

Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden 7.0

Version 1.0
vom 19.09.2023

Anschrift des Herausgebers

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

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: methoden@iqwig.de

Internet: www.iqwig.de

Inhaltsverzeichnis

	Seite
Abkürzungsverzeichnis.....	v
1 Dokumentation der Anhörung	1
2 Würdigung der Anhörung	2
2.1 Würdigung allgemeiner Anmerkungen.....	2
2.2 Würdigung der Stellungnahmen zu Abschnitt 1.3 „Evidenzbasierte Medizin“	3
2.3 Würdigung der Stellungnahmen zu Abschnitt 1.4 „Gesundheitsökonomie“	3
2.4 Würdigung der Stellungnahmen zu Abschnitt 2.1.3.2 „AbD-Konzept“	3
2.5 Würdigung der Stellungnahmen zu Abschnitt 2.1.3.3 „Recherche nach Indikationsregistern“	5
2.6 Würdigung der Stellungnahmen zu Abschnitt 2.1.4 „Kosten-Nutzen-Bewertung gemäß § 35b SGB V“	5
2.7 Würdigung der Stellungnahmen zu Abschnitt 2.2.3 „Gewährleistung der fachlichen Unabhängigkeit“	5
2.8 Würdigung der Stellungnahmen zu Abschnitt 3.2.1 „Umgang mit möglichem Reporting Bias“	5
2.9 Würdigung der Stellungnahmen zu Abschnitt 3.3.4 „Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V“	6
2.10 Würdigung der Stellungnahmen zu Abschnitt 3.5 „Diagnostische Verfahren“	13
2.11 Würdigung der Stellungnahmen zu Kapitel 4 „Kosten-Nutzen-Bewertung medizinischer Interventionen“	14
2.12 Würdigung der Stellungnahmen zu Abschnitt 4.2 „Kosten-Nutzen-Bewertung von Arzneimitteln gemäß § 35b SGB V“	14
2.13 Würdigung der Stellungnahmen zu Abschnitt 4.3.2 „Intervention und Komparator(en)“	15
2.14 Würdigung der Stellungnahmen zu Abschnitt 4.3.3 „Gesundheitsökonomische Zielgröße“	16
2.15 Würdigung der Stellungnahmen zu Abschnitt 4.3.4 „Perspektive“	16
2.16 Würdigung der Stellungnahmen zu Abschnitt 4.3.5 „Zeithorizont“	17
2.17 Würdigung der Stellungnahmen zu Abschnitt 4.5 „Effektivität sowie weitere klinische Inputparameter“	17
2.18 Würdigung der Stellungnahmen zu Abschnitt 4.6 „Nutzwerte (im Fall einer Kosten-Nutzen-Analyse)“	18
2.19 Würdigung der Stellungnahmen zu Abschnitt 4.7 „Kosten“	18

2.20	Würdigung der Stellungnahmen zu Abschnitt 4.7.2 „Schritte zur Kostenbestimmung“	19
2.21	Würdigung der Stellungnahmen zu Abschnitt 4.8 „Inflationsbereinigung und Diskontierung“	19
2.22	Würdigung der Stellungnahmen zu Abschnitt 4.9 „Gesundheitsökonomisches Studiendesign“	20
2.23	Würdigung der Stellungnahmen zu Abschnitt 4.10 „Ergebnisdarstellung“	20
2.24	Würdigung der Stellungnahmen zu Abschnitt 4.14 „Informationsbeschaffung und Datenquellen in der KNB“	21
2.25	Würdigung der Stellungnahmen zu Abschnitt 7.4 „Auswahl der Evidenz“	22
2.26	Würdigung der Stellungnahmen zu Kapitel 8 „Informationsbeschaffung“	22
2.27	Würdigung der Stellungnahmen zu Abschnitt 8.1.1 „Recherche in bibliografischen Datenbanken“	22
2.28	Würdigung der Stellungnahmen zu Abschnitt 8.1.2 „Suche in Studienregistern“ ..	23
2.29	Würdigung der Stellungnahmen zu Abschnitt 8.2.1 „Recherche nach systematischen Übersichten“	23
2.30	Würdigung der Stellungnahmen zu Abschnitt 8.4 „Recherche nach Leitlinien“	24
2.31	Würdigung der Stellungnahmen zu Abschnitt 9.1 „Qualitätsbewertung von Einzelstudien“	24
2.32	Würdigung der Stellungnahmen zu Abschnitt 9.1.1 „Kriterien für den Einschluss von Studien“	25
2.33	Würdigung der Stellungnahmen zu Abschnitt 9.1.3 „Rangordnung verschiedener Studienarten / Evidenzgrade“	26
2.34	Würdigung der Stellungnahmen zu Abschnitt 9.3.2 „Statistische Signifikanztests und Konfidenzintervalle“	26
2.35	Würdigung der Stellungnahmen zu Abschnitt 9.3.6 „Adjustierung und multifaktorielle Verfahren“	26
2.36	Würdigung der Stellungnahmen zu Abschnitt 9.3.7 „Metaanalysen“	28
2.37	Würdigung der Stellungnahmen zu Abschnitt 9.3.8 „Indirekte Vergleiche“	29
2.38	Würdigung der Stellungnahmen zu Abschnitt 9.3.12 „Umgang mit variablen Beobachtungszeiten“	29
2.39	Literaturverzeichnis	31
3	Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll	34
3.1	Teilnehmerliste der wissenschaftlichen Erörterung	34
3.2	Liste der Stellungnahmen, zu denen keine Vertreterin bzw. kein Vertreter an der wissenschaftlichen Erörterung teilgenommen hat.....	36
3.3	Tagesordnung der wissenschaftlichen Erörterung	36
3.4	Protokoll der wissenschaftlichen Erörterung	36

3.4.1	Begrüßung und Einleitung	37
3.4.2	Tagesordnungspunkt 1: Anwendungsbegleitende Datenerhebung	38
3.4.3	Tagesordnungspunkt 2: Kein Zusatznutzen auf der Basis von KNBs.....	42
3.4.4	Tagesordnungspunkt 3: Besonderheiten von Orphan Drugs in KNBs.....	48
3.4.5	Tagesordnungspunkt 4: 80 %-Regel	55
3.4.6	Tagesordnungspunkt 5: Umgang mit variablen Beobachtungszeiten	58
3.4.7	Tagesordnungspunkt 6: Signifikanz und multiples Testen.....	67
3.4.8	Tagesordnungspunkt 7: Verschiedenes	69
Anhang A	Dokumentation der Stellungnahmen	73

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AbD	anwendungsbegleitende Datenerhebung
AUC	Fläche unter der Kurve
EMA	European Medicines Agency (Europäische Arzneimittel-Agentur)
EORTC QLU-C10D	European Organisation for Research and Treatment of Cancer Quality of Life Utility-Core 10 Dimensions
G-BA	Gemeinsamer Bundesausschuss
GKV	gesetzliche Krankenversicherung
KNB	Kosten-Nutzen-Bewertung
PREFER	Patient Preferences in Benefit-Risk Assessments during the Drug Life Cycle
PRO	Patient-reported Outcome (patientenberichteter Endpunkt)
pU	pharmazeutischer Unternehmer
QALY	Quality-adjusted Life Year (qualitätsadjustiertes Lebensjahr)
RECORD	REporting of studies Conducted using Observational Routinely-collected health data
RECORD-PE	RECORD für Pharmakoepidemiologie
ROC	Receiver Operating Characteristic
VAS	visuelle Analogskala

1 Dokumentation der Anhörung

Am 06.12.2022 wurde der Entwurf der Allgemeinen Methoden Version 7.0 veröffentlicht und zur Anhörung gestellt. Bis zum 28.02.2023 konnten schriftliche Stellungnahmen eingereicht werden. Innerhalb dieser Frist gingen 38 Stellungnahmen per E-Mail ein. Darüber hinaus erhielten wir 1 Mitteilung, dass keine Stellungnahme abgegeben wird. Alle fristgerecht eingereichten Stellungnahmen sind im Anhang abgebildet.

Unklare Aspekte in den schriftlichen Stellungnahmen wurden in einer wissenschaftlichen Erörterung am 28.04.2023 im KOMED, Köln, diskutiert. Das Wortprotokoll der Erörterung befindet sich in Abschnitt 3.4.

2 Würdigung der Anhörung

Im Stellungnahmeverfahren zum Dokument „Allgemeine Methoden: Entwurf für Version 7.0“ vom 06.12.2022 gingen bis zum 28.02.2023 Stellungnahmen von 38 Organisationen, Institutionen, Firmen und Einzelpersonen ein. Alle Stellungnahmen wurden in der Überarbeitung des Methodenpapiers berücksichtigt. Wiederholte Argumente aus früheren Anhörungen werden hier nicht wiederholt gewürdigt. Ebenso werden Argumente zu Abschnitten, die in der vorliegenden Version des Methodenpapiers gar nicht verändert wurden, hier nicht gewürdigt, sofern diese Argumente nicht zu einer Anpassung des Methodenpapiers geführt haben. Sofern sich aus den Stellungnahmen Änderungen im Methodenpapier ergeben haben, ist dies im vorliegenden Dokument explizit beschrieben. Redaktionelle Hinweise wie z. B. Fehler im Literaturverzeichnis haben wir aufgenommen und umgesetzt; dies wird im Weiteren aber nicht näher ausgeführt.

2.1 Würdigung allgemeiner Anmerkungen

In einer Stellungnahme wird erfragt, warum im Verfahren des ThemenChecks Medizin die HTA-Berichte in ThemenCheck-Berichte umbenannt wurden.

In § 139b Abs. 5 SGB V hat der Gesetzgeber 2015 die Grundlage für ein bürgerorientiertes Verfahren geschaffen, bei dem Versicherte und interessierte Einzelpersonen Vorschläge für die Bewertung von Untersuchungs- und Behandlungsverfahren einreichen können. Der gesetzliche Auftrag wurde vom IQWiG im Jahr 2016 mit der Bezeichnung ThemenCheck Medizin umgesetzt. Zu ausgewählten Themen erstellen externe Sachverständige seitdem wissenschaftliche Bewertungen, die nicht nur eine Nutzenbewertung, sondern auch eine gesundheitsökonomische Bewertung und eine Auseinandersetzung mit ethischen, sozialen und weiteren Aspekten beinhalten. Diese Berichte wurden in der Vergangenheit u. a. auf dem Deckblatt der Berichte als „HTA-Bericht“ gekennzeichnet.

Das IQWiG erstellt weitere Berichte, etwa Bewertungen nicht medikamentöser Verfahren oder von Arzneimitteln, bei denen es sich – unter der Berücksichtigung nationaler und internationaler Standards – um HTA-Berichte handelt, auch wenn diese auf dem Deckblatt nicht als solche bezeichnet werden. Um dem Missverständnis vorzubeugen, dass nur für den ThemenCheck Medizin des IQWiG HTA-Berichte erstellt werden, und um für die Lesenden deutlicher zu machen, welche Berichte aus dem bürgerorientierten Verfahren nach § 139b Abs. 5 SGB V hervorgehen, wurden die im Rahmen des ThemenCheck Medizin erstellten HTA-Berichte in ThemenCheck-Berichte umbenannt.

2.2 Würdigung der Stellungnahmen zu Abschnitt 1.3 „Evidenzbasierte Medizin“

In 1 Stellungnahme wird angeregt, die Verbindung zwischen klinischer Expertise sowie „der Evidenz, dem klinischen Zustand und den Umständen, in denen sich eine Person befindet, und ihren Präferenzen und Einstellungen“ deutlicher anzusprechen.

Die Anregung wurde durch eine Anpassung umgesetzt.

2.3 Würdigung der Stellungnahmen zu Abschnitt 1.4 „Gesundheitsökonomie“

In einigen Stellungnahmen wird es als irreführend empfunden, dass die Preisverhandlungen als Ziel der Kosten-Nutzen-Bewertung (KNB) genannt wurden.

Die KNB ist kein regulärer Teil der Erstattungsbetragsverhandlung. Die gesetzlichen Regelungen, wann eine KNB initiiert werden kann, sind in Abschnitt 4.1 beschrieben. Sofern eine KNB beauftragt wird, können die Ergebnisse im Sinne von zusätzlichen ökonomischen Informationen neben den Ergebnissen der frühen Nutzenbewertung zur Orientierung bei den Erstattungsbetragsverhandlungen herangezogen werden.

2.4 Würdigung der Stellungnahmen zu Abschnitt 2.1.3.2 „AbD-Konzept“

In mehreren Stellungnahmen (auch zum Abschnitt 3.3.4) wird die frühzeitige Einbindung insbesondere des pharmazeutischen Unternehmers bei der Erarbeitung eines Konzepts für eine anwendungsbegleitende Datenerhebung (AbD) gefordert. Zudem wird in 1 Stellungnahme angemerkt, dass Aspekte aus dem Stellungnahmeverfahren unklar bleiben, weil bspw. die 4 Wochen Stellungnahmefrist für ein AbD-Konzept nicht erwähnt werden.

Die Erstellung eines AbD-Konzepts verläuft, wie auch bei anderen Produkten des IQWiG (bspw. der Dossierbewertung) üblich, ohne Einbindung des pharmazeutischen Unternehmers. Das anschließende Beteiligungsverfahren zu einem AbD-Konzept liegt im Verantwortungsbereich des G-BA und wird durch ihn durchgeführt. Der Ablauf, einschließlich der Stellungnahmefrist, ist in seiner Verfahrensordnung geregelt [8]. Über dieses Verfahren ist auch die frühzeitige Einbindung des betroffenen pharmazeutischen Unternehmers möglich. Es ergibt sich kein Änderungsbedarf für das Methodenpapier.

