



Stellungnahme der Bundesärztekammer

zu den Allgemeinen Methoden des IQWiG Entwurfsversion 4.0 in der Fassung vom 09.03.2011

Berlin, 08.04.2011

Bundesärztekammer
Herbert-Lewin-Platz 1
10623 Berlin

Zu den Allgemeinen Methoden des IQWiG in der Entwurfsversion 4.0 in der Fassung vom 09.03.2011 nimmt die Bundesärztekammer wie folgt Stellung:

Die Bundesärztekammer begrüßt die regelmäßige Überprüfung und Überarbeitung der methodischen Grundlagen des IQWiG. Die Bundesärztekammer hatte hierzu bereits mehrfach in Stellungnahmen Hinweise gegeben. Zu der jetzt vorgelegten Entwurfsversion der „Allgemeinen Methoden“ zur Nutzenbewertung in der Version 4.0 nimmt die Bundesärztekammer wie folgt Stellung:

Präambel

In der überarbeiteten Präambel werden jetzt auch Gremien und Organe des IQWiG vorgestellt (Stiftungsrat und Stiftungsvorstand). Hierzu ist anzumerken, dass

- ▶ Strukturen des Instituts nicht in der Präambel abgehandelt werden sollten, sondern, wie in Version 3.0, in dem dafür geeigneteren Kapitel 1,
- ▶ als weitere zu nennende Gremien der Beirat und das Kuratorium des Instituts fehlen,
- ▶ die Zusammensetzung des Stiftungsrats auf Seiten der „Leistungserbringer“ die Bundesärztekammer vermissen lässt.

Löschen des ehemaligen Abschnitts 1.2 „Struktur“

Die Erläuterung der Struktur des Instituts mag nicht unmittelbar Bestandteil seiner Methoden sein, die Kenntnisse der Struktur können jedoch das Verständnis für die Arbeitsweise erleichtern. Insofern gehen durch die Streichung dieses Abschnitts möglicherweise nützliche Informationen verloren. Es sollte daher wenigstens ein Hinweis in dem Text eingefügt werden, dass Informationen zu Struktur und Organisation des Instituts auf der Internetseite des IQWiG verfügbar sind.

Überarbeiteter Abschnitt 1.2.6

Im Zusammenhang mit der Erläuterung der Eigenschaften klinischer Studien werden die Begriffe „Verzerrung“ und „Concealment“ benutzt, dabei wird aber nur der Begriff der „Verzerrung“ erläutert. Auf den Klammerzusatz mit „Concealment“ sollte entweder an dieser Stelle verzichtet werden, oder es sollte ein gleichfalls deutschsprachiger Begriff mit Erläuterungen ergänzt werden (Verblindung, verdeckte Zuordnung, ...).

Unter Nr. 1 ist vom **Nachweis** von Kausalität die Rede, unter Nr. 2 vom **Beweis** der Kausalität - hier sollte eine einheitliche Begriffsverwendung angestrebt werden, vorzugsweise zugunsten des Begriffs Nachweis, da eine Beweisführung im mathematischen oder naturwissenschaftlich-physikalischen Sinne hier wohl nicht gemeint ist.

Hauptanliegen des Kapitels ist nach wie vor, den hohen Stellenwert von RCTs für die Beurteilung des Nutzens einer Intervention oder Maßnahme darzulegen. Dabei wird erläutert, warum sich hohe Ergebnissicherheit und Alltagsnähe nicht ausschließen müssen, um offenbar wiederkehrenden Forderungen nach Berücksichtigung von Studien auch ohne Randomisierung entgegenzutreten. Zu vermissen ist bei diesen Ausführungen je-

doch die Würdigung des Umstands, dass für viele Fragestellungen gar keine RCTs zur Verfügung stehen, zumindest nicht in solchen Zeiträumen, wie sie der G-BA zur Beantwortung von Fragen zu Leistungsausschluss oder -gewährung innerhalb der GKV benötigt und zur Klärung als Auftrag an das Institut vergibt. Insofern gehen die ausgiebigen Ausführungen nicht auf alle der in den letzten Jahren zu beobachtenden Probleme der Entscheidungsfindung bei der Nutzenbewertung ein.

