aus der Perspektive des Konsumenten der Leistung. Der Nutzen entsteht beim Konsument und wird zumeist durch die Entscheidung über Alternativen durch den Konsumenten bestimmt. Wenn aufgrund von Informationsasymmetrien ein Individuum (eine Gruppe von Individuen) keine Entscheidung zwischen Handlungsalternativen für sich selbst treffen kann, können die Entscheidungsrechte an Dritte (eine Person oder Gruppe) übertragen werden, welche eine Wahlhandlung zwischen zwei Alternativen für den Konsumenten vornehmen. Ein Nutzen wird also aus der (erwarteten) Wirksamkeit abgeleitet, entspricht aber nicht den Wirksamkeitsparametern. Nutzen ist eine bewertende Aussage des Wissens über kausale Effekte einer Intervention. Bei der Ableitung des Nutzens aus den Effektgrößen sind deshalb mehrere Wertentscheidungen notwendig. Unklar ist, in welchem funktionalen Zusammenhang

ein Effekt einen Nutzen beim Patienten begründet. Werden mehrere Endpunkte für die Nutzenbewertung herangezogen, müssen Nutzen und Schaden abgewogen werden. Präferenzgewichte sollten zuverlässig und korrekt die relative Wichtigkeit der Endpunkte abbilden. Die relative Wichtigkeit eines Zielkriteriums, bedeutet, dass sich die Relevanz eines Endpunktes immer in Relation zu der Bedeutung alternativer Zielkriterien bemisst. Das Bewerten ist Gegenstand der Ökonomie und Statistik, insbesondere der Entscheidungstheorie. Zielsetzung ist die systematische Aggregation unterschiedlicher Zielerreichungsgrade, d.h. die Zusammenfassung aller Informationen über die Effekte in einem (eindimensionalen) Nutzenmaß. Ergebnis der Bewertung ist die Bereitstellung von Informationen, d.h. indikationsspezifischen oder indikationsübergreifenden Kennzahlen, für eine Entscheidung. Der rationale Vergleich von Interventionen erfolgt anhand dieses eindimensionalen Nutzenmaßes, als notwendige Bedingung für die Ermittlung des Zusatznutzens einer Intervention über die Vergleichsintervention. Nettonutzen ist das positive Ergebnis der vergleichenden Abwägung von Nutzen und Schaden einer Intervention hinsichtlich der Zielkriterien bei der Behandlung von Patienten bzw. Patientengruppen. Zusatznutzen ist das positive Ergebnis der vergleichenden Abwägung von Interventionen hinsichtlich der bewerteten Nettonutzen für Patienten bzw. Patientengruppen. Folgende Aspekte können dabei im Mittelpunkt stehen: Relevanz der Zielkriterien, Werturteile im Hinblick auf die Gewichtung und Vergleichbarkeit der Effekte sowie Annahmen über die Aggregation multipler Zielkriterien.

- (3) Die **Entscheidung** über die Intervention(en) wird auf Basis der Informationen über den Nutzen getroffen. Eine rational begründete Entscheidung über Interventionen basiert auf vorab definierten Zielen und vorhandenen Werturteilen/Wertmaßstäben. Die Entscheidungskriterien können abhängig sein, von den Zielen der Effektivität und der Wirtschaftlichkeit einer Intervention. Gegenstand der Entscheidung können die alternativen Interventionen in einem Indikationsgebiet sein (indikationsspezifische Entscheidung). Denkbar ist auch eine Entscheidung über alle Indikationsgebiete (indikationsübergreifende Entscheidung). Der Gegenstand der Entscheidung (z. B. Erstattungsfähigkeit oder Preisfestsetzung) konkretisiert die Anforderungen an das Skalenniveau (z. B. ordinale oder kardinale Skalen). Soll die Wirtschaftlichkeit über alle Handlungsalternativen garantiert werden, dann muss notwendigerweise ein indikationsübergreifendes Maß des Gesamtnutzens den Kosten gegenübergestellt werden. Die Ökonomie und Statistik befasst sich in der Entscheidungstheorie mit der Frage nach der optimalen Entscheidung. Die rational begründete Entscheidung orientiert sich neben den kausal begründeten Effekten (Zielerreichungsgraden) auch an den

Präferenzen der Betroffenen, als Ausdruck der Werte (Wertvorstellungen, Werturteilen oder Wertmaßstäben). Zielsetzung der Entscheidung ist die rationale Festlegung auf die optimale Handlungsalternative (oft unter Unsicherheit). Folgende Aspekte können dabei interessant sein: Konkretisierung der Entscheidungslogik, Auswahl der Wirtschaftlichkeitsprinzipien (Minimumprinzip und Maximumprinzip), Unterstützung /Konfliktlösung der Verhandlungspartner und Berücksichtigung ethischer und moralischer Rahmenbedingungen.