In 1 Stellungnahme wird kritisiert, dass die Informationsbeschaffung über bestehende Datenplattformen und Indikationsregister sehr allgemein dargestellt sei und nicht klar sei, ob bspw. nur nationale oder auch internationale Register betrachtet werden. Zudem sollten die Versorgungsbedingungen näher definiert werden, die für eine Verwendung der Datenplattformen und Indikationsregister zu fordern sind.

Die spezifische Methodik und die Ergebnisse der Informationsbeschaffung nach Indikationsregistern sind in den AbD-Konzepten beschrieben. Dabei werden auch die

Einschlusskriterien für ein potenziell relevantes Indikationsregister aufgeführt. Diese umfassen in der Regel, dass ein potenziell relevantes Register funktionsfähig ist und aktuell Daten von Patientinnen und Patienten der zu untersuchenden Patientenpopulation umfasst. Zudem müssen in dem Indikationsregister auch Daten von Patientinnen und Patienten aus Deutschland dokumentiert werden. Letzteres Einschlusskriterium soll sicherstellen, dass der deutsche Versorgungskontext im Rahmen der AbD ausreichend abgebildet ist. Inwiefern auch internationale Register bzw. ausländische Zentren den deutschen Versorgungskontext ausreichend abbilden können, hängt sehr von der spezifischen Fragestellung der AbD ab und ist im Einzelfall zu bewerten. In Abschnitt 3.3.4 des Methodenpapiers werden die regelhaften Einschlusskriterien für potenziell relevante Indikationsregister aufgeführt.

In 1 Stellungnahme wird angemerkt, dass in der Prozessbeschreibung des AbD-Konzepts die weiteren Aufgaben des IQWiG wie das Review des Studienprotokolls für die AbD ergänzt werden sollen.

Die in der Stellungnahme genannten weiteren Arbeiten werden in Form von Addenda zum AbD-Konzept erstellt und stellen somit separate Produkte dar. Sie sind somit nicht mehr Teil der Erstellung des AbD-Konzepts, sondern werden nach Beschluss des G-BA zur Forderung einer AbD durchlaufen. Für welche Arbeiten das IQWiG im Auftrag des G-BA eingebunden werden kann, ist in Abschnitt 3.3.4 (Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V [AbD-Konzepte]) der Allgemeinen Methoden beschrieben. Es ergibt sich kein Änderungsbedarf für das Methodenpapier.

In 1 Stellungnahme wird kritisiert, dass der Auswahlprozess und Einbezug von externen Sachverständigen und Registerbetreibern zu allgemein dargestellt sei. Zudem bleibe intransparent, wie genau bei der Auswahl von Registern und der Befragung von Registerbetreibenden vorgegangen wird.

Im Methodenpapier wird beschrieben, dass für die Erarbeitung eines AbD-Konzepts externe Sachverständige für die medizinische Expertise sowie Betroffene bzw. Patientenorganisationen für die Patientenperspektive eingebunden werden, und zwar auf die gleiche Weise wie auch für Dossierbewertungen. Die Identifikation der externen Sachverständigen erfolgt über eine institutseigene Sachverständigendatenbank. Wie Registerbetreiber eingebunden werden, ist im Methodenpapier und auch in den AbD-Konzepten selbst ebenfalls beschrieben. Es werden die Registerbetreibenden aller potenziell relevanten und grundsätzlich geeigneten Register kontaktiert. Informationen werden zum einen über einen Fragebogen eingeholt, der den Registerbetreibenden zur Verfügung gestellt wird, zum anderen über einen persönlichen Kontakt, bspw. in einem Videomeeting. Die Antworten aus dem Fragebogen sind im Anhang eines AbD-Konzepts transparent dargestellt. Es ergibt sich kein Änderungsbedarf für das Methodenpapier.

2.5 Würdigung der Stellungnahmen zu Abschnitt 2.1.3.3 „Recherche nach Indikationsregistern“

In verschiedenen Stellungnahmen werden präzisere Angaben zur Recherche nach Indikationsregistern gewünscht. Darunter fallen neben der Darstellung der Informationsquellen auch die Kriterien bei der Auswahl der Indikationsregister.

Auf eine detaillierte Darstellung der Informationsquellen wird im Methodenpapier verzichtet. Im Methodenpapier werden lediglich die allgemeinen Grundsätze beschrieben, die konkrete Umsetzung findet sich in der jeweilig erstellten Recherche nach Indikationsregistern. Pharmazeutische Unternehmen werden dabei als Informationsquelle in der Regel nicht berücksichtigt. Die Kriterien bei der Auswahl der Indikationsregister wurden im Methodenpapier in Abschnitt 3.3.4 ergänzt.

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

In einigen Stellungnahmen werden Klärungen zum Verfahrensprozess der KNB einschließlich spezifischer Verantwortlichkeiten im Hinblick auf die Einreichung eines Dossiers gefordert.

Zu diesem Punkt wird auf die Würdigung in Abschnitt 2.11 verwiesen.

2.7 Würdigung der Stellungnahmen zu Abschnitt 2.2.3 „Gewährleistung der fachlichen Unabhängigkeit“

Im Zusammenhang mit der Beurteilung der fachlichen Unabhängigkeit wird in 1 Stellungnahme die Terminologie „Darlegung von Beziehungen“ kritisiert. Dieser Ausdruck entspräche nicht internationalen Standards.

Das Institut verwendet im Entwurf der Allgemeinen Methoden Version 7.0 nicht die Formulierung „Darlegung von Beziehungen“, sondern „Offenlegung von Beziehungen“. Diese ergibt sich aus dem § 139a Abs. 3 SGB V, welcher die Aufgaben des Instituts regelt bzw. festlegt, dass externe Sachverständige, an die wissenschaftliche Forschungsaufträge vergeben werden, „... alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen offenzulegen [haben]“.

2.8 Würdigung der Stellungnahmen zu Abschnitt 3.2.1 „Umgang mit möglichem Reporting Bias“

In den Stellungnahmen wird einerseits die Verkürzung des Abschnitts 3.2.1 begrüßt, andererseits jedoch die dadurch fehlende Klarheit bezüglich der für die Bewertung geforderten Informationsdichte sowie der Folgen, die sich aus einer relevanten Menge nicht verfügbarer Daten ergeben, bemängelt.

Die in den Stellungnahmen aufgeworfene Kritik ist berechtigt. Der Hinweis zur Informationsdichte wurde durch Verweise auf Kapitel 9 zur Informationsbeschaffung, insbesondere zur Identifikation unpublizierter Studien und Daten durch die Beschaffung von Studienberichten (Clinical Study Reports), adressiert. Die sich aus einer mangelnden Datenverfügbarkeit ergebenden Folgen wurden deutlicher skizziert. Darüber hinaus wurde die eigentliche Zielsetzung des Abschnitts durch eine Änderung der Benennung in „Konsequenzen einer unvollständigen Datenverfügbarkeit“ betont.

2.9 Würdigung der Stellungnahmen zu Abschnitt 3.3.4 „Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V“

Allgemeine Anmerkung: Mehrere Stellungnahmen beziehen sich allgemein auf das Verfahren und den Verfahrensablauf der anwendungsbegleitenden Datenerhebung (inklusive der Fristen für mögliche Beteiligungen) und die anschließende erneute Nutzenbewertung sowie die erneuten Preisverhandlungen. Das Verfahren der anwendungsbegleitenden Datenerhebung ist ein Verfahren des G-BA und in seiner Verfahrensordnung beschrieben [8]. Auf entsprechende Stellungnahmen wird im Folgenden nicht eingegangen.

In 1 Stellungnahme wird angemerkt, dass die Begriffe „geplante oder laufende Datenerhebungen“ sowie „studienindividuelle Datenerhebung“ näher erläutert werden sollten.

Im Rahmen der Erarbeitung eines AbD-Konzepts wird geprüft, ob die wissenschaftlichen Fragestellungen, die mit der AbD beantwortet werden sollen, teilweise oder vollständig in bereits geplanten oder laufenden Datenerhebungen zum jeweiligen Arzneimittel adressiert werden. Unter diesen Datenerhebungen sind deshalb jegliche Arten von Studien zu verstehen, die für die Beantwortung der Fragestellung (oder einer Teilfragestellung) der geplanten AbD geeignet sind. Da die Fragestellung der AbD von dem Vergleich des Nutzens eines Arzneimittels mit demjenigen einer zweckmäßigen Vergleichstherapie handelt, müssen die geplanten oder laufenden Datenerhebungen vergleichend angelegt sein. Dafür sind verschiedene Studiendesigns denkbar, bspw. prospektive Kohortenstudien oder auch randomisierte kontrollierte Studien. Sofern es sich um einen nicht randomisierten Vergleich handelt, muss, wie in den Stellungnahmen festgestellt, aus den verfügbaren Informationen über die geplante oder laufende Datenerhebung hervorgehen, dass alle notwendigen Daten (zu Ein- / Ausschlusskriterien für die relevante Patientenpopulation, zu allen relevanten Confoundern, zu patientenrelevanten Endpunkten) in ausreichender Qualität erhoben werden.

Der Begriff „studienindividuelle Datenerhebung“ ist unabhängig vom zugrunde liegenden Studiendesign gemeint, sondern soll lediglich eine Abgrenzung bieten zu Situationen, in denen für eine AbD eine bereits bestehende Datenplattform, z. B. ein Indikationsregister, genutzt werden kann.

Insgesamt ergibt sich hier kein Änderungsbedarf im Methodenpapier.

In 1 Stellungnahme wird angemerkt, dass der Zeitraum der AbD nicht nur akzeptabel sein sollte, sondern dass die Daten für das Arzneimittel und die zweckmäßige Vergleichstherapie (zVT) auch zu ausreichend gleichen Zeiten erhoben werden, um „historische Vergleiche“ zu verhindern.

Die Formulierung des akzeptablen Zeitraums in Abschnitt 3.3.4 des Methodenpapiers bezieht sich auf die Machbarkeit einer AbD, nämlich dahin gehend, dass ausreichend Patientinnen und Patienten rekrutiert und beobachtet werden können, um mit einer AbD grundsätzlich aussagekräftige Ergebnisse für eine Nutzenbewertung generieren zu können. Der Zeitraum bezieht sich nicht auf die zeitlich parallele Erhebung der Daten. Es ist zwar zu bevorzugen, dass die Daten für das zu bewertende Arzneimittel und die zweckmäßige Vergleichstherapie prospektiv und zeitlich parallel erhoben werden. In Einzelfällen kann auch ein Rückgriff auf retrospektive Daten notwendig sein, bspw. wenn aufgrund von starken Präferenzen nicht ausreichend Patientinnen und Patienten zur Verfügung stehen, die die zweckmäßige Vergleichstherapie erhalten. Dies ist aber nur dann möglich, wenn die Anforderungen an die Vollständigkeit und Qualität der erhobenen Daten hinreichend erfüllt sind. Insgesamt ergibt sich kein Änderungsbedarf im Methodenpapier.

In mehreren Stellungnahmen wird kritisiert, dass die Anforderungen an Studiendurchführung und Registerqualität Anpassungen an den Registern erfordern, die für die Register teilweise nicht umsetzbar seien, und Register nicht auf eine umfassende Datenerhebung ausgelegt seien.

Im Rahmen einer Konzepterstellung wird geprüft, ob ein Register für eine AbD als primäre Datenquelle infrage kommt. Für ein solches Register wird im AbD-Konzept zudem der notwendige Anpassungsbedarf formuliert. Sollte die Prüfung ergeben, dass die Anpassungen nicht realistisch möglich sind, würde ein Register nicht als geeignete primäre Datenquelle bewertet. Die 1. AbD zu Onasemnogen-Abeparvovec läuft derzeit im SMARtCARE-Register [10]. Insofern trifft die Aussage in den Stellungnahmen, die Anforderungen seien für die Register nicht umsetzbar, nicht grundsätzlich zu. Die Verantwortung für die Qualität der erhobenen Daten liegt zudem beim pU. Sollte sich für eine AbD herausstellen, dass ein Register nicht als primäre Quelle geeignet ist, kann der pU stets eine studienindividuelle Datenerhebung in einer eigens für die AbD aufgestellten Datenplattform durchführen. Insgesamt ergibt sich kein Änderungsbedarf im Methodenpapier.

In mehreren Stellungnahmen wird gefordert, dass bei der Erstellung eines AbD-Konzepts auch Sachverstand aus den Bereichen klinische Entwicklung, Studienmanagement, Datenmanagement und IT-Lösungen in der klinischen Forschung einfließen solle.

Ein AbD-Konzept beschreibt lediglich die methodischen Grundzüge der AbD. Das AbD-Konzept bietet die Grundlage für die Entscheidung des G-BA zur Forderung einer AbD. Die Verantwortung für die konkrete Planung und praktische Umsetzung der AbD liegt hingegen beim pU. Nach der Konzepterstellung führt der G-BA ein Beteiligungsverfahren durch. Dadurch ist gewährleistet, dass der pU und andere Sachverständige entsprechende Expertise in die Diskussion einbringen können. Es ergibt sich kein Änderungsbedarf im Methodenpapier.