Es sollte außerdem bedacht werden, dass der neu eingefügte Abschnitt 7.3.4 (Bewertung subjektiver Endpunkte bei offenen Studiendesigns) zwar nicht unmittelbar im Widerspruch zur herausgehobenen Bedeutung von RCTs gesehen werden muss, diese jedoch relativiert. Damit bleibt die Frage einer verlässlichen und reproduzierbaren Bewertungsgrundlage bzw. ein möglicher ad-hoc-Umgang des Instituts in Bewertungssituationen vakant.

Überarbeiteter Abschnitt 2.1

Abbildung 1 „Ablauf der Berichterstellung“ entspricht nicht mehr in allen Punkten den Erläuterungen im Text. So fehlen die (in der entsprechenden Abbildung der Version 3.0 noch vorhandenen) Textfelder, wonach der vorläufige Berichtplan und der Vorbericht beim Auftraggeber/Kuratorium/Stiftungsvorstand vorgelegt werden.

Die Auswahl externer Sachverständiger mithilfe einer Sachverständigendatenbank erscheint als pragmatischer Weg zur Bewältigung der durch kurze Fristen gekennzeichneten Nutzenbewertungen nach § 35a SGB V. Die Auswahl von Sachverständigen soll anhand einer Kriterienliste geschehen, die auf Selbstauskünften der Bewerber basiert. Ein Verifizierungsverfahren dieser Angaben durch das Institut ist nicht regelhaft vorgesehen - ob hierzu eine Notwendigkeit besteht, wird sich vermutlich erst nach ersten Erfahrungen mit dem Verfahren erweisen. Es sollte überlegt werden, die ausführlichere Beschreibung des Prozederes, die auf den Internetseiten des Instituts verfügbar ist, auch in den „Allgemeinen Methoden“ darzustellen.

Überarbeiteter Abschnitt 3.1.1

Die stärkere Bezugnahme auf den Aspekt der klinischen Wirksamkeit bei der Definition des patientenrelevanten medizinischen Nutzens von Behandlungsmaßnahmen ist zu begrüßen (die Bundesärztekammer hatte dies in ihrer Stellungnahme vom 13.12.2007 zur Version 3.0 des Methodenpapiers angeregt). Die in diesem Zusammenhang zitierte und als „höchstrichterliche Rechtsprechung“ zusammengefasste Entscheidung des Bundessozialgerichts vom 6. Mai 2009 (Protonentherapie bei Brustkrebs), dass „bestimmte (Nutzen-)Aspekte erst dann notwendigerweise zu bewerten sind, wenn die therapeutische Wirksamkeit hinreichend belegt ist“, eignet sich allerdings nur bedingt als Referenz für einen differenzierten Umgang mit diesen Begriffen bzw. der Problematik. Dies ergibt sich allein schon aus einzelnen Formulierungen der BSG-Entscheidung, die eine Unschärfe bei der Trennung zwischen Nutzen und Wirksamkeit erkennen lassen: „Danach ist eine Methode nur dann entsprechend dem allgemein anerkannten Stand der medizinischen Erkenntnisse erforderlich, wenn ihr **Nutzen** und die medizinische Notwendigkeit

der Methode belegt sind. Fehlen solche Belege für die therapeutische **Wirksamkeit** eines Behandlungskonzepts, so ...“ (BSG v. 6.5.2009, B 6 A 1/08 R).

Die Möglichkeit, einen „endpunktbezogen »geringeren Schaden« (im Sinne einer Verringerung von Nebenwirkungen) bei Betrachtung der Effekte auf alle anderen Endpunkte in die abwägende Feststellung eines »Zusatznutzens« einzubeziehen, ist zu begrüßen. Dieses Vorgehen sollte nicht nur als „möglich“ im Sinne einer Option dargestellt werden, sondern regelmäßig Anwendung finden.

Überarbeiteter Abschnitt 3.1.2

Die stärkere Ausdifferenzierung der Bewertung des Umgangs mit Surrogatparametern ist zu begrüßen. Die in der letzten Stellungnahme der Bundesärztekammer vom 13.12.2007 vorgeschlagene Differenzierung zwischen intermediären und Proxy-Zielgrößen ist an dieser Stelle jedoch zu wiederholen.