Dieser umfassende Prozess über alle drei Phasen kann als Verfahren der Nutzenbewertung bezeichnet werden. Aufbauend auf dieser Klassifizierung kann unterschieden werden in die „Nutzenbewertung im engeren Sinne“, die sich mit der eigentlichen Übertragung der klinischen Effekte in ein Nutzenmaß befasst, und einer „Nutzenbewertung im weiteren Sinne“, welche die Nutzenbewertung durch die Messung der Effekte und die Entscheidung anhand der Nutzenmaße als notwendige Bestandteile des Verfahrens komplementiert.

In Abschnitt 1.2 wird die evidenzbasierte Medizin als die Grundlage der Methoden erörtert. Der Beitrag der EbM liegt darin, dass sich die medizinische Behandlung von Patienten nicht „auf Meinungen und Übereinkünfte stützt“ (S. 4), sondern auf Basis von „Belege(n), die mit möglichst objektiven wissenschaftlichen Methoden erhoben wurden“ (S. 4). Aus der Perspektive der Haushalte und Unternehmen ist die Umsetzung dieser Zielsetzung wünschenswert. Eine transparente Nutzenbewertung und Entscheidung über die Erstattungsfähigkeit und Preisfestsetzung, die sich konsistent am Patientennutzen und der Wirtschaftlichkeit orientiert, maximiert die Wohlfahrt für die Bevölkerung und schafft Planungssicherheit für das innovative Gesundheitsunternehmen. Es bedarf geeigneter Methoden und Instrumente, die über das gesamte Bewertungsverfahren, die Evidenz über die Effekte und die Werte berücksichtigt, um so einer „subjektiv geprägten und damit häufig verzerrten Bewertung vorbeugen“ zu können (S. 4). Dazu gehören die Messung klinischer und nicht-klinischer Effekte, die Bewertung dieser gemessenen Zielkriterien (klinische oder nicht-klinische Endpunkte) sowie die Anwendung einer spezifischen Entscheidungslogik. Die dafür notwendigen Methoden und Kriterien müssen ferner den anerkannten internationalen Standards der EbM, aber auch der Gesundheitsökonomie entsprechen.

Das vorliegende Methodenpapier konzentriert sich sehr stark auf die Frage- und Problemstellungen des Messens und damit auf die evidenzbasierte Medizin, ohne in ausreichendem Maß die gesundheitsökonomischen Fragestellungen aufzugreifen. Die Instrumente der Entscheidungstheorie werden bei der Problemlösung nicht ausreichend berücksichtigt. Diese Stellungnahme konzentriert sich

auf die Frage- und Problemstellungen des Bewertens (Phase 2) und damit im Wesentlichen auf den gesundheitsökonomischen Beitrag für die Nutzenbewertung. Die offenen Frage- und Problemstellungen des Entscheidens (Entscheidungslogik der Effizienzgrenze) werden hier nicht vertieft.

Empfehlungen von Arbeitsschritten bei der Nutzenbewertung

In der evidenzbasierten Medizin gibt es die prominente Forderung nach der Berücksichtigung der Patientenpräferenzen, jedoch gibt es keine Anhaltspunkte wie dies geschehen soll. In der Ökonomie und Entscheidungstheorie gibt es ein reichhaltiges Spektrum an Methoden, Verfahren und Instrumenten. Folgende Fragestellungen und Arbeitsschritte sind notwendig, um ein Nutzenbewertungsverfahren an beiden Disziplinen auszurichten.