In mehreren Stellungnahmen wird infrage gestellt, inwiefern eine AbD bei seltenen Erkrankungen durchführbar sei. Insbesondere bei sehr seltenen Erkrankungen stelle die Rekrutierung einer ausreichenden Teilnehmerzahl für die Darstellung einer relevanten Effektstärke eine erhebliche Herausforderung dar.

Dass eine Erkrankung selten ist, bedeutet nicht grundsätzlich, dass eine AbD nicht realistisch durchführbar ist. In den AbD-Konzepten wird geprüft, inwiefern eine AbD realisierbar ist, dabei wird auch die zur Verfügung stehende Patientenzahl berücksichtigt. Wenn diese Prüfung ergibt, dass eine AbD in Deutschland aufgrund der zu geringen Patientenzahl nicht realistisch durchführbar ist, wird weiterhin geprüft, ob eine Ausweitung der AbD auf Zentren außerhalb von Deutschland möglich ist. Sollte sich ergeben, dass eine AbD aufgrund der zur Verfügung stehenden Patientenzahlen nicht realistisch durchführbar ist, wird dies entsprechend im Fazit des AbD-Konzepts beschrieben. Es ergibt sich kein Änderungsbedarf im Methodenpapier.

In mehreren Stellungnahmen wird kritisiert, dass die Anforderungen an die Effektstärke, ab der aufgrund des nicht randomisierten Studiendesigns erst eine Aussage zum Nutzen oder Schaden einer Intervention abgeleitet werden könne (z. B. im Bereich von 2 bis 5 für das relative Risiko [bzw. 0,2 bis 0,5 bei mortalitätssenkenden Interventionen] bezogen auf die untere bzw. obere Konfidenzintervallgrenze), zu hoch seien und wissenschaftlich nicht anerkannt seien. Zudem wird in mehreren Stellungnahmen gefordert, dass nicht allein das Effektmaß und die Lage des Konfidenzintervalls in einem solchen Bereich eine Entscheidung über einen Zusatznutzen begründen sollten, sondern stattdessen weitere Faktoren zu berücksichtigen seien wie die Konsistenz der Effekte, die Objektivität der Endpunkte und kontextuelle Aspekte (bspw. besondere Therapiesituation oder Art der Erkrankung).

In nicht randomisierten Studien gibt es verschiedene Biasquellen, die sich nicht abschließend vermeiden lassen. Die Berücksichtigung einer verschobenen Nullhypothese soll gewährleisten, dass vor dem Hintergrund der Unsicherheit trotzdem eine belastbare Aussage zum Zusatznutzen auf Basis solcher Daten möglich ist. Die Grenze von 2 bis 5 für das relative Risiko gibt nur einen Rahmen, in dem der Effekt sein muss, unsichere Daten wegen residualen Confoundings, Qualität der erhobenen Daten, ggf. andere Biasquellen (z. B. Festlegung des Zeitpunkts für den Beobachtungsstart). Die konkrete Schwelle ergibt sich gemäß A19-43 aus der konkreten Datenlage. Diese hängt aber von der spezifischen Datensituation ab und lässt sich nicht algorithmisch festlegen. Es ergibt sich kein Änderungsbedarf im Methodenpapier.

In mehreren Stellungnahmen wird kritisiert, dass die Anforderungen an die systematische Identifikation und Berücksichtigung von Confoundern zu hoch seien und dass stattdessen eine überschaubare Anzahl von zu kontrollierenden Confoundern herangezogen werden sollte, unter Beachtung der tatsächlich erhebbaren und nicht der theoretisch möglichen Parameter.

Confounding ist eine der wesentlichen Quellen für Verzerrung in nicht randomisierten Studien. Um diese Verzerrung bestmöglich zu vermeiden, ist eine systematische Herangehensweise an die Identifikation möglicher Confounder notwendig. Ein mögliches Verfahren ist in der Literatur beschrieben [23]. In der Erörterung haben die Stellungnehmenden auf Nachfrage keine anderen Verfahren zur systematischen Confounderidentifikation genannt. Die systematisch identifizierten möglichen Confounder sollten dann für die AbD vollständig erhoben und in der Auswertung der Daten berücksichtigt werden. Eine Einschränkung auf Basis von Verfügbarkeit der erhobenen oder erhebbaren Confounder stellt eine selbst potenziell zu Verzerrung führende, selektive Erhebung dar und ist somit für das Ziel, nämlich der möglichst großen Verringerung der Verzerrung durch Confounding, nicht ausreichend.

In mehreren Stellungnahmen wird angemerkt, dass der Abschnitt im Wesentlichen dem Rapid Report A19-43 des IQWiG entspreche und verschiedene Aspekte, die sich teilweise durch die vergangenen Verfahren ergeben haben, in Abschnitt 3.3.4 des Methodenpapiers nicht aufgegriffen seien. Dazu gehöre, dass ein Verhältnis der Patientinnen und Patienten von 1:1 zwischen den Behandlungsgruppen nicht wahrscheinlich sei. Zudem sollte der Umgang mit anderen Quellen für systematische Verzerrungen, wie die Festlegung des Zeitpunkts des Beobachtungsstarts, in Abschnitt 3.3.4 des Methodenpapiers aufgenommen werden.

Die in den Stellungnahmen genannten Aspekte wie die Berücksichtigung des Verteilungsverhältnisses der Patientinnen und Patienten auf die Behandlungsgruppen und die Berücksichtigung eines gleichen Beobachtungsstarts bezogen auf beide Behandlungsgruppen werden regelhaft in den aktuellen AbD-Konzepten adressiert [14-17]. Diesen Aspekten wird zugestimmt und das Methodenpapier entsprechend ergänzt.

In mehreren Stellungnahmen wird kritisiert, dass bei der Konzepterstellung mögliche Veränderungen, bspw. durch die Zulassung weiterer Therapien oder durch Änderung in den Versorgungsanteilen über die Zeit, nicht berücksichtigt werden.

Die Arzneimittel, für die ein AbD-Konzept beauftragt wird, sind entweder noch nicht zugelassen oder erst seit kurzer Zeit auf dem Markt verfügbar. Informationen zu möglichen Veränderungen über die Zeit bspw. bezüglich der Versorgungsanteile werden bei der Konzepterstellung berücksichtigt, sofern sie belastbar sind. In der Regel liegen sie zum Zeitpunkt der Konzepterstellung jedoch nicht vor und können demnach nicht berücksichtigt werden. Es ergibt sich kein Änderungsbedarf für das Methodenpapier.

In mehreren Stellungnahmen wird gefordert, dass die Realisierbarkeit einer AbD im Methodenpapier näher konkretisiert und operationalisiert werden müsse.

Die Kriterien, die bei der Einschätzung der Realisierbarkeit einer AbD berücksichtigt werden, sind zum einen im Rapid Report A19-43 und zum anderen im Methodenpapier beschrieben. Inwiefern die (Nicht-)Erfüllung dieser Kriterien im Rahmen eines AbD-Konzepts dazu führt, dass eine AbD realistisch durchführbar ist oder nicht, lässt sich nicht algorithmisch festlegen und hängt von der spezifischen Fragestellung der AbD ab. Es ergibt sich kein Änderungsbedarf im Methodenpapier.

In 1 Stellungnahme wird sich auf die Fallzahlschätzung im Rahmen der AbD-Konzepterstellung bezogen, dass diese transparent dargestellt und für mehr als einen Endpunkt durchgeführt werden solle und dass verschiedene Szenarien berechnet werden sollen.

Wie die Fallzahlplanung in klinischen Studien bezieht sich die orientierende Fallzahlschätzung in einem AbD-Konzept in der Regel auf einen Endpunkt. Es wird derjenige Endpunkt gewählt, für den unter Berücksichtigung der zur Verfügung stehenden Patientenzahlen aussagekräftige Ergebnisse in der AbD erwartbar sind. Die der Schätzung zugrunde liegenden Annahmen werden in den AbD-Konzepten dargestellt und begründet. Um der Unsicherheit der in die Schätzung der Fallzahl zugrunde liegenden Daten zu berücksichtigen, werden in den aktuellen AbD-Konzepten zudem regelhaft verschiedene Szenarien für die Fallzahlschätzung berechnet [15-17]. Es ergibt sich kein Änderungsbedarf für das Methodenpapier.

In 1 Stellungnahme wird angemerkt, dass die Festlegung der Endpunkte nicht durch die Verfügbarkeit dieser Daten in einem möglicherweise geeigneten Register beeinflusst werden sollte und dass eine Fokussierung auf einzelne verfügbare Endpunkte nicht die Basis einer Nutzenbewertung mithilfe einer AbD sein dürfe.

Dieser Stellungnahme wird zugestimmt. Wie in den veröffentlichten AbD-Konzepten ersichtlich, werden im Rahmen der Erarbeitung eines AbD-Konzepts bei der Definition der wissenschaftlichen Fragestellung die für die Bewertung des Zusatznutzens patientenrelevanten Endpunkte beschrieben, und zwar unabhängig von den im jeweiligen Register verfügbaren Daten. Vielmehr ist in den AbD-Konzepten sogar klar beschrieben, wenn die entsprechenden Register aufgrund dessen angepasst werden müssen, weil relevante Endpunkte nicht erhoben sind. Die orientierende Fallzahlschätzung basiert wiederum auf einem patientenrelevanten Endpunkt, für den sich auf Basis der verfügbaren Informationen und unter Berücksichtigung der für die AbD zur Verfügung stehenden Patientenzahlen Stichprobengrößen ergeben, die für den Nachweis eines Vorteils des zu bewertenden Arzneimittels benötigt werden.

In 1 Stellungnahme wird angemerkt, dass der Begriff der zweckmäßigen Vergleichstherapie nicht von Zulassungsstudien übertragbar auf versorgungsnahe Daten und die AbD sei, da sich bei einer nicht randomisierten, nicht interventionellen Studie die Verfügbarkeit und Auswahlmöglichkeiten für Therapien unkontrollierter ändern können als bei einer interventionellen Zulassungsstudie.

Der G-BA legt für die AbD eine zweckmäßige Vergleichstherapie fest, unabhängig davon, ob es sich um ein Orphan Drug handelt oder nicht. Insofern ist der Begriff der zweckmäßigen Vergleichstherapie auch im Kontext der AbD richtig. Wie in Zulassungsstudien auch kann es in nicht interventionellen Studien (wie einer AbD) im Verlauf der Beobachtung zu Änderungen der Behandlung bei Patientinnen und Patienten kommen. Diese können für die Interpretierbarkeit der Ergebnisse problematisch sein, bspw. wenn Patientinnen und Patienten aus dem Vergleichsarm auf eine Behandlung mit dem zu bewertenden Arzneimittel wechseln. Andererseits sind Situationen denkbar, in denen ein Wechsel der Behandlung kein Problem für die Interpretierbarkeit der Ergebnisse darstellt. Für die Auswertung der Ergebnisse aus der AbD ist eine ITT-Auswertung zu bevorzugen. Therapiewechsel können ggf. im Rahmen von Sensitivitätsanalysen adressiert werden. Dies ist vom pU entsprechend im Studienprotokoll bzw. im statistischen Analyseplan vorab zu spezifizieren. Es ergibt sich kein Änderungsbedarf für das Methodenpapier.

In mehreren Stellungnahmen wird gefordert, dass im Methodenpapier bzw. in den AbD-Konzepten bereits eine abschließende Liste der relevanten Confounder, klare Kriterien für die Vergeblichkeit einer AbD sowie detailliertere Angaben zur Methodik der Auswertung (genaue Beschreibung der statistischen Verfahren) aufgenommen werden. In 1 Stellungnahme wird weiterhin vorgeschlagen, dass einerseits prinzipielle von der AbD unabhängige Qualitätskriterien für Register im Methodenpapier ergänzt werden. Zudem wird hier kritisiert, dass zur Dauer der AbD auch die Planung mit Erstellung des Studien- und Registerprotokolls, die Umsetzung oder Anpassung der Datenplattform gehöre und dies im Methodenpapier entsprechend ergänzt werden müsse.

Wie im Methodenpapier beschrieben, soll das AbD-Konzept zum einen die methodischen Anforderungen an die Durchführung der AbD einschließlich der Auswertung zum Zwecke der Nutzenbewertung gemäß § 35a SGB V gegenüber einer zVT beschreiben, zum anderen soll mithilfe des AbD-Konzepts geprüft werden, ob die geplante AbD prinzipiell realisierbar ist. Im Methodenpapier und in den AbD-Konzepten können nur die grundlegenden methodischen Anforderungen dargestellt werden. Ein AbD-Konzept ersetzt keine konkrete Studienplanung. Dazu gehören auch genauere Festlegungen, bspw. zu Vergeblichkeitskriterien. Diese liegt im Verantwortungsbereich des pU und ist im Detail vorab im Studienprotokoll bzw. im statistischen Analyseplan darzulegen. Es ergibt sich kein Änderungsbedarf im Methodenpapier.

In mehreren Stellungnahmen wird gefordert, dass in einem AbD-Konzept auch geprüft werden soll, ob nicht alternative Studienformen zu aussagekräftigeren Daten führen können als eine AbD. Dabei werden beispielsweise indirekte Vergleiche mit externen bzw. „historischen“ Kontrollen genannt.