Überarbeiteter Abschnitt 3.1.3

In diesem nur geringfügig geänderten Abschnitt ist aus der Gegenüberstellung von Schaden und Nutzen medizinischer Interventionen nach wie vor die Tendenz erkennbar, in Zweifelsfällen eher ablehnenden bzw. dilatorischen Entscheidungen den Weg zu bahnen (vergleiche die Stellungnahme der Bundesärztekammer vom 13.12.2007).

Überarbeiteter Abschnitt 3.1.4

Trotz der passagenweisen Überarbeitung dieses Abschnitts bleibt im Kern als Regelanforderung der Beleg eines statistisch signifikanten Effekts durch eine Metaanalyse von Studien mit endpunktbezogen geringer Ergebnisunsicherheit oder durch mindestens zwei voneinander unabhängig durchgeführte Studien mit endpunktbezogen geringer Ergebnisunsicherheit bestehen. Die Bundesärztekammer hatte dies bereits in ihrer Stellungnahme vom 13.12.2007 als zu rigide eingestuft, zumindest bezogen auf lebensbedrohliche Erkrankungen. Positiv ist anzumerken, dass nunmehr für die Entscheidungsperspektive des Instituts auf das Prinzip der Risikovorsorge unter Annahme eines Schadenpotenzials verzichtet werden soll.

Neuer Abschnitt 3.3.3

In diesem Abschnitt werden überwiegend die gesetzlichen Vorgaben der Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V wiedergegeben. Mit Blick auf die Nutzenbewertung wird auf die „im vorliegenden Methodenpapier beschriebenen Standards der evidenzbasierten Medizin“ verwiesen; die Bewertung der Kosten soll „auf Basis der Standards der Gesundheitsökonomie“ erfolgen. Zumindest der letztgenannte Aspekt erscheint sehr pauschal, da die Gesundheitsökonomie, wie jeder andere wissenschaftliche Fachbereich auch, von unterschiedlichen Lehrmeinungen und entsprechenden Diskursen geprägt ist. Wünschenswert an dieser Stelle wäre der Verweis auf die institutseigenen Me-

thoden zur Bewertung von Kosten-Nutzen-Verhältnissen, die hierzu allerdings einer Aktualisierung bedürften.

Abschnitt 3.5

Unverändert sollen für die Bewertung diagnostischer Verfahren „dieselben Nutzenkategorien wie bei der Bewertung therapeutischer Interventionen“ gelten, nämlich „Mortalität, Morbidität und gesundheitsbezogene Lebensqualität“. Dies erscheint nach wie vor problematisch, da einerseits die Studiensituation bei diagnostischen Verfahren in aller Regel ungünstiger ist als bei therapeutischen Verfahren, andererseits die ärztliche Perspektive, die für Therapieentscheidungen maßgeblich ist, ausgeblendet wird. Komplett negiert wird der Selbstwert einer diagnostischen Erkenntnis für Arzt und Patient, die unabhängig von der Zielstellung einer therapeutischen Konsequenz existiert.

Überarbeiteter Abschnitt 4

Der neue Titel dieses Abschnitts „Methoden der Versorgungsqualität“ ist, im Gegensatz zum früheren Titel „Leitlinien und Disease-Management-Programme“, nicht selbsterklärend. Es ist anhand des Titels nicht erkennbar, dass es in diesem Abschnitt primär um das Thema Leitlinien geht.

Der Versuch einer eigenen Definition von Leitlinien zu Beginn des Abschnitts wirkt angesichts mittlerweile seit vielen Jahren auf nationaler und europäischer Ebene vorliegender und weithin akzeptierter Formulierungen (siehe etwa unter auf den Webseiten des ÄZQ unter <http://www.leitlinien.de>) redundant.