Messen:

- (1) Welche Bewertungskriterien werden bei der Bewertung der Handlungsalternativen zugrunde gelegt? Bewerten der Relevanz von Zielkriterien und Nutzendimensionen, d.h. konkret Identifikation der wichtigsten Bewertungskriterien (patientenrelevante Endpunkte).
- (2) Wie wird das Studiendesign bewertet? Wie wird die Qualität der Daten bewertet? Wie wird die Qualität der Ergebnisse bewertet? Bewerten der Ergebnissicherheit, d.h. Einordnung des Evidenzgrad, der Repräsentativität und der Effektschätzer bzw. Konfidenzintervalle.

Bewerten:

- (3) Welche Ober- und Untergrenzen werden für die Berechnung der Zielerreichungsgrade festgelegt? Festlegung der Annahmen über die Vergleichbarkeit, d.h. Konkretisierung der Realisationspotentials und Bewertung der Interaktionen zwischen den Bewertungskriterien (patientenrelevante Endpunkte).
- (4) Welche Wertmaßstäbe werden der Bevölkerung, den Versicherten und den Patienten unterstellt? Wer trifft die Wertentscheidungen im Bewertungsverfahren? Sind die Wertentscheidungen transparent? Welche Evidenz ist Grundlage dieser Wertentscheidungen?
- (5) Bewerten der Präferenzstruktur über das einzelne Bewertungskriterium, d.h. Gewichtung der Ausprägungen von Bewertungskriterien (patientenrelevante Endpunkte). Bewerten der Präferenzstruktur zwischen den Bewertungskriterien, d.h. Gewichtung der Bewertungskriterien (patientenrelevante Endpunkte). Wie werden die Bewertungskriterien bei der Bewertung der Handlungsalternativen gewichtet?

Entscheiden:

- (6) Welche Annahmen liegen der Aggregation kausaler Effekte zugrunde? Konkretisierung der Annahmen zur Operationalisierung bzw. Aggregation des Nutzens, d.h. Festlegung der Nutzenfunktion.
- (7) Welche Annahmen gelten für die Entscheidungsregel? Festlegung auf eine Entscheidungsregel, d.h. Identifikation einer Allokationslogik.

Anforderungen an das Skalenniveau

Die Entscheidung, basierend auf der frühen Nutzenbewertung über die Erstattungsfähigkeit bedarf einer unterschiedlichen Informationsqualität gegenüber der Entscheidung über die Preise im Kontext der Kosten-Nutzen Bewertung.

Nach § 5 Abs. 4 Satz 1 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) ist im Dossier und damit bei der Nutzenbewertung festzulegen, „in welchem Ausmaß ein Zusatznutzen vorliegt“. Die 6 Kategorien sind in § 5 Abs. 7 der AM-NutzenV vorgegeben: (1) erheblicher Zusatznutzen, (2) beträchtlicher Zusatznutzen, (3) geringer Zusatznutzen, (4) nicht quantifizierbarer Zusatznutzen, (5) kein Zusatznutzen belegt und (6) geringerer Zusatznutzen. Bei der Operationalisierung des Zusatznutzens geht das IQWiG davon aus, dass in der Arzneimittel-Nutzenbewertungsverordnung eine „Hierarchisierung dieser Zielgrößen intendiert“ ist (S. 206). Zudem kann davon ausgegangen werden, dass für die Bestimmung der 6 oben genannten Kategorien, der Zusatznutzen in einem eindimensionalen Nutzenmaß quantifiziert werden muss. (Ansonsten würde die Kategorie (4) keinen Sinn ergeben).

Geht man davon aus, dass sich die Teilnutzen unterschiedlicher patientenrelevanter Endpunkte auf ein Maß bzw. eine Dimension reduzieren lassen, dann muss festgelegt werden, welches Skalenniveau für die Entscheidung erforderlich ist. Man kann zwischen kardinalem und ordinalem Skalenniveau unterscheiden. Die kardinale Nutzenmessung misst der Nutzendifferenz von Intervention A und Intervention B eine Bedeutung zu. Die Nutzendifferenz zwischen zwei Interventionen hat in der ordinalen Nutzentheorie keine wesentliche Aussagekraft.