In den AbD-Konzepten wird geprüft, ob es laufende oder geplante Datenerhebungen gibt, die für die Beantwortung der Fragestellung der Nutzenbewertung geeignet sind. Die in den vergangenen AbD-Konzepten identifizierten laufenden oder geplanten Datenerhebungen waren durchweg nicht für die Beantwortung der Fragestellung, nämlich den Vergleich des Nutzens des zu bewertenden Arzneimittels mit dem der zweckmäßigen Vergleichstherapie, geeignet. Dies ist in fast allen Fällen darauf zurückzuführen, dass die Datenerhebungen nicht vergleichend angelegt waren. Eine darüber hinausgehende Prüfung, ob alternative Studienformen zu aussagekräftigeren Daten führen, ist nicht Gegenstand eines AbD-Konzepts. Darzulegen, ob für eine AbD auch retrospektive Daten, z. B. aus einem Register, genutzt werden können, liegt im Verantwortungsbereich des pU. Es ergibt sich kein Änderungsbedarf im Methodenpapier.

In 1 Stellungnahme wird vorgeschlagen, dass bestehenden Datenplattformen bei der Konzepterstellung für eine AbD die Möglichkeit gegeben werden sollte, sich für die AbD unter Darstellung ihres Potenzials zu bewerben.

Bei der Konzepterstellung wird eine Recherche nach bestehenden Registern durchgeführt. Diese schließt verschiedene Suchquellen ein, darunter verschiedene Datenbanken sowie optionale Befragungen von Experten oder gezielte Websuchen. Im Beteiligungsverfahren besteht darüber hinaus auch die Möglichkeit, weitere Datenplattformen zu nennen. Daraus haben sich in den vergangenen Verfahren bisher keine weiteren potenziell relevanten Datenquellen ergeben. Es ergibt sich kein Änderungsbedarf für das Methodenpapier.

In 1 Stellungnahme wird vorgeschlagen, in Abschnitt 3.3.4 bezüglich der Adjustierung für Confounder neben dem Verweis auf den Rapid Report A19-43 zu ergänzen, dass beim Vorliegen von zeitabhängigem Confounding ggf. spezifische Kausalinferenz-Verfahren zur Kontrolle dieses Confounding-Typs erforderlich sind wie z. B. Marginal Structural Models.

Eine solche Ergänzung ist sinnvoll. Diese wurde aber nicht in Abschnitt 3.3.4 vorgenommen, sondern in Abschnitt 9.3.6, auf den wiederum in Abschnitt 3.3.4 verwiesen wird.

In 1 Stellungnahme werden Vorschläge zur Diskussion der Vor- und Nachteile verschiedener Auswertungsverfahren zur Adjustierung für Confounder gemacht, die sich primär auf den Rapid Report A19-43 beziehen. Es wird betont, dass sich die Wahl der konkreten Auswertungsmethode an den Zielen der jeweiligen Auswertung orientieren sollte.

Gegenstand dieses Stellungnahmeverfahrens sind die Allgemeinen Methoden und nicht der Rapid Report A19-43. Wir werden die Vorschläge zur Diskussion der Vor- und Nachteile verschiedener Auswertungsverfahren zur Adjustierung für Confounder berücksichtigen, falls es in der Zukunft zu einem Update des Rapid Reports A19-43 kommen sollte. Der allgemeine Aspekt, dass sich die Wahl der konkreten Auswertungsmethode an den Zielen der jeweiligen Auswertung orientieren sollte, wurde in Abschnitt 9.3.6 ergänzt.

2.10 Würdigung der Stellungnahmen zu Abschnitt 3.5 „Diagnostische Verfahren“

In 1 Stellungnahme wird darauf hingewiesen, dass „der G-BA nicht zwingend die höchste Evidenzstufe fordert, sondern auch andere Evidenzstufen mit Studien zur Testgenauigkeit berücksichtigt“. Auch unterscheide die Verfahrensordnung des G-BA, ob es um „diagnostischen oder therapeutischen Nutzen“ gehe.

In der Nutzenbewertung diagnostischer Methoden lässt sich kein methodischer Unterschied zwischen G-BA und IQWiG erkennen. Beide Organisationen halten – im Einklang mit internationalen Standards der evidenzbasierten Medizin – unter bestimmten Umständen auch Evidenz zur Testgenauigkeit für ausreichend. Dies lässt sich auch am konkreten Beispiel belegen [9]. Die Verfahrensordnung des G-BA enthält keine grundsätzlich andere Nutzendefinition für diagnostische im Vergleich zu therapeutischen Methoden, sondern verweist auf die Besonderheit, dass diagnostischer Nutzen sich regelhaft allein durch therapeutische Konsequenzen ergibt. Daher brauche es vor allem Unterlagen „zum Nachweis der therapeutischen Konsequenz einer diagnostischen Methode“ (§ 10, Satz 2(1) G-BA-VerfO [8]). Letztlich gehe es auch bei Diagnostik darum, dass bei der Bewertung „die Verwendung patientenrelevanter Zielgrößen (wie z. B. Mortalität, Morbidität, Lebensqualität)“ berücksichtigt werde. Insgesamt ergibt sich hier kein Änderungsbedarf im Methodenpapier.

In derselben Stellungnahme wird auch die Sorge geäußert, dass zu hohe methodische Ansprüche an einen neuen diagnostischen Test in der Coronapandemie erhebliche Nachteile herbeigeführt hätten: Hätte man Studien zur diagnostisch-therapeutischen Kette abgewartet, „wären Coronatests erst viel später verfügbar gewesen“.

Es war in der Tat sinnvoll, nicht auf Studienergebnisse zu warten, die den Einfluss von Coronatests auf patientenrelevante Zielgrößen, z. B. infektionsbedingte Morbidität und Mortalität, prüfen. Denn es ist selbstverständlich, dass in einer pandemischen Situation zur akuten Gefahrenabwehr Evidenzanforderungen abgesenkt werden müssen [12].

In 1 Stellungnahme wird das PRISMA-DTA-Instrument (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Diagnostic Test Accuracy Studies) in Abschnitt 3.5 des Methodenpapiers vermisst.

Dieses Instrument [20] wird bereits in Abschnitt 9.2.3 des Methodenpapiers aufgeführt, sodass in Abschnitt 3.5 kein Ergänzungsbedarf besteht.

In 1 Stellungnahme wird angeregt, Ergebnisse zur Testgüte zusätzlich auch in Form einer Receiver-Operating-Characteristic-Kurve (ROC-Kurve) und der zugehörigen Area under the curve (AUC) zu betrachten und darzustellen.

ROC-Kurven mitsamt der zugehörigen AUC sind sehr nützliche Darstellungs- und Analyseformen zu Testgüte-Fragestellungen. Auch wenn beide Begriffe in Abschnitt 3.5 nicht genannt werden, schließt dies nicht aus, dass entsprechende Analysen in der Berichtsarbeit des IQWiG verwendet werden. Da Testgüte-Metaanalysen und ihre Darstellung als summarische ROC-Kurve ausführlich in Abschnitt 9.3.7 des Methodenpapiers beschrieben werden, erscheint eine Ergänzung in Abschnitt 3.5 nicht erforderlich.

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

In den Stellungnahmen wird gefordert, einen Abschnitt zum Verfahrensprozess der KNB zu ergänzen. In 1 weiteren Stellungnahme wird darauf hingewiesen, die Verantwortlichkeiten für die KNB (insbesondere im Hinblick auf die Einreichung eines Dossiers) zu beschreiben.

Das Nähere zum Verfahren der Bewertung des Kosten-Nutzen-Verhältnisses von Arzneimitteln regelt der G-BA in seiner Verfahrensordnung (5. Kapitel 2. Abschnitt). Dort ist auch die Möglichkeit einer Einreichung eines Dossiers durch einen pharmazeutischen Hersteller (pU) beschrieben. Auf diese Möglichkeit wird auch im Methodenpapier explizit hingewiesen: Für eine zeitnahe und aussagekräftige KNB müssen die vom pU für die KNB erstellten Dossiers die notwendigen Daten beinhalten, z. B. auch Daten, die für die Erstellung der Modelle notwendig sind und die vom Inhalt der vorgelagerten Nutzenbewertung nicht erfasst werden.

2.12 Würdigung der Stellungnahmen zu Abschnitt 4.2 „Kosten-Nutzen-Bewertung von Arzneimitteln gemäß § 35b SGB V“

In 1 Stellungnahme wird angeregt, den inkrementellen Charakter der Kosten-Nutzen-Bewertung als Form einer vollständigen gesundheitsökonomischen Evaluation hervorzuheben. Des Weiteren wird darauf hingewiesen, dass es sich bei einer Ausgaben-Einfluss-Analyse um einen anderen Studientyp als eine Kosten-Effektivitäts-Analyse bzw. eine Kosten-Nutzwert-Analyse handelt und diese daher nicht ergänzend, sondern zusätzlich erfolgt.

Entsprechende redaktionelle Anpassungen wurden in der Tabelle in Abschnitt 4.2 und im Text vorgenommen.

2.13 Würdigung der Stellungnahmen zu Abschnitt 4.3.2 „Intervention und Komparator(en)“

In einigen Stellungnahmen wird darauf hingewiesen, dass im Entwurf zum Methodenpapier 7.0 nicht ersichtlich sei, wie im Falle einer patientenindividuellen Therapie oder Therapie nach Maßgabe des Arztes als zVT vorgegangen werden soll.

Für die Umsetzung der patientenindividuellen Therapie oder Therapie nach ärztlicher Maßgabe im Rahmen der KNB ist zu prüfen, inwieweit ein Therapiemix herangezogen werden kann. Dabei werden insbesondere die Umsetzung im Rahmen der vorangegangenen Nutzenbewertung und ggf. Versorgungsanteile zur Gewichtung der einzelnen Therapieoptionen berücksichtigt.

In einigen Stellungnahmen wird angemerkt, dass im Entwurf zum Methodenpapier 7.0 den Besonderheiten von Orphan Drugs im Hinblick auf etwaige KNB nicht Rechnung getragen wird.

Es wurde in Abschnitt 4.3.2 zusätzlich auf Orphan Drugs als Sonderfall im Kontext der KNB eingegangen. Darin wird einerseits beschrieben, dass eine KNB zu Orphan Drugs grundsätzlich nicht ausgeschlossen ist und in welchen Situationen bzw. unter welchen Voraussetzungen es dazu kommen kann. Eine KNB zu Orphan Drugs lässt sich nicht in jeder Situation sinnvoll durchführen. Häufig liegen bei Markteinführung nur eingeschränkte Daten zu Wirksamkeit und Sicherheit vor oder es mangelt an Daten zum natürlichen Krankheitsverlauf und epidemiologischen Daten zu seltenen Erkrankungen [22]. Die in den Stellungnahmen vielfach kritisierte Auswahl von Vergleichsinterventionen für Orphan Drugs im Rahmen einer KNB obliegt nicht dem IQWiG, sondern dem Auftraggeber.

In 1 Stellungnahme wird thematisiert, dass unklar sei, wie Jahrestherapiekosten zu berechnen sind. Des Weiteren wird darauf hingewiesen, dass die Darstellung der Therapiekosten im Rahmen der KNB sich grundsätzlich von der geforderten Darstellung der Jahrestherapiekosten im Dossier gemäß § 35a SGB V unterscheidet.

Der Hinweis im Methodenpapier auf die Jahrestherapiekostenberechnung im Dossier zur Nutzenbewertung gemäß § 35a SGB V diente der Erläuterung der Bedeutung der geforderten gesonderten Analysen unter Berücksichtigung der Kosten aller Therapieoptionen der zVT. Tatsächlich ist die Berechnung der Therapiekosten im Rahmen der KNB nicht zwangsläufig gleichzusetzen mit den Berechnungen der Jahrestherapiekosten im Dossier zur Nutzenbewertung gemäß § 35a SGB V. Bei der Berechnung der Therapiekosten im Rahmen der KNB werden z. B. in Anlehnung an die tatsächliche Therapiedauer ggf. längere Zeiträume betrachtet.

2.14 Würdigung der Stellungnahmen zu Abschnitt 4.3.3 „Gesundheitsökonomische Zielgröße“

Die Verwendung von qualitätsadjustierten Lebensjahren (QALYs) als Maß für den Gesamtnutzen wird in einigen Stellungnahmen begrüßt und in anderen Stellungnahmen kritisiert. Es wird gefordert, die Vor- und die Nachteile des QALY-Konzepts in einem separaten Abschnitt zu diskutieren.

Gemäß den gesetzlichen Vorgaben ist im Auftrag u. a. festzulegen, welche Art des Nutzens und welches Maß des Gesamtnutzens bei der KNB zu berücksichtigen sind [1]. Im Methodenpapier wurden verschiedene Optionen für die Wahl der gesundheitsökonomischen Zielgröße aufgeführt und dabei wurde auch auf die kritische Diskussion zum QALY-Konzept mit entsprechenden Quellen hingewiesen. Des Weiteren wurde die Diskussion um das QALY-Konzept bereits in früheren Versionen des Methodenpapiers aufgegriffen.