Die Aussage, dass das „Ressort Versorgungsqualität des IQWiG seinen Aufgabenbereich in der wissenschaftlichen Identifikation und Beschreibung von Versorgungsstandards auf Basis hochwertiger Leitlinien sieht“, gehört eher in einen Geschäftsbericht, weniger in ein Methodenpapier. Die Aufgabe des IQWiG sollte sich aus seiner gesetzlichen Grundlage (vgl. § 139a SGB V) ergeben und nicht durch die Sichtweise einzelner Ressorts bestimmt werden. Mit Blick auf Leitlinien lautet hier der klare gesetzliche Auftrag „Bewertungen evidenzbasierter Leitlinien für die epidemiologisch wichtigsten Krankheiten“ sowie „Abgabe von Empfehlungen zu Disease-Management-Programmen“ (§ 139a SGB V Abs. 3 Nr. 3 SGB V). Die in Aussicht gestellte Identifikation von Versorgungsstandards auf der Basis hochwertiger Leitlinien ist etwas völlig anderes, indem Leitlinien nur noch Mittel zum Zweck sein sollen und nicht mehr der eigentliche Gegenstand der Befassung. Die Aufgabe, die sich das Ressort Versorgungsqualität des IQWiG hier selber erteilt, erinnert an im Zuge des GKV-Gesundheitsreformgesetzes 2000 in einem neuen § 137e Abs. 3 SGB V formulierte Aufgabe für den damaligen Koordinierungsausschuss, „auf der Grundlage evidenzbasierter Leitlinien die Kriterien für eine im Hinblick auf das diagnostische und therapeutische Ziel ausgerichtete zweckmäßige und wirtschaftliche Leistungserbringung für mindestens 10 Krankheiten je Jahr [zu] beschließen, bei denen Hinweise auf unzureichende, fehlerhafte oder übermäßige Versorgung bestehen und deren Beseitigung die Morbidität und Mortalität der Bevölkerung nachhaltig beeinflussen kann“. Diese, dem Vorläufergremium des G-BA zugeordnete Aufgabe verschwand schon nach kurzer Zeit wieder durch das GKV-Modernisierungsgesetz. Die

Gründe für seinen raschen Sinneswandel sind seitens des Gesetzgebers nicht dokumentiert worden, es liegt aber die Vermutung nahe, dass nicht nur die Quantität der Aufgabe (10 Krankheiten pro Jahr) an methodische Grenzen stieß.

Von dieser grundsätzlichen Überlegung abgesehen bedeutet die beabsichtigte Verknüpfung des Untersuchungsgebiets an das Vorhandensein „hochwertiger Leitlinien“ eine Limitierung wider besseres Wissen. Auch hier hatte sich bekanntlich der Gesetzgeber in einer ähnlich gelagerten (Fehl-)Vorstellung selber korrigieren müssen, indem er - ebenfalls per GKV-Modernisierungsgesetz - seinen Anspruch revidieren musste, Patienten in strukturierten Behandlungsprogrammen nach „evidenzbasierten Leitlinien“ behandeln lassen zu wollen (§ 137f Abs. 2 Nr. 1 SGB V alt, Stand 2002-2004), anstatt wie seither und bis heute gültig „nach dem aktuellen Stand der medizinischen Wissenschaft unter Berücksichtigung von evidenzbasierten Leitlinien oder nach der jeweils besten, verfügbaren Evidenz...“

Dass „hohe methodische Qualität von Leitlinien nicht notwendigerweise mit der inhaltlichen Qualität der darin enthaltenen Empfehlungen korreliert“ (siehe Abschnitt 4.3.1), ist zutreffend, die Schlussfolgerung, die Parameter daher als vollständig voneinander entkoppelt betrachten zu müssen und komplett neue Methoden zur Bewertung von Leitlinieninhalten zu entwickeln, erscheint aber ambitioniert. Unter praktischen Aspekten erscheint es fraglich, mit welchen Ressourcen das Institut die inhaltliche Qualität der Empfehlungen prüfen möchte. Nicht umsonst ist in der Version 3.0 der IQWiG-Methoden festgestellt worden: „Die inhaltliche Prüfung umfasst neben der Bewertung von Vollständigkeit und Aktualität der einbezogenen Literatur auch die Interpretation und Bewertung der Studienergebnisse. Da dieses Verfahren sehr aufwendig ist, muss sich die inhaltliche Prüfung auf die in Auftrag gegebenen Fragestellungen des G-BA bzw. auf die wesentlichen Kernempfehlungen der Leitlinie beschränken.“

Insgesamt bietet der neu konzipierte Abschnitt 4 zwar eine Reihe wertvoller Überlegungen zum anspruchsvollen Thema „Critical Appraisal“ (bezeichnenderweise taucht dieser Begriff - außer im Literaturverzeichnis - im Methodenpapier gar nicht auf), hat sich aber vom eigentlichen gesetzlichen Auftrag des Instituts für den Bereich Leitlinien entfernt. Das Thema „Abgabe von Empfehlungen zu Disease-Management-Programmen“ ist bei der Neukonzeption des Abschnitts ebenfalls aus dem Blickfeld geraten.