Bei der Entscheidung über die Erstattungsfähigkeit ist die Information über die Rangreihung der Interventionen ausreichend. (Eine Intervention A hat einen höheren Nutzen als eine Intervention B; der resultierende Zusatznutzen ist Grundlage der Entscheidung über den Festbetrag und die Erstattung.) Für diese Erstattungsentscheidung ist eine ordinale Nutzenfunktion für den Zusatznutzen ausreichend,

d.h. der Nutzen der Interventionen A und B muss lediglich in eine Rangreihe gebracht werden. Das ordinale Skalenniveau sagt über den Vergleich der Interventionen A und B nur aus, welche der Interventionen einen Zusatznutzen (höheren Nutzen) hat. Die ordinale Nutzenfunktion macht keine Aussagen darüber, „wie viel höher“ der Nutzen ist. Mit Hilfe der ordinalen Nutzenfunktion können die Präferenzen analytisch dargestellt werden, ohne dass die Nutzenvorstellungen der Patienten oder Experten quantifiziert werden müssen.

Bei der Entscheidung über die Festsetzung des Preises einer Gesundheitstechnologie ist ein kardinales Skalenniveau notwendig. Die kardinale Nutzenfunktion quantifiziert den Nutzen, d.h. der Nutzen einer Intervention kann als Zahl dargestellt werden. Die Nutzenmessung ist die Voraussetzung für eine rationale Entscheidung, welche die Wirtschaftlichkeit als Rationale in die Entscheidung integriert. Zielsetzung dieser Entscheidung ist die Nutzenmaximierung, d.h. die Optimierung der Gesundheitsversorgung unter Beachtung der Budgetrestriktion.

Zusammenfassung

Diese Stellungnahme befasst sich nicht mit den Methoden und Kriterien der Messung klinischer und nicht-klinischer Erfolgsparameter, der Begründung unterschiedlicher Studiendesigns, biometrischen Anforderungen oder der sachgerechten Umsetzung der EbM. Auch die theoretischen Grundlagen der im Methodenpapier vorgeschlagenen Entscheidungslogik sollten hier nicht vertieft werden – es wird nicht darauf eingegangen, ob die Effizienzgrenze als Entscheidungsregel geeignet ist oder nicht. Diese Stellungnahme konzentriert sich auf die Bewertung der gemessenen klinischen und nicht-klinischen Effekte. Die Stellungnahme erfolgt vor dem Hintergrund der in der Gesundheitsökonomie, noch spezifischer der Entscheidungstheorie, diskutierten Methoden und Instrumente.

Damit konzentrieren sich die Ausführungen auf die Bewertung der gemessenen Effekte als Kern der Nutzenbewertung. Es geht darum, wie die Multidimensionalität des Patientennutzens (Mortalität, Morbidität, Lebensqualität) bei der Entscheidung berücksichtigt werden kann. Zudem geht es darum, wie die patientenrelevanten Endpunkte, bei der Abwägung des Nutzen und Schadens, der Aggregation eines Maßes des Gesamtnutzens und bei der Beurteilung des Ausmaßes des Zusatznutzens in die Entscheidung eingehen.

Das zur Diskussion stehende Methodenpapier 4.2 (Version vom 18.06.2014) unterstellt, dass eine klinische Studie prospektiv die Wirkung und den Wert einer Intervention am Patienten bzw. Patientengruppen untersucht. Diese Annahme ist falsch, da die Konkretisierung des Nutzens einer Intervention weitere Werturteile in die Untersuchung einbeziehen muss.

Eine klinische Studie kann nur die Wirkung einer Intervention untersuchen bzw. Wissen über die Wirksamkeit generieren. Die Bewertung des Studiendesigns resultiert in einer Gewichtung der Ergebnisse im Hinblick auf unterschiedliche Evidenzgrade. Die Bewertung der Effektstärken resultiert ferner in einer Gewichtung der Zielkriterien im Hinblick auf die Ergebnisunsicherheit. Diese Bewertung des Studiendesigns oder der Studienergebnisse endet nicht mit dem Wissen über den Nutzen, sondern erweitert nur das Wissen über die Wirksamkeit. Eine Dokumentation des Zusatznutzens oder des Gesamtnutzens ist auf Basis der kausal begründeten Effekte, ohne die Berücksichtigung von Wertmaßstäben oder Präferenzen, nicht möglich.