Im Methodenpapier wird für den Fall einer KNB bezogen auf mehrere Endpunkte der Nutzenbewertung die Möglichkeit einer Gewichtung der Ergebnisse unter Berücksichtigung von Verfahren der Präferenzhebung von Patientinnen und Patienten beschrieben. In 1 Stellungnahme wird eine Konkretisierung der Methodik für eine Präferenzhebung, was ihr zufolge die Akzeptanz der Ergebnisse im Rahmen einer KNB seitens des IQWiG erhöhen könnte, gefordert. In 1 anderen Stellungnahme wird vorgeschlagen, auf eine Gewichtung zu verzichten, und im Falle mehrerer, relevanter Endpunkte nur eine sehr beschränkte Anzahl Endpunkte als gesundheitsökonomische Zielgröße der KNB zu beauftragen, die sich auch sinnvoll in einer modellbasierten Evaluation abbilden ließe.

Im Methodenpapier wurde auf die wissenschaftliche Diskussion zur Nutzbarkeit der Verfahren der Präferenzhebung, der Herausforderungen und zu berücksichtigender Aspekte mit entsprechenden Referenzen verwiesen. In diesen werden insbesondere auch methodische Herausforderungen thematisiert [4,11,18]. Auch die Europäische Arzneimittelbehörde (EMA) hat in ihrem Qualifizierungsgutachten zum Rahmenwerk des Patient Preferences in Benefit-Risk Assessments during the Drug-Life-Cycle(PREFER)-Projekts u. a. darauf hingewiesen, dass jede Patientenpräferenzstudie, unabhängig davon, ob sie dem Rahmenwerk entspricht, entsprechend seinen Zielen und seinem spezifischen Anwendungsfall bewertet werden [6]. Vor dem Hintergrund der dargestellten Diskussion erfolgt im Methodenpapier keine weitere Festlegung auf ein spezifisches Verfahren oder methodisches Vorgehen. Dies ist u. a. abhängig von der konkreten Fragestellung und kann ggf. im Rahmen der Auftragserteilung erfolgen.

2.15 Würdigung der Stellungnahmen zu Abschnitt 4.3.4 „Perspektive“

In mehreren Stellungnahmen wird kontrovers für bzw. gegen die regelhafte oder auftragsbezogene Berücksichtigung verschiedener Perspektiven argumentiert.

Im Methodenpapier wird die Perspektive der Versichertengemeinschaft der gesetzlichen Krankenversicherung (GKV) als Referenzfall genannt (siehe § 35b SGB V und Verfahrensordnung des G-BA). Die Betrachtung ergänzender Perspektiven ist nicht ausgeschlossen und grundsätzlich auftragsabhängig.

In 1 Stellungnahme wird angemerkt, dass Doppelzählungen bei Kosten vermieden und Kosten unterschiedlicher Perspektiven nicht miteinander verrechnet werden sollen.

Im Methodenpapier wird in Abschnitt 4.3.4 Perspektive ein entsprechender Hinweis zur Sensibilisierung aufgenommen.

2.16 Würdigung der Stellungnahmen zu Abschnitt 4.3.5 „Zeithorizont“

In einigen Stellungnahmen wird die Festlegung eines festen Zeithorizonts von 5 Jahren für die Szenarioanalyse kritisiert.

Im Methodenpapier wurden redaktionelle Anpassungen zur Klarstellung vorgenommen. Wie im Methodenpapier in Abschnitt 4.3.5 dargelegt, richtet sich die Festlegung des Zeithorizonts in der Basisfallanalyse nach der Fragestellung (Indikation) bzw. dem Auftrag. In den ergänzenden Sensitivitätsanalysen wird jedoch zwecks Vergleichbarkeit ein fixer Zeithorizont gewählt, der einer mittelfristigen Betrachtung entspricht. Weitere Analysen mit anderen Zeiträumen wie ein kurzfristiger Zeithorizont von 1 Jahr können ergänzend durchgeführt werden.

2.17 Würdigung der Stellungnahmen zu Abschnitt 4.5 „Effektivität sowie weitere klinische Inputparameter“

In einigen Stellungnahmen wird die Prämisse, dass das Ergebnis der KNB das Ergebnis der vorangegangenen Nutzenbewertung nicht infrage stellen soll, kritisiert.

Hintergrund ist, dass im § 35b SGB V festgelegt ist, dass die KNB unter Berücksichtigung des therapeutischen Zusatznutzens für die Patienten im Verhältnis zu den Kosten erfolgen soll. Des Weiteren sind mit dem Beschluss zur KNB der Zusatznutzen sowie die Therapiekosten bei Anwendung des jeweiligen Arzneimittels festzustellen [1]. Im Methodenpapier wurde beschrieben, dass sich die modellbasierte KNB nicht als Basis für einen Nachweis des Zusatznutzens im Sinne des § 35a SGB V eignet. Grund ist, dass sich die Zielsetzung modellbasierter KNB von der Zielsetzung klinischer Studien unterscheidet. Diese wurde im Methodenpapier ausführlich erläutert und auch in der wissenschaftlichen Erörterung am 28.04.2023 thematisiert. Im Rahmen der Erörterung wurden jedoch keine alternativen Vorschläge zur Berücksichtigung des Zusatznutzens in der KNB vorgebracht. Es erfolgt keine Änderung im Methodenpapier.

2.18 Würdigung der Stellungnahmen zu Abschnitt 4.6 „Nutzwerte (im Fall einer Kosten-Nutzen-Analyse)“

Es wird in einigen Stellungnahmen angemerkt, dass direkte Erhebungsmethoden wie Time-Trade-Off oder Standard Gamble aus Gründen des Mehraufwandes und der kognitiven Belastung im Rahmen von klinischen Studien nicht möglich sind.

Im Methodenpapier ist bereits der Hinweis auf die visuelle Analogskala (VAS) enthalten, die im Rahmen der KNB eingesetzt werden kann. Hierbei handelt es sich um Bewertungen von Patientinnen und Patienten, die ihren eigenen Gesundheitszustand auf einer Skala z. B. von 0 bis 100 einordnen. Da es sich um ein einziges Item handelt, stellt dieser Ansatz weder einen hohen Mehraufwand noch eine große kognitive Belastung dar. Auf die wissenschaftliche Diskussion um die Verwendung der VAS wurde im Methodenpapier hingewiesen.

In mehreren Stellungnahmen wird angemerkt, dass die Zugrundelegung von Bewertungen von Patientinnen und Patienten zu den Gesundheitszuständen aufgrund einer fehlenden Datenverfügbarkeit (da konträr zum gesundheitsökonomischen Standard) ein wenig praktikabler Ansatz ist und auch ein Nachweis, dass keine Unterschiede zwischen Patientinnen und Patienten und der Allgemeinbevölkerung bestehen, deshalb schwierig ist. Des Weiteren wurden verschiedene Klarstellungen bzw. Konkretisierungen der Vorgaben in den Stellungnahmen gefordert.

Im Methodenpapier wurden Anpassungen und Ergänzungen zur Klarstellung vorgenommen. Nutzwerte sollten auf Bewertungen von Patientinnen und Patienten basieren. Des Weiteren sei auf aktuelle Forschung im Kontext des European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Utility-Core 10 Dimensions (QLU-C10D) Instruments hingewiesen, ein Value Set auf Basis der Bewertungen von onkologischen Patientinnen und Patienten zu entwickeln. [7]. Nutzwerte, die auf Bewertungen der Allgemeinbevölkerung basieren, sind in Ausnahmefällen möglich.

In einigen Stellungnahmen wird darauf hingewiesen, dass eine Bewertung von Gesundheitszuständen durch Patientinnen und Patienten in besonderen Fällen (z. B. Alzheimer, Pädiatrie) aufgrund des Zustands ggf. nicht möglich ist.

Dieser Aspekt wurde im Methodenpapier adressiert.

2.19 Würdigung der Stellungnahmen zu Abschnitt 4.7 „Kosten“

In mehreren Stellungnahmen werden zum Teil missverständliche bzw. unklare Formulierungen im Methodenpapier angemerkt.

Im Methodenpapier wurden redaktionelle Anpassungen zur Klarstellung beim Umgang mit Zeitkosten sowie der Bewertung von Ressourcen vorgenommen.

2.20 Würdigung der Stellungnahmen zu Abschnitt 4.7.2 „Schritte zur Kostenbestimmung“

In 1 Stellungnahme wird gefragt, wie im Rahmen der Perspektive der GKV-Versichertengemeinschaft mit den Kosten informeller Pflege und Unterstützungsleistungen umzugehen ist, da diese zwar unentgeltlich erbracht werden, aber teilweise das Potenzial einer Kassenleistung bergen würden.

Informelle Pflege bzw. Unterstützungsleistungen sind in der Perspektive der GKV-Versichertengemeinschaft im Sinne von Zeitkosten der Angehörigen als direkte Kosten zu berücksichtigen und gesondert auszuweisen. Im Unterabschnitt „Identifizierung der Ressourcen“ wurde der Deutlichkeit halber ein Hinweis aufgenommen, dass Zeitkosten der Angehörigen in der besagten Perspektive zu berücksichtigen sind.

2.21 Würdigung der Stellungnahmen zu Abschnitt 4.8 „Inflationsbereinigung und Diskontierung“

In mehreren Stellungnahmen werden Anpassungsvorschläge zu den Diskontraten unterbreitet. Diese beziehen sich auf deren Höhe (Vorschlag einer Anpassung an die durchschnittlichen Zinsen der letzten 12 Monate), auf die Annahme einer differenziellen Diskontierung (im Basisfall oder zumindest in einer regelhaften Sensitivitätsanalyse) und die Wahl der Diskontraten in Abhängigkeit von der jeweiligen Situation.

Die Umsetzung des vorgeschlagenen Ansatzes einer Anpassung der Diskontraten an den Durchschnitt der langfristigen Zinsen für Bundesanleihen (Approximation des Kapitalmarktes) über die letzten 12 Monate hätte zur Folge, dass die gesundheitsökonomischen Evaluationen aus Deutschland aufgrund der sich ständig verändernden Diskontierungsrate nicht mehr untereinander vergleichbar wären [21]. Darüber hinaus liegt die Rate in Höhe von 3 % innerhalb der Spanne der international verwendeten Diskontraten und fördert somit auch deren Vergleichbarkeit [5].

Die in 2 Stellungnahmen geforderte differenzielle Diskontierung (Annahme einer geringeren Diskontrate für den Nutzen als für die Kosten) wurde bereits als mögliche Sensitivitätsanalyse ins Methodenpapier aufgenommen (einzige Änderung gegenüber der Vorversion) und ist damit nicht ausgeschlossen.

Die Wahl von Diskontraten in Abhängigkeit der jeweiligen Situation stellt keine international übliche Praxis dar [3]. Lediglich in wenigen Ländern wie dem Vereinigten Königreich findet bei Betrachtung sehr langer Zeithorizonte (z. B. ab 30 Jahren) eine Verringerung der sonst üblichen Raten ab einem bestimmten Zeitpunkt statt.

Es erfolgt keine Änderung im Methodenpapier.

2.22 Würdigung der Stellungnahmen zu Abschnitt 4.9 „Gesundheitsökonomisches Studiendesign“

Es wird in 1 Stellungnahme angemerkt, begleitende gesundheitsökonomische Evaluationen (Piggy-back-Studien) und entscheidungsanalytische Modelle als mögliches gesundheitsökonomisches Studiendesign im Kontext der Fragestellung differenziert zu diskutieren.

Im Methodenpapier wurde ergänzt: „Aufgrund der Möglichkeit, Daten aus unterschiedlichen Datenquellen im Rahmen eines gesundheitsökonomischen Entscheidungsmodells zusammenzuführen (u. a. auch unter Miteinbezug einer Piggy-back-Studie) und dabei auch langfristige Konsequenzen abzubilden, stellen modellbasierte gesundheitsökonomische Evaluationen den Regelfall für die KNB dar.“ Wie in der Stellungnahme richtig anmerkt wird, liegt der Vorteil entscheidungsanalytischer Modelle darin, dass mehrere Datenquellen zusammengeführt werden können. In diesem Kontext können auch Piggy-back-Studien in entscheidungsanalytische Modelle miteinfließen und sind nicht ausgeschlossen. Allerdings wird davon ausgegangen, dass diese in den seltensten Fällen einen ausreichend langen Zeitraum betrachten und die Kostenerhebung auf Deutschland übertragbar ist.

In einigen Stellungnahmen wird gefordert, als weiteres Szenario für die Extrapolation auch das Ansteigen des Effekts nach Studienende zu berücksichtigen.

Im Methodenpapier wird dieses Szenario nicht mit aufgenommen. In der Literatur wird kein vergleichbarer ansteigender Effekt nach Studienende beschrieben. Beispielsweise bei Kühne et al. [19] werden folgende Extrapolationsszenarien genannt: (1) konstanter Behandlungseffekt (entspricht dem optimistischen Szenario im Methodenpapier, z. B. bei kontinuierlichen Therapien), (2) abnehmender Effekt im Zeitverlauf (entspricht dem optimistischen Szenario im Methodenpapier, z. B. bei zeitlich begrenzten Therapien), (3) kein Behandlungseffekt nach Studienende (entspricht dem konservativen Szenario im Methodenpapier) sowie (4) Angleichung an den Behandlungseffekt der zVT, sodass kein Unterschied mehr zwischen den Therapien besteht (entspricht dem konservativen Szenario im Methodenpapier). Da im Methodenpapier nur Mindestanforderungen an die zu berücksichtigenden Szenarien gestellt werden, ist die Betrachtung weiterer Szenarien keineswegs ausgeschlossen. Deren Verwendung ist allerdings jeweils in Abhängigkeit der Datenlage zu beurteilen.