Überarbeiteter Abschnitt 5.2.1

Es sei darauf hingewiesen, dass die kritische Einschätzung des Instituts, wonach in den Gesundheitsinformationen für Bürger und Patienten „die durchgehende Verwendung maskuliner Personenbezeichnungen (generisches Maskulinum) zu einer gedanklichen Unterrepräsentation von Frauen führt, worin eine sprachliche Benachteiligung von Frauen zu sehen ist“, offenbar nicht für die Zielgruppe seiner Methodentexte gilt (Auszug aus der Präambel des vorliegenden Methoden-Entwurfs 4.0: „In diesem Dokument wird bei der Angabe von Personenbezeichnungen jeweils die männliche Form angewandt. Dies erfolgt mit dem Ziel einer besseren Lesbarkeit“).

Überarbeiteter Abschnitt 5.3.4

Die stärker differenzierte Darstellung des Aktualisierungsprozesses der vom Institut verfassten Gesundheitsinformationen wird begrüßt, insbesondere der Verzicht auf den in der Version 3.0 noch verwendeten groben Hinweis, wonach „sich das Institut überdies über methodologische Entwicklungen auf dem Laufenden“ halte, „um entscheiden zu können, wann die Evidenz für bestimmte Resultate so überzeugend geworden ist, dass es keiner weiteren Aktualisierungen“ bedürfe.

Überarbeiteter Abschnitt 5.5.3

Als Beitrag zur Evaluation der Gesundheitsinformationen des Instituts wird auf einen von der Universität Bielefeld zur Evaluierung von Online-Modellprojekten entwickelten Online-Fragebogen verwiesen. Unverändert gegenüber den Methoden 3.0 ist von der „Auswertung der ersten 2561 ausgefüllten Fragebögen“ die Rede. Die Formulierung „ersten“ erweckt den Eindruck, als wollte das Institut mit den Auswertungen weiterer Fragebögen fortfahren, dies ist aber offenbar nicht gemeint oder zwischenzeitlich nicht der Fall gewesen.

Des weiteren erscheint auch die Aussage, wonach es bisher kein Instrument gäbe, „das sich als zuverlässiger Indikator für die Qualität von Gesundheitsinformationen oder -websites erwiesen hat“, insofern überprüfungswürdig, als die zur Unterstützung dieser Feststellung zitierte Literatur aus den Jahren 1998, 2000, 2002 und 2004 stammt und damit möglicherweise nicht mehr den neuesten Stand auf diesem sehr dynamischen Gebiet berücksichtigt. Verwiesen sei in diesem Zusammenhang auf das gemeinsame Portal Patienten-Information.de von Bundesärztekammer und Kassenärztlicher Bundesvereinigung, bei dem unter dem Stichwort „Linkpolicy“ Kriterien aufgelistet werden, welche zur Bewertung der Qualität von Gesundheitsinformationen bzw. diesbezüglichen Internetangeboten herangezogen werden können.

Überarbeiteter Abschnitt 7.1.4

In diesem Abschnitt wird an mehreren Stellen die Bezeichnung „Behandler“ verwendet, wo „Arzt/Ärztin“ sinngemäß zutreffend wäre, siehe z. B. „Weiterhin ist es erforderlich, dass die Zielvariable unabhängig vom (unverblindeten) Behandler ist bzw. unabhängig vom Behandler verblindet erhoben wird ...“. Der Begriff „Behandler“ anstelle von Arzt/Ärztin ist durch seine Verwendung im Dritten Reich als Diskriminierung jüdischer Ärzte belastet und sollte dementsprechend nicht mehr verwendet werden. Im September 1938 war es zum Entzug der Approbationen für jüdische Ärzte und einer anschließenden Bezeichnung als „Behandler“ gekommen. Vertiefende Informationen zu diesem Thema sind z. B. im Internet unter <http://www.jahrestag-approbationsentzug.de/> zu finden.