Angenommen man könnte davon ausgehen, dass die Messung der patientenrelevanten Endpunkte zu 100% die Wirksamkeit einer Intervention A im Vergleich zu einer Intervention B wiedergeben. Dann kann auf Basis dieses Wissens keine Aussagen über den Nutzen, den Nettonutzen oder den Zusatznutzen gemacht werden. Insofern ist eine rationale Entscheidung nur auf Basis dieser klinischen Evidenz nicht möglich. Solange ein Nutzenmaß über die Erstattungsfähigkeit (Zusatznutzen) oder Preisfestsetzung (Gesamtnutzen in Relation zu den Kosten) informieren soll, gilt es, die gemessenen Effekte der patientenrelevanten Endpunkte in ein wissenschaftlich begründetes Nutzenmaß zu überführen. Um den Nutzen oder Wert einer Intervention zu analysieren, bedarf es mehrerer Wertentscheidungen. Ohne diese Wertentscheidungen zu treffen, ist es nicht möglich, einen Nutzenwert zu generieren oder eine systematische und transparente Entscheidung zu treffen.

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 Methodenpapiers. Offen bleibt, wer diese Wertentscheidung trifft. Offen bleibt auch, welche Evidenz als Basis dieser Wertentscheidungen herangezogen wird. Gegenwärtig werden diese Wertentscheidungen nicht transparent gemacht, ohne wissenschaftliche Evidenz nur auf Basis impliziten Wissens.

Wertentscheidungen können explizit oder implizit getroffen werden. Wertentscheidungen können auf Basis subjektiver oder objektiver Daten getroffen werden. Zunehmend wird die Forderung laut, dass 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.

Literatur / Quellen

Drummond MF, Sculpher MJ, Torrance GW, O'Brian BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.

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

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Hrsg.) (2013): Dokumentation und Würdigung der Stellungnahmen zur „Aktualisierung einiger Abschnitte der Allgemeinen Methoden Version 4.0 sowie neue Abschnitte zur Erstellung der Allgemeinen Methoden Version 4.1“ Version 1.0 vom 28.11.2013

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Hrsg.): Internet Glossar; URL: <https://www.iqwig.de/de/glossar.2727.html> (letzter Abruf: 06.08.2014)

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Hrsg.): „Hintergrund: Was sind Patientenpräferenzen?“ URL: https://www.iqwig.de/de/methoden/grundsatz/hintergrund_was_sind_patientenpraferenzen.3757.html; (letzter Abruf: 06.08.2014)

Saaty, T. L. (1990), How to make a decision: The analytic hierarchy process, in: European Journal of Operational Research, 48(1):9-26.

A.2.3 – Vach, Werner

Von: [Werner Vach](#)
An: [Methoden](#)
Thema: Stellungnahme zu "Allgemeine Methoden: Entwurf für Version 4.2"
Datum: Montag, 28. Juli 2014 21:49:22
Anlagen: KSS500_20140728_11590368_001.pdf
KSS500_20140728_11590456_002.pdf

Sehr geehrte Damen und Herren,

gerne möchte ich die Möglichkeit der Stellungnahme zu "Allgemeine Methoden: Entwurf für Version 4.2" nutzen.

Meine Anmerkungen beziehen sich ausschließlich auf den Abschnitt 3.5., und sie sollten als Anregungen verstanden werden.

1. Das IQWiG betont zu Recht das enge Zusammenspiel von Diagnose und Therapie, welches dem Nutzen diagnostischer Verfahren in vielen Fällen zugrunde liegt. Dies gilt, wie das IQWiG auch ausführt, gerade für randomisierte Studien, die auf eine direkte Messung des Nutzens durch Betrachtung patientenrelevanter Endpunkte abzielen. Allerdings spricht das IQWiG auch hier weiterhin von einer Nutzenbewertung des diagnostischen Verfahrens selber, was meine Erachtens etwas irreführend ist, weil in diesen Studien eben nicht diagnostische Verfahren, sondern eine Kombination von diagnostischen Verfahren mit bestimmten Therapieentscheidungen (die sehr genau oder sehr vage festgelegt sein könnten) untersucht wird.

Ich denke, dass eine Präzisierung dessen, was auf seinen Nutzen eigentlich bei der Einbeziehung randomisierter Studien untersucht wird, sehr hilfreich ist, um zu vermeiden, dass die Ergebnisse von Nutzenbewertungen diagnostischer Maßnahmen falsch interpretiert werden und er zu unerwünschten

Konsequenzen für das deutsche Gesundheitswesen kommt. Da in diesen Studien Kombinationen von diagnostischen Verfahren und Therapieentscheidungen untersucht werden, sollte eine Nutzenbewertung darauf abzielen, eine Entscheidung vorzubereiten, eben diese Kombination, und nicht das Verfahren alleine, in den Leistungskatalog der GKV aufzunehmen.