2.23 Würdigung der Stellungnahmen zu Abschnitt 4.10 „Ergebnisdarstellung“

In verschiedenen Stellungnahmen werden die Vorgaben zur Darstellung der ergänzenden Szenarioanalyse zVT-Kosten kritisch thematisiert. Es wird u. a. darauf hingewiesen, dass in den Fällen, in denen die zVT mehrere Therapieoptionen umfasst (ODER-Verknüpfung), nicht davon auszugehen ist, dass diese gleich effektiv sind. Es werden verschiedene Indikationsgebiete benannt, in denen Wirkstoffe bzw. Therapieoptionen der zVT einen Zusatznutzen gegenüber

einem anderen Wirkstoff bzw. einer Therapieoption der zVT aufweisen. Es wird weiterhin argumentiert, dass die Kostenvariation auf der Kostenseite über Sensitivitätsanalysen untersucht werden könne.

Gemäß den Vorgaben der Verfahrensordnung des G-BA erfolgt der Nachweis des Zusatznutzens des zu bewertenden Arzneimittels regelhaft im Vergleich zur vom G-BA bestimmten zweckmäßigen Vergleichstherapie (zVT). Sind mehrere Alternativen für die Vergleichstherapie gleichermaßen zweckmäßig, genügt der Nachweis des Zusatznutzens gegenüber einer dieser Therapieoptionen als Grundlage für den Beschluss zum Zusatznutzen. Die Kosten sind im Dossier zur Nutzenbewertung gemäß § 35a SGB V dagegen für alle vom G-BA als zVT bestimmten Therapieoptionen anzugeben (5. Kapitel 1. Abschnitt § 9 der Verfahrensordnung des G-BA) [8]. Die Darstellung der Kosten erfolgt somit unabhängig von der für den Nachweis des Zusatznutzens gewählten Option der zVT oder der vorliegenden Evidenz für verschiedene Therapieoptionen der zVT im Vergleich untereinander. Vor diesem Hintergrund sind zusätzliche Informationen zur Veränderung des ICERs in Abhängigkeit von der Variation der Kosten von besonderer Bedeutung zur Unterstützung der Erstattungs-betragsverhandlungen. Durch Anpassungen der Formulierungen im Methodenpapier in den Abschnitten 4.3.2 und 4.10 wurden die Anforderungen an die ergänzenden Analysen geschärft. Gefordert wird eine Sensitivitätsanalyse, welche insbesondere die Kosten der günstigsten und der teuersten Therapieoption der zVT berücksichtigt. Die im Entwurf zum Methodenpapier 7.0. vorgeschlagene grafische Darstellung entfällt.

2.24 Würdigung der Stellungnahmen zu Abschnitt 4.14 „Informationsbeschaffung und Datenquellen in der KNB“

In 1 Stellungnahme wird angemerkt, dass unklar ist, was „optionale fokussierte Informationsbeschaffung“ in der Situation bedeutet, wenn keine Bewertungen von Gesundheitszuständen verfügbar sind. Konkret lautet die Frage: „Bedeutet das, dass eine Bewertung der Gesundheitszustände und somit die Durchführung einer Kosten-Nutzwert-Analyse als optional gelten?“

Offenbar liegt hier ein Missverständnis vor. Mit optionaler fokussierter Informationsbeschaffung ist gemeint, dass die Bewertung der Gesundheitszustände vorzugsweise direkt aus den klinischen Studien zu verwenden ist. Wenn diese nicht verfügbar oder erschöpfend sind, wird auf Literatur zurückgegriffen, die mittels fokussierter Informationsbeschaffung identifiziert wurde. Entsprechend steht im Methodenpapier in Abschnitt 4.1.4: „Die Bestimmung der Nutzwerte kann im Rahmen klinischer Studien erfolgen. Alternativ oder ergänzend kann eine fokussierte Informationsbeschaffung nach Nutzwerten für die im Modell betrachteten Gesundheitszustände erforderlich sein.“

2.25 Würdigung der Stellungnahmen zu Abschnitt 7.4 „Auswahl der Evidenz“

In 2 Stellungnahmen wird vorgeschlagen, für den Einschluss einer systematischen Übersichtsarbeit einen Mindestwert für die AMSTAR-2-Bewertung zu definieren. Weiterhin wird in 1 Stellungnahme angeregt, die Zahl der berücksichtigten Übersichtsarbeiten zu einem Thema beispielsweise auf 3 zu beschränken.

Bei der Erstellung von Gesundheitsinformationen werden systematische Übersichtsarbeiten eingeschlossen, wenn die Bewertung nach AMSTAR 2 eine mindestens moderate Ergebnissicherheit (Overall Confidence Rating) ausweist. Das wird ergänzt.

Die Begrenzung der Zahl der berücksichtigten Übersichtsarbeiten zu einem Thema halten wir nicht für sinnvoll, weil diese bezüglich beispielsweise Studienpopulation, Einschlusskriterien und berichteten Endpunkten nicht immer deckungsgleich sind. Der Umgang mit mehreren Übersichtsarbeiten zum gleichen Thema wird bereits in Abschnitt 7.4 beschrieben.

In 1 Stellungnahme wird angeregt, bei fehlenden systematischen Übersichtsarbeiten auch Primärstudien zu berücksichtigen.

Dieser Aspekt wird bereits in Abschnitt 7.3 adressiert.

In 1 Stellungnahme wird angeregt, die Instrumente zur Bewertung des Verzerrungspotenzials von systematischen Übersichtsarbeiten näher zu beschreiben.

Die Bewertung ist Teil der Bewertung mittels AMSTAR 2 (Domänen 7, 8 und 10) und bedürfen unseres Erachtens keiner weiteren Ausführung.

2.26 Würdigung der Stellungnahmen zu Kapitel 8 „Informationsbeschaffung“

In 1 Stellungnahme wird vorgeschlagen, auch die „Specialized Registers“ relevanter themenbezogener Cochrane Review Groups und deren Erfahrungen mit Suchstrategien und Suchstrings zu nutzen.

Die Inhalte der „Specialized Registers“ von Cochrane finden sich in CENTRAL. Diese Datenbank wird i. d. R. in die Informationsbeschaffung nach Studien einbezogen. Es ergibt sich kein Änderungsbedarf für das Methodenpapier.

2.27 Würdigung der Stellungnahmen zu Abschnitt 8.1.1 „Recherche in bibliografischen Datenbanken“

In 1 Stellungnahme wird eine detaillierte Beschreibung der Anforderungen an die Prüfung und ggf. Validierung von Studienfiltern und Klassifikatoren gefordert, um die vom IQWiG zugrunde gelegten Anforderungen an diesen Prozess transparent und nachvollziehbar zu machen.

Laut Methodenpapier ist zu prüfen, ob validierte Suchfilter vorliegen und diese zuverlässig anzuwenden sind. Ob Suchfilter zuverlässig anzuwenden sind, ist projektspezifisch je nach Fragestellung zu prüfen. Anforderungen an Suchfilter finden sich in der Literatur beschrieben. Zur Erläuterung hierfür wurden Literaturquellen im Methodenpapier ergänzt.

In 1 Stellungnahme wird darauf hingewiesen, dass im Methodenpapier unklar bleibt, was mit einem unabhängigen Set relevanter Referenzen gemeint ist und wo dieses herkommt.

Es handelt sich bei dem Set um die anfänglich bekannten Referenzen. Aus diesen werden 2 Sets gebildet (Set zur Entwicklung und Set zur Qualitätssicherung der Suchstrategie). Da der Abschnitt im Methodenpapier Unklarheiten aufwies, wurde er redaktionell überarbeitet.

2.28 Würdigung der Stellungnahmen zu Abschnitt 8.1.2 „Suche in Studienregistern“

In 1 Stellungnahme wird auf einen falsch eingefügten Link hingewiesen; statt auf CTIS wird auf die Suchoberfläche von EU-CTR verwiesen.

Der Hinweis ist korrekt und wurde im Methodenpapier angepasst.

In 1 Stellungnahme wird darauf hingewiesen, dass die Dossievorlage für Modul 4 aufgrund der neuen EMA-Datenbank „Clinical Trials Information System“ (CTIS) angepasst werden muss. Zudem könne ab 2025 eine Suche in EU-CTR entfallen.

Die Anpassung der Dossievorlagen fällt nicht in den Verantwortungsbereich des IQWiG. Wir werden die Information bezüglich der notwendigen Anpassung an die Geschäftsstelle des Gemeinsamen Bundesausschusses weiterleiten. Ob ab 2025 die Suche in EU-CTR entfallen kann, ist unklar, da das Register als Archiv weitergeführt wird. Deshalb erfolgte keine Anpassung im Methodenpapier.

2.29 Würdigung der Stellungnahmen zu Abschnitt 8.2.1 „Recherche nach systematischen Übersichten“

In 1 Stellungnahme wird darauf hingewiesen, dass eine Beschränkung auf maximal drei basissystematische Übersichten (Basis-SÜ) sinnvoll erscheint. Zudem wird darauf hingewiesen, dass es im Folgesatz heißen müsste: „von den verwendeten Basis-SÜ“.

Wir stimmen überein, dass sich die Auswahl auf wenige Basis-SÜ beschränken sollte. Eine Entscheidung bezüglich der Anzahl von Basis-SÜ sollte jedoch projektspezifisch fallen und nicht im Methodenpapier vorab festgelegt werden. Die Änderung des Folgesatzes wurde, wie in der Stellungnahme vorgeschlagen, übernommen.

2.30 Würdigung der Stellungnahmen zu Abschnitt 8.4 „Recherche nach Leitlinien“

In 1 Stellungnahme wird dargestellt, dass es unklar ist, ob die zusätzlich genannten Datenbanken bei der fokussierten Recherche nach Leitlinien ebenfalls fester Bestandteil der umfassenden Leitlinienrecherche sind.

Die bibliografische Datenbank MEDLINE wird in der umfassenden Leitliniensuche nicht durchsucht, ersetzt jedoch in der fokussierten Informationsbeschaffung zum Teil die Suche bei den fachübergreifenden und -spezifischen Leitlinienanbietern der umfassenden Informationsbeschaffung. Die weiteren genannten Datenquellen der fokussierten Informationsbeschaffung sind Bestandteil der umfassenden Leitlinienrecherche. Eine Anpassung im Methodenpapier ist nicht erforderlich.

In 1 Stellungnahme wird darauf hingewiesen, dass die nationalen Versorgungsleitlinien nur über Leitlinien.de aktuell und gültig abrufbar seien. In der Stellungnahme wird dafür plädiert, dass auch diese Webseite an dieser Stelle explizit mit aufgeführt wird.

Die Suche auf AWMF online ist für die fokussierte Informationsbeschaffung ausreichend, um nationale Versorgungsleitlinien zu identifizieren. Jedoch wird darauf geachtet, auf der Originalseite des Leitlinien-Anbieters zu prüfen, ob es die aktuellste Version ist. Eine Anpassung im Methodenpapier ist nicht erforderlich.

In 1 Stellungnahme kommt die Frage auf, welche expliziten Fragestellungen für eine ausschließliche Recherche in MEDLINE qualifiziert wären, da in dieser Datenbank tatsächlich nur ein Bruchteil der verfügbaren Leitlinien überhaupt anzufinden sind.

Es handelt sich um ein Missverständnis, MEDLINE wird immer in Kombination mit den genannten Websites durchsucht. Eine alleinige Suche in MEDLINE ist nicht vorgesehen, sondern wird immer mit anderen Informationsquellen kombiniert. Eine Anpassung im Methodenpapier ist nicht erforderlich.

2.31 Würdigung der Stellungnahmen zu Abschnitt 9.1 „Qualitätsbewertung von Einzelstudien“

In 1 Stellungnahme wird vorgeschlagen, in der Liste von Stellungnahmen, die zur Verbesserung der Qualität von Publikationen erarbeitet wurden, das REporting-of-studies-Conducted-using-Observational-Routinely-collected-health-data(RECORD)-Statement für Beobachtungsstudien mit Routinedaten und das RECORD-PE-Statement für pharmakoepidemiologische Studien zu ergänzen.

Dieser Vorschlag wurde umgesetzt und die Liste um das RECORD- und das RECORD-PE-Statement mit entsprechenden Literaturziten erweitert.

2.32 Würdigung der Stellungnahmen zu Abschnitt 9.1.1 „Kriterien für den Einschluss von Studien“

In mehreren Stellungnahmen wird kritisiert, dass durch die Neuformulierung des 2. Absatzes das genaue Vorgehen bei Erfüllung und Nichterfüllung der 80 %-Regel unklar sei. Insbesondere sei unklar, ob auch Studien, die die 80 %-Regel nicht erfüllen, eingeschlossen werden können, falls gesicherte Informationen vorliegen, dass eine Übertragung der Ergebnisse auf die interessierende Fragestellung trotzdem valide ist.

Das Institut erkennt an, dass der Text präzisiert werden sollte. In der wissenschaftlichen Erörterung am 28.04.2023 wurde das genaue Vorgehen noch einmal erläutert. Der Satz zu den geeigneten Informationen über eine Effektmodifikation soll sich sowohl auf die Situation „80 %-Regel erfüllt“ als auch auf die Situation „80 %-Regel nicht erfüllt“ beziehen. Die inhaltlich einzige Neuerung gegenüber dem bisherigen Vorgehen ist, dass in Einzelfällen auch bei Erfüllung der 80 % die Ergebnisse nicht uneingeschränkt herangezogen werden (siehe auch nächsten Punkt). Der Text dieses Abschnitts wurde präzisiert.