Neuer Abschnitt 7.3.4

Infolge der in diesem Abschnitt verwendeten Definition von subjektiven Endpunkten, darunter der Erhebung und Einschätzung „von den behandelnden bzw. endpunkterhebenden Personen“, die auf eine Stufe gestellt werden soll mit „patientenberichteten Endpunkten“, wird ein beträchtlicher Anteil von RCTs (und darauf fußender Metaanalysen) in den

Verdacht „hochgradiger“ Verzerrung geraten. Die prospektive, projektspezifische Festlegung adjustierter Entscheidungsgrenzen als methodischer Lösungsansatz des Instituts dürfte dabei großen Spielraum für die Bewertung bieten. Hierin könnte ein Risiko für eine zu variable Berücksichtigung von Studien durch das Institut liegen. Unklar erscheint, auf welche empirischen Daten das Institut sich für die Wahl eines Grenzwertes projektspezifisch stützen möchte; der Verweis auf Daten, „wie sie z. B. Wood et al. liefern“ [die zentrale Referenzliteraturstelle für den in diesem Abschnitt thematisierten Zweifel an der Qualität von RCTs], beschreibt dies nur vage.

Es sollte außerdem aus systematischen Gründen überlegt werden, die Ausführungen zum Umgang mit Verzerrung in den Abschnitt 7.3.11 (Darstellung von Verzerrungsarten) einzugliedern.

Fazit

Die Bundesärztekammer erkennt weitere Verbesserungen in den Allgemeinen Methoden im Vergleich zu den Vorversionen. Dies betrifft etwa

- die stärkere Ausdifferenzierung der Bewertung des Umgangs mit Surrogatparametern,
- die stärkere Bezugnahme auf den Aspekt der klinischen Wirksamkeit bei der Definition des patientenrelevanten medizinischen Nutzens von Behandlungsmaßnahmen,
- die Absicht, die Verringerung von Nebenwirkungen stärker in die abwägende Feststellung eines Zusatznutzens einer therapeutischen Methode einzubeziehen,
- die differenziertere Darstellung des Aktualisierungsprozesses der vom Institut verfassten Gesundheitsinformationen.

Andere Inhalte der Methoden erscheinen nach wie vor problematisch, etwa

- das Festhalten an vergleichsweise rigiden Bewertungsmechanismen unter Verwendung von RCTs und Metaanalysen trotz erfahrungsgemäß unbefriedigenden Studiensituationen. Hinsichtlich der Beurteilung zur Ergebnissicherheit und der klinischen Relevanz von Studienergebnissen auf Basis der gegenwärtig besten Evidenz sollten unter Würdigung der jeweiligen Zielgrößen prinzipiell Studien aller Evidenzgrade beim „Assessment“ berücksichtigt und dargestellt werden, um eine möglichst vollständige Information des Auftraggebers G-BA zu gewährleisten und diesem ein möglichst umfassendes „Appraisal“ zu ermöglichen.
- die zu wenig differenzierte Übertragung der Ansprüche der Bewertung des Nutzens therapeutischer auf diagnostische Maßnahmen,
- der Gegenüberstellung von Schaden und Nutzen medizinischer Interventionen mit der Tendenz, im Zweifel eher ablehnenden oder dilatorischen Entscheidungen den Weg zu bahnen,
- das neu formulierte Kapitel zu Leitlinien weist viele interessante Überlegungen zur inhaltlichen Bewertung von Leitlinien auf, hat sich dabei aber vom gesetzlichen Auftrag entfernt,

- eine noch geringe und damit erweiterungsfähige Detailtiefe weisen die im Zuge des Arzneimittelmarkt-Neuordnungsgesetzes (AMNOG) neu formulierten Abschnitte zur Nutzenbewertung von Arzneimitteln nach § 35a SGB V auf. Die Bundesärztekammer geht von einer zeitnahen ausführlichen methodischen Auseinandersetzung des IQWiG auch mit Fragen gesundheitsökonomischer Analysen und Kosten-Nutzen-Bewertungen aus.

Berlin, 08.04.2011

gez.
Dr. rer. nat. Ulrich Zorn, MPH
Bereichsleiter im Dezernat 3