Das heißt, nur die diagnostische Maßnahme inklusive der Therapieentscheidung nach Kriterien, die denen der randomisierten Studien ähnlich sind, ist aufzunehmen. Wenn die Therapieentscheidungen in der Praxis in Deutschland anders verlaufen als in den randomisierten Studien, besteht die Gefahr, dass die Aufnahme des diagnostischen Verfahrens in den Leistungskatalog nicht zur erwarteten Verbesserung der Versorgung führt.

Hinzu kommt, dass in randomisierten Studien häufig ein neues diagnostisches Verfahren mit einem bisherigen, im Leistungskatalog bereits aufgenommen Verfahren, verglichen wird, und in den Studien in jedem Patienten in der Regel nur eines der beiden Verfahren angewandt wird. Führt nun eine Nutzenbewertung des neuen Verfahrens letztendlich zur Aufnahme in den Leistungskatalog, so besteht die Gefahr, dass dann nicht nur das neue, sondern zusätzlich das alte Verfahren in vielen Patienten angewandt wird. Für diese Praxis liegt aber keine Nutzenbewertung vor. Es ist daher sicherzustellen, dass allen Beteiligten klar ist, dass aus diesen Studien eine Nutzenbewertung hinsichtlich des Einsatzes des neuen Verfahrens (in Kombination mit einer Therapieentscheidung) ohne gleichzeitige Verwendung des bisherigen Standardverfahrens erfolgt.

Aus diesen Gründen würde ich die oben genannte Präzisierung anregen.

2. Das IQWiG weist zu Recht auf die Unterschiede zwischen verschiedenen Designs für randomisierte Studien zur Nutzenbewertung diagnostischer Verfahren

hin. Es erscheint mir aber wichtig, schon im Methodenpapier auf einen grundlegenden Unterschied zwischen Strategie-/Diskordanzdesigns auf der einen, und Interaktionsdesigns und Anreicherungsdesigns auf der anderen Seite, hinzuweisen.

Bei Strategie-/Diskordanzdesigns werden zwei verschiedene diagnostische Verfahren miteinander verglichen. Die daraus folgende Therapieentscheidung ist in der Regel an sich etabliert (z.B. unterschiedliche Therapien nach Stadium der Erkrankung), und es geht um die Frage, welches Verfahren besser geeignet ist, die Entscheidung zu unterstützen. Studien mit diesen Designs werden daher in der Regel dazu benutzt werden können, den Zusatznutzen eines neuen diagnostischen Verfahrens gegenüber einer etablierten Standarddiagnostik zu bestimmen.

In Interaktionsdesigns und Anreicherungsdesigns kommt hingegen nur ein diagnostisches Verfahren zum Einsatz. Sie sind daher nicht geeignet, einen Zusatznutzen gegenüber einer etablierten Standarddiagnostik nachzuweisen. Diese Designs werden typischerweise auch für einen anderen Zweck benutzt. Hier geht es in der Regel um eine neue Diagnostik und eine neuartige Therapieentscheidung, häufig auch um eine neue Therapie. Aus der Sicht der Nutzenbewertung vergleichen diese Studien also implizit den bisherigen Standard (d.h. keine Diagnostik und die gleiche Therapie für alle Patienten)

mit einem neuen Ansatz, bestehend aus einer Diagnostik für alle Patienten und einer alternativen Therapie für einen Teil der Patienten. In dieser Situation - die eben typisch ist für neuartige diagnostische Verfahren, die keine etablierten Verfahren ersetzen sollen - können diese Studien ihren Beitrag zur Nutzenbewertung leisten.

Ich denke, dass diese Unterscheidung und Einordnung nützlich ist, um Missverständnisse zu vermeiden. Siehe hierzu auch die nächsten beiden Punkte.

3. Anreicherungsdesigns und Interaktionsdesigns unterscheiden sich dadurch, dass in Anreicherungsdesigns eine Entscheidung, in welcher Untergruppe die neue Therapieoption nicht erprobt werden soll, vorab gefällt wird, während in Interaktionsdesigns die neue Therapie in allen Patienten erprobt wird. Dadurch haben Interaktionsdesigns die zusätzliche Eigenschaft, dass sie möglicherweise zu der Erkenntnis führen können, dass die Patienten in ihrer Gesamtheit von der neuen Therapie profitieren, und somit das diagnostische Verfahren gar nicht notwendig ist. Dies ist in einem Anreicherungsdesign nicht möglich.