In mehreren Stellungnahmen wird beschrieben, dass eine zusätzliche Betrachtung von Effektmodifikationen bei 80%iger Erfüllung der Einschlusskriterien nicht notwendig und nicht sachgerecht sei.

Studien, die die Einschlusskriterien zu mindestens 80 % erfüllen, werden regelhaft in die Bewertung eingeschlossen. In Einzelfällen liegen jedoch Informationen über Effektmodifikationen durch die betreffenden Einschlusskriterien vor. Es kann z. B. – wie in der wissenschaftlichen Erörterung am 28.04.2023 auch erläutert – der Fall vorliegen, dass gerade die Patientinnen und Patienten, die die Einschlusskriterien nicht erfüllen, für einen Großteil der Ereignisse eines Endpunkts verantwortlich sind. Hier behält sich das Institut vor, trotz Erfüllung der 80 %-Regel die Ergebnisse nicht uneingeschränkt einzubeziehen.

In mehreren Stellungnahmen wird kritisiert, dass der Satz „Falls geeignete Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium (Population bzw. Interventionen) vorliegen, muss über die Berücksichtigung der Studie anhand der Stärke der Effektmodifikation und des Anteils der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen, bzw. des Grades der Abweichung der Interventionen entschieden werden.“ zu vage sei und näher spezifiziert werden solle.

Die Frage, was geeignete Informationen über das Vorhandensein oder Nichtvorhandensein von Effektmodifikationen sind, hängt sehr stark vom Kontext und von vielen Faktoren ab. Die Stärke einer möglichen Effektmodifikation muss gegen den Erfüllungsgrad der Einschlusskriterien – qualitativ und quantitativ – abgewogen werden. Eine weitere Spezifizierung ist in allgemeiner Form nicht möglich. Das Institut wird bei Anwendung dieses

Sachverhalts in seinen Bewertungen jeweils immer ausführlich begründen und darlegen, welche Informationen wie zur Entscheidungsfindung beigetragen haben.

2.33 Würdigung der Stellungnahmen zu Abschnitt 9.1.3 „Rangordnung verschiedener Studienarten / Evidenzgrade“

In 1 Stellungnahme wird angeregt, in Abschnitt 9.1.3 eine Einschätzung zu GRADE-Bewertungen für nicht randomisierte Studien mithilfe von ROBINS-I zu ergänzen.

Der kurze Abschnitt 9.1.3 dient nur einer groben Orientierung ohne Einschätzungen zu konkreten Bewertungsinstrumenten. Details zur Bewertung nicht randomisierter Studien finden sich in anderen Abschnitten der „Allgemeinen Methoden“ (Abschnitte 3.8, 9.1.4, 9.3.6).

2.34 Würdigung der Stellungnahmen zu Abschnitt 9.3.2 „Statistische Signifikanztests und Konfidenzintervalle“

In 1 Stellungnahme werden die Kürzungen und Umformulierungen in Abschnitt 9.3.2 begrüßt. In mehreren Stellungnahmen wird dagegen betont, dass für ergänzende Analysen oder Sensitivitätsanalysen eine Adjustierung für multiples Testen nicht erforderlich ist. Es wird vorgeschlagen, die Verwendung von Adjustierungen für multiples Testen genauer zu spezifizieren und mit Beispielen zu erläutern.

Offensichtlich gibt es hier Befürchtungen, dass das Institut zukünftig für ergänzende Analysen oder Sensitivitätsanalysen die Anwendung von Adjustierungen für multiples Testen plant. In der wissenschaftlichen Erörterung am 28.04.2023 wurde festgestellt, dass dies eine falsche Interpretation des 2. Teils des Satzes „Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an [59], z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.“ aus dem Entwurf für die Version 7.0 der „Allgemeinen Methoden“ vom 06.12.2022 darstellt. Um dieses Missverständnis zu vermeiden, wurde der Halbsatz wieder umformuliert und die ursprüngliche Formulierung aus der Version 6.1 der „Allgemeinen Methoden“ vom 24.01.2022 mit ergänzender Erläuterung wiederhergestellt: „Im Rahmen der Nutzenbewertung (siehe Abschnitt 3.1) versucht das Institut, den Fehler 1. Art für jeden einzelnen Nutzenaspekt getrennt zu kontrollieren, indem beispielsweise bei Überlebenszeitdaten die zeitübergreifende Analyse primär herangezogen wird.“

2.35 Würdigung der Stellungnahmen zu Abschnitt 9.3.6 „Adjustierung und multifaktorielle Verfahren“

In 1 Stellungnahme wird darauf hingewiesen, dass bei anwendungsbegleitenden Datenerhebungen (AbDs) unterschiedliche Erhebungszeitpunkte zum Arzneimittel und zur zweckmäßigen Vergleichstherapie problematisch sein können.

Es wurde 1 Satz in den „Allgemeinen Methoden“ ergänzt, dass bei AbDs unterschiedliche Erhebungszeitpunkte beachtet werden müssen.

In mehreren Stellungnahmen wird kritisiert, dass vom Institut zur Identifikation von Confoundern nur das Verfahren nach Pufulete et al. [23] akzeptiert wird und als Verfahren zur Confounderadjustierung nur die Propensity-Score-Methode.

Es ist nicht richtig, dass das Institut zur Identifikation von Confoundern nur das Verfahren nach Pufulete et al. [23] akzeptiert und als Verfahren zur Confounderadjustierung nur die Propensity-Score-Methode. Es wird in den Allgemeinen Methoden als eine Möglichkeit zur Confounderidentifikation das Verfahren nach Pufulete et al. [23] zitiert. Bezüglich der Verfahren zur Confounderadjustierung werden neben der Propensity-Score-Methode explizit auch andere Verfahren wie multifaktorielle Regressionsmodelle und Instrumentenvariablen genannt. Es besteht daher kein Änderungsbedarf.

In mehreren Stellungnahmen wird gefordert, dass das Institut für die Durchführung von AbDs und die verschiedenen Schritte der Datenanalyse ganz konkrete Methoden und Kriterien inklusive Schwellenwerte, z. B. für die Überprüfung von Positivität, Überlappung und Balanciertheit, vorgeben und beschreiben soll.

Die „Allgemeinen Methoden“ haben nicht den Charakter eines Statistik-Lehrbuchs. Wie in der Präambel dargelegt, sollen die „Allgemeinen Methoden“ 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. Die „Allgemeinen Methoden“ stellen daher einen Rahmen dar, der im konkreten Einzelfall projektspezifisch ausgefüllt werden muss. Es besteht daher kein Änderungsbedarf.

In 1 Stellungnahme wird darauf hingewiesen, dass relevante Confounder in den zur Verfügung stehenden Registern möglicherweise nicht erhoben wurden, und es wird gefordert, dass sich die Confounderadjustierung an den tatsächlich erhobenen Confoundern orientieren sollte.

Es ist richtig, dass relevante Confounder in den zur Verfügung stehenden Registern möglicherweise nicht erhoben wurden. Eine adäquate Confounderadjustierung ist dann aber nicht möglich und somit auch keine adäquate Effektschätzung. Es besteht daher kein Änderungsbedarf.

In 1 Stellungnahme wird darauf hingewiesen, dass durch die Verwendung des Begriffs „Interventionen“ bei der Erläuterung von Positivität nicht interventionelle Studien bei AbDs nicht ausdrücklich miteingeschlossen werden.

Um klarzustellen, dass hier nicht interventionelle Studien nicht ausgeschlossen werden sollen, wurde der Begriff „Interventionen“ durch „Behandlungsformen“ ersetzt.

In mehreren Stellungnahmen wird kritisiert, dass es einen Widerspruch gibt zwischen Abschnitt 9.3.8 zu indirekten Vergleichen und Abschnitt 9.3.6 zu multifaktoriellen Verfahren.

In der wissenschaftlichen Erörterung am 28.04.2023 wurde festgestellt, dass dieser angebliche Widerspruch darauf basiert, dass in Abschnitt 9.3.8 indirekte Vergleiche ohne Brückenkomparator, d. h. indirekte Vergleiche bei unverbundenen Netzwerken, abgelehnt und in Abschnitt 9.3.6 in Zusammenhang mit AbDs aber nicht randomisierte Vergleiche akzeptiert werden. Bei AbDs kommt ja insbesondere die Propensity-Score-Methode infrage; diese Methode könnte auch bei unverbundenen Netzwerken angewendet werden. Die Erklärung für diese unterschiedliche Bewertung in den Abschnitten 9.3.8 und 9.3.6 ist, dass in Abschnitt 9.3.8 typische indirekte Vergleiche gemeint sind, bei denen zumindest zum Teil nur aggregierte Daten vorhanden sind, in Abschnitt 9.3.6 bei AbDs aber vorausgesetzt wird, dass vollständige IPD-Information vorliegt. Nur bei vollständiger IPD-Information ist es möglich, bei unverbundenen Netzwerken mithilfe von multifaktoriellen Verfahren eine ausreichende Confounderkontrolle zu erreichen. Es wurden entsprechende Erläuterungen in den Abschnitten 9.3.6 und 9.3.8 der „Allgemeinen Methoden“ ergänzt.

In 1 Stellungnahme wird darauf hingewiesen, dass eine adäquate Confounderadjustierung nur bei ausreichend großen Fallzahlen möglich ist, und es wird eine Lösung für AbDs bei kleinen Fallzahlen gefordert.

Es ist richtig, dass eine adäquate Confounderadjustierung nur bei ausreichend großen Fallzahlen möglich ist. Eine adäquate Effektschätzung wiederum ist nur bei einer ausreichenden Confounderadjustierung möglich. Daher kann eine AbD nur bei ausreichend hohen Fallzahlen erfolgreich sein. Was „ausreichend hoch“ bedeutet, ist von Fall zu Fall unterschiedlich. In der Regel dürfte jedoch eine Fallzahl unter 100 nicht ausreichend sein. Es besteht kein Änderungsbedarf.

2.36 Würdigung der Stellungnahmen zu Abschnitt 9.3.7 „Metaanalysen“

In 1 Stellungnahme wird darauf hingewiesen, dass die Verwendung des I^2 -Maßes zunehmend in die Kritik geraten ist, und es wird empfohlen, Vorschläge zur Heterogenitätseinteilung auf Basis des Heterogenitätsparameters τ zu ergänzen.

In Abschnitt 9.3.7 B der „Allgemeinen Methoden“ wurden die Limitationen des I^2 -Maßes ergänzt sowie ein Vorschlag zur Heterogenitätseinteilung auf Basis des Heterogenitätsparameters τ mit entsprechenden Literaturzitaten.

In mehreren Stellungnahmen wird der Wunsch geäußert, dass das Institut konkretere Empfehlungen zur Anwendung bayesscher Verfahren macht, z. B. zu A-priori-Verteilungen für den Heterogenitätsparameter τ in bayesschen Metaanalysen mit zufälligen Effekten.

Das Institut führt zurzeit Untersuchungen zur Bestimmung von A-priori-Verteilungen für den Heterogenitätsparameter τ in bayesschen Metaanalysen mit zufälligen Effekten durch. Weitere Details hierzu werden in zukünftigen Versionen der „Allgemeinen Methoden“ enthalten sein.

2.37 Würdigung der Stellungnahmen zu Abschnitt 9.3.8 „Indirekte Vergleiche“

Bei nicht erfüllter Konsistenzannahme für einen indirekten Vergleich kann kein Beleg für einen Nutzen oder Schaden abgeleitet werden. In 1 Stellungnahme wird darauf hingewiesen, dass es gekennzeichnet werden muss, wenn die Konsistenzannahme aufgrund fehlender Informationen als nicht erfüllt bewertet werden muss.

Es ist richtig, dass unterschieden werden muss, ob Evidenz dafür vorhanden ist, dass es keinen Nutzen gibt, oder ob man nicht auf einen Nutzen schließen kann, da die Evidenzlage unzureichend ist. Dies ist aber kein spezielles Problem von indirekten Vergleichen, sondern gilt für die Nutzenbewertung ganz allgemein. Es ergibt sich daher kein Änderungsbedarf.

In mehreren Stellungnahmen wird kritisiert, dass es einen Widerspruch gibt zwischen Abschnitt 9.3.8 zu indirekten Vergleichen und Abschnitt 9.3.6 zu multifaktoriellen Verfahren.

Zu diesem Punkt wird auf die Würdigung in Abschnitt 2.35 verwiesen.

2.38 Würdigung der Stellungnahmen zu Abschnitt 9.3.12 „Umgang mit variablen Beobachtungszeiten“

In mehreren Stellungnahmen wird eine genauere Differenzierung des Sachverhalts unterschiedlicher Beobachtungsdauern versus verkürzte Beobachtungsdauern angeregt und empfohlen, dass Situationen berücksichtigt werden, in denen trotz verkürzter Beobachtung vergleichbare Beobachtungsdauern zwischen den Studienarmen vorliegen.

Dem Vorschlag folgend wurden Konkretisierungen vorgenommen, die deutlich machen, welche Analysen im Fall verkürzter, aber nicht relevant unterschiedlicher Beobachtungszeiten adäquat sein können.