Es ist für mich aber nicht nachvollziehbar, dass deshalb Studien mit Anreicherungsdesign keine tragfähigen Schlüsse hinsichtlich des Nutzens des diagnostischen Verfahrens (in Kombination mit der Therapieentscheidung) erlauben. Wie oben ausgeführt, beruht eine Nutzenbewertung aufgrund eines Interaktionsdesigns oder Anreicherungsdesigns auf dem impliziten Vergleich mit der Situation, dass alle Patienten die Standardtherapie bekommen. Und in beiden Fällen kann man zu dem Schluss kommen, dass es eine durch das diagnostische Verfahren bestimmbare Untergruppe gibt, die von einem Therapiewechsel profitiert. Und in beiden Fällen ergibt sich somit ein Nutzen für das Gesamtkollektiv, wenn im Rest der Patienten an der Standardtherapie festgehalten wird. Daher gibt es aus der Sicht der Nutzenbewertung keinen Unterschied zwischen den Designs.

Es ist sachlich richtig, dass bei einer Nutzenbewertung aufgrund eines Anreicherungsdesigns keine Information darüber vorliegt, ob die neue Therapieoption nicht auch im Rest der Patienten einen Vorteil bietet. Und es kann daher passieren, dass man zu einem späteren Zeitpunkt feststellt, dass man auf das diagnostische Verfahren verzichten kann, weil die neue Therapieoption für alle Patienten von Vorteil ist. Und dass daher das neue Verfahren eigentlich überflüssig und daher "nutzlos" ist. Aber es ist ja Ziel einer Nutzenbewertung zu untersuchen, ob ein neues Verfahren im Vergleich zur bisherigen Praxis einen Vorteil für die Patienten bedeutet, und nicht im Vergleich zu einer möglichen, zukünftigen Praxis, zu der es eben noch keine Studien gibt.

4. Das IQWiG weist zurecht darauf hin, dass die statistischen Analyse von Studien mit Interaktionsdesigns einer sorgfältigen, prospektiven Planung bedarf, um die Aussagekraft einordnen zu können. Ich denke, es wäre aber - aus mindestens zwei Gründen - wichtig, wenn das IQWiG auch dazu Stellung nimmt, welche Form der Analyse nach seiner Ansicht für eine Nutzenbewertung angemessen wäre.

a) Eine methodische Herausforderung bei der Einbeziehung von Studien mit Interaktionsdesigns in eine Nutzenbewertung ist die Notwendigkeit des Nachweises der Existenz einer Subgruppe, die von der neuen Therapie profitiert, als auch des Nachweises eine Subgruppe, bei der ein Wechsel nicht empfohlen werden kann. Der erste Schritt ist noch vergleichsweise einfach, aber der zweite erfordert eine Festlegung, wie man das Problem des Nachweises von "keinem Unterschied" angehen möchte.

b) Es gibt viele Studien, bei denen alle Patienten randomisiert werden, und ein diagnostisches Verfahren (Biomarker) in allen Patienten angewandt wird, die aber nicht im Sinne eines Interaktionsdesigns ausgewertet werden.

In einer Übersichtsarbeit von Freidlin et al (Journal of Clinical Oncology 31, 3158-3161)

aus dem letzten Jahr werden sechs übliche Auswertungsstrategien betrachtet, von denen nur eine standardmäßig die Analyse sowohl der markerpositiven als auch der markernegativen Patienten vorsieht. Es ist daher damit zu rechnen, dass häufig Studien mit Interaktionsdesign vorliegen, ihre Analyse aber nicht direkt den Anforderungen einer Nutzenbewertung genügt.

Ich hoffe, dass diese Anregungen auf fruchtbaren Boden fallen. Für Rückfragen stehe ich gerne zur Verfügung.

MfG

Werner Vach

--

Werner Vach, professor
Clinical Epidemiology Group
Center for Medical Biometry and Medical Informatics
Medical Center - University of Freiburg
Stefan-Meier-Str. 26, D-79104 Freiburg, Germany