In mehreren Stellungnahmen wird infrage gestellt, dass eine einheitliche Erhebungsdauer für alle Endpunkte in einer Studie bzw. die Verwendung des Treatment-Policy-Estimand für die Fragestellung der Nutzenbewertung erforderlich / zielführend ist.

Die Berücksichtigung von patientenberichteten Endpunkten (PROs) nur während der Behandlungsphase widerspricht dem Intention-to-treat-Prinzip und führt zu systematischen Verzerrungen. In der Nutzenbewertung ist – wie für alle anderen patientenrelevanten Endpunkte – auch für PROs der Treatment-Policy-Estimand ausschlaggebend. Die Berücksichtigung von PROs auch nach Behandlungsabbruch oder -wechsel ist grundsätzlich sinnvoll und notwendig, um Aussagen zu längerfristigen Auswirkungen neuer Therapien auf Symptomatik und Lebensqualität machen zu können [13]. Eine Änderung / Ergänzung im Methodenpapier wird als nicht notwendig erachtet.

In mehreren Stellungnahmen wird angeführt, dass die Erhebung von PROs nach Behandlungsabbruch in der Realität nur schwer möglich bzw. nicht durchführbar sei. Dabei wird insbesondere auch mit der Belastung der Patienten durch die PRO-Erhebung argumentiert.

Aus den Stellungnahmen lassen sich keine Quellen identifizieren, die dieses Argument stützen. Vielmehr beschreibt eine mit den Stellungnahmen von Abbvie und der DKG vorgelegte Publikation [2] Maßnahmen, die dazu beitragen können, patientenberichtete Endpunkte über einen längeren Zeitraum zu erheben und gleichzeitig die Belastung der Patientinnen und Patienten zu reduzieren (z. B. durch eine klar definierte/kommunizierte Rationale, Einsatz von ePROM, durchdachte Wahl / Anzahl der Erhebungszeitpunkte und -instrumente). Dabei wird betont, dass die Befürchtung, dass Patienten oder Patientinnen per se z. B. zu krank sind, um Fragebögen auszufüllen, nicht durch Evidenz belegt ist, und die daraus folgende Zurückhaltung gegenüber PRO-Erhebungen kontraproduktiv für die Generierung von (notwendiger) Evidenz ist. Auch in der wissenschaftlichen Erörterung der Stellungnahmen zum Entwurf der Allgemeinen Methoden 7.0 wurden keine Daten vorgelegt bzw. berichtet, die zeigen, dass eine Erhebung von PROs nach Behandlungsabbruch oder -wechsel wegen zu hoher Belastung der Patientinnen und Patienten nicht möglich sei. Eine Änderung / Ergänzung im Methodenpapier wird als nicht notwendig erachtet.

In mehreren Stellungnahmen wird eine Konkretisierung dahin gehend, was mit „systematisch verkürzter Beobachtungsdauer“ gemeint ist, empfohlen, zudem wurde die Frage gestellt, wie mit PRO-Daten umzugehen ist, die mortalitätsbedingt unterschiedlich lang erhoben und beobachtet worden sind.

Grundsätzlich ist eine Erhebung der PRO-Daten – wie für andere Endpunkte auch – über den gesamten Studienverlauf erforderlich. Eine systematische Verkürzung liegt dann vor, wenn diese Erhebung geplant vorzeitig (d. h. vor dem Studienende) beendet wird, z. B. durch ein vorzeitiges Erhebungsende bei Progression der Erkrankung, Behandlungsabbruch oder -wechsel. Die Ausführung im Methodenpapier, dass Auswertungen zu einer „dauerhaften Veränderung“ im Falle sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauern in der Regel nicht für die Nutzenbewertung herangezogen werden,

bezieht sich nicht auf unterschiedliche Beobachtungsdauern, die (bei Erhebung bis zum Studienende) mortalitätsbedingt entstehen können. Sofern diese Konstellation eintritt, ist dies bei der Interpretation der Daten zu berücksichtigen.

2.39 Literaturverzeichnis

1. SGB V Handbuch; Sozialgesetzbuch V; Krankenversicherung. Altötting: KKF-Verlag; 2020.
2. Aiyegbusi OL, Roydhouse J, Rivera SC et al. Key considerations to reduce or address respondent burden in patient-reported outcome (PRO) data collection. Nat Commun 2022; 13(1): 6026. <https://dx.doi.org/10.1038/s41467-022-33826-4>.
3. Attema AE, Brouwer WBF, Claxton K. Discounting in Economic Evaluations. Pharmacoeconomics 2018; 36(7): 745-758. <https://dx.doi.org/10.1007/s40273-018-0672-z>.
4. Bouvy JC, Cowie L, Lovett R et al. Use of Patient Preference Studies in HTA Decision Making: A NICE Perspective. Patient 2020; 13(2): 145-149. <https://dx.doi.org/10.1007/s40271-019-00408-4>.
5. Edlin R, McCabe C, Hulme C et al. Cost Effectiveness Modelling for Health Technology Assessment: A Practical Course. Cham: Adis; 2015.
6. European Medicines Agency. Qualification Opinion of IMI PREFER [online]. 2022 [Zugriff: 14.08.2023]. URL: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-imi-prefer_en.pdf.
7. Gamper EM, King MT, Norman R et al. The EORTC QLU-C10D discrete choice experiment for cancer patients: a first step towards patient utility weights. J Patient Rep Outcomes 2022; 6(1): 42. <https://dx.doi.org/10.1186/s41687-022-00430-5>.
8. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. URL: <https://www.g-ba.de/informationen/richtlinien/42/>.
9. Gemeinsamer Bundesausschuss. Mutterschafts-Richtlinien: Nichtinvasive Pränataldiagnostik zur Bestimmung des fetalen Rhesusfaktors [online]. 2020 [Zugriff: 14.08.2023]. URL: <https://www.g-ba.de/beschluesse/4439/>.
10. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Onasemnogen-Abeparvovec (spinale Muskelatrophie); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen [online]. 2021 [Zugriff: 14.08.2023]. URL: https://www.g-ba.de/downloads/39-261-4702/2021-02-04_AM-RL-XII_awD_Onasemnogen-Abeparvovec_D-549_BAnz_WZ.pdf.

11. Gongora-Salazar P, Rocks S, Fahr P et al. The Use of Multicriteria Decision Analysis to Support Decision Making in Healthcare: An Updated Systematic Literature Review. *Value Health* 2023; 26(5): 780-790. <https://dx.doi.org/10.1016/j.ival.2022.11.007>.
12. Greenhalgh T, Fisman D, Cane DJ et al. Adapt or die: how the pandemic made the shift from EBM to EBM+ more urgent. *BMJ Evid Based Med* 2022; 27(5): 253-260. <https://dx.doi.org/10.1136/bmjebm-2022-111952>.
13. Haslam A, Herrera-Perez D, Gill J et al. Patient Experience Captured by Quality-of-Life Measurement in Oncology Clinical Trials. *JAMA Netw Open* 2020; 3(3): e200363. <https://dx.doi.org/10.1001/jamanetworkopen.2020.0363>.
14. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzept für eine anwendungsbegleitende Datenerhebung – Fedratinib; Rapid Report [online]. 2022 [Zugriff: 11.07.2023]. URL: https://www.iqwig.de/download/a21-142_anwendungsbegleitende-datenerhebung-fedratinib_rapid-report_v1-0.pdf.
15. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Brexucabtagen Autoleucler (rezidierte oder refraktäre B-Zell-Vorläufer ALL); Bewertung gemäß § 35a SGB V; AbD-Konzept [online]. 2023 [Zugriff: 31.07.2023]. URL: https://www.iqwig.de/download/a22-118_brexucabtagen-autoleucler_abd-konzept_v1-0.pdf.
16. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Etranacogen Dezaparvovec (Hämophilie B) – Bewertung gemäß § 35a SGB V; AbD-Konzept [online]. 2023 [Zugriff: 11.07.2023]. URL: https://www.iqwig.de/download/a22-83_etranacogen-dezaparvovec-haemophilie-b_abd-konzept_v1-1.pdf.
17. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzept für eine anwendungsbegleitende Datenerhebung – Valoctocogen Roxaparvovec; Rapid Report [online]. 2023 [Zugriff: 11.07.2023]. URL: https://www.iqwig.de/download/a22-20_anwendungsbegleitende-datenerhebung-valoctocogen-roxaparvovec_rapid-report_v1-0.pdf.
18. Janssens R, Russo S, van Overbeeke E et al. Patient Preferences in the Medical Product Life Cycle: What do Stakeholders Think? Semi-Structured Qualitative Interviews in Europe and the USA. *Patient* 2019; 12(5): 513-526. <https://dx.doi.org/10.1007/s40271-019-00367-w>.
19. Kühne F, Schomaker M, Stojkov I et al. Causal evidence in health decision making: methodological approaches of causal inference and health decision science. *Ger Med Sci* 2022; 20: Doc12. <https://dx.doi.org/10.3205/000314>.
20. McInnes MDF, Moher D, Thombs BD et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA* 2018; 319(4): 388-396. <https://dx.doi.org/10.1001/jama.2017.19163>.

21. O'Mahony JF, Paulden M, McCabe C. NICE's Discounting Review: Clear Thinking on Rational Revision Meets Obstacle of Industrial Interests. *Pharmacoeconomics* 2021; 39(2): 139-146. <https://dx.doi.org/10.1007/s40273-020-00990-8>.
22. Pearson I, Rothwell B, Olaye A et al. Economic Modeling Considerations for Rare Diseases. *Value Health* 2018; 21(5): 515-524. <https://dx.doi.org/10.1016/j.jval.2018.02.008>.
23. Pufulete M, Mahadevan K, Johnson TW et al. Confounders and co-interventions identified in non-randomized studies of interventions. *J Clin Epidemiol* 2022; 148(148): 115-123. <https://dx.doi.org/10.1016/j.jclinepi.2022.03.018>.

3 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll

3.1 Teilnehmerliste der wissenschaftlichen Erörterung

Name	Organisation / Institution / Firma / privat
Bauer, Robert	Daiichi Sankyo Deutschland GmbH
Böhme, Sarah	Pfizer Pharma GmbH
Büttner, Reinhard	Nationales Netzwerk Genomische Medizin, Lungenkrebs (nNGM)
Caruso, Simone	Bayer Vital GmbH
Denzer-Lippmann, Melanie	Novartis Pharma GmbH
Dintsios, Ch.-Markos	Bayer Vital GmbH
Eisele, Lewin	Janssen-Cilag GmbH
Ellis, Aneurin	Bristol-Myers Squibb GmbH & Co. KGaA
Engelking, Franziska	Boehringer Ingelheim Pharma GmbH & Co. KG
Fecker, Julian	Roche Pharma GmbH
Friede, Tim	Deutsche Gesellschaft für Kardiologie – Herz- und Kreislaufforschung e. V./ Deutsche Region der Internationalen Biometrischen Gesellschaft (IBS-DR)/ Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e. V.
Friederich, Andrea	Amgen GmbH
Gaupel, Ann-Christin	Merck Healthcare Germany GmbH
Heuck, Alexander	AbbVie Deutschland GmbH & Co. KG
Hogger, Sara	AMS Advanced Medical Services GmbH
Hussein, Rugzan Jameel	Kassenzahnärztliche Bundesvereinigung
Kucka, Annett	Ipsen Pharma GmbH
Kumke, Thomas	UCB Pharma GmbH
Lambertz, Lisa	Novartis Pharma GmbH
Loske, Sonja	SmartStep Consulting GmbH
Musiolik, Katharina	Ecker + Ecker GmbH
Neumeyer-Gromen, Angela	Deutsche Krankenhausgesellschaft e. V.
Rasch, Andrej	Verband Forschender Arzneimittelhersteller e. V.
Rasokat, Anna	Nationales Netzwerk Genomische Medizin, Lungenkrebs (nNGM)
Rauchensteiner, Stephan	Pfizer Pharma GmbH
Richter, Swantje	AMS Advanced Medical Services GmbH
Sauerbruch, Tilman	Deutsche Gesellschaft für Innere Medizin e. V.
Schaser, Vanessa	AbbVie Deutschland GmbH & Co. KG
Schepers, Cornelia	Boehringer Ingelheim Pharma GmbH & Co. KG
Schrom, Eva	Amgen GmbH
Sebastian, Martin	Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e. V.
Spranger, Anne	Bundesverband Medizintechnologie e. V.

Name	Organisation / Institution / Firma / privat
Stausberg, Jürgen	Deutsches Netzwerk Versorgungsforschung e. V.
Wille, Hans	Arzneimittelkommission der deutschen Ärzteschaft
Wolf, Jürgen	Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e. V.
Kuss, Oliver	wissenschaftlicher Beirat
Rörtgen, Thilo	Protokollant
Bender, Ralf	IQWiG
Flintrop, Jens	IQWiG
Kaiser, Thomas	IQWiG (Moderation)
Koch, Klaus	IQWiG
Kunigkeit, Christopher	IQWiG
Lange, Stefan	IQWiG
Mostardt, Sarah	IQWiG
Potthast, Regine	IQWiG
Preukschat, Daniela	IQWiG
Sauerland, Stefan	IQWiG
Schaubert, Kristina	IQWiG
Schwalm, Anja	IQWiG
Selbach, Claudia	IQWiG
Skipka, Guido	IQWiG
Thomas, Stefanie	IQWiG
Vervölgyi, Volker	IQWiG
Waffenschmidt, Siw	IQWiG

3.2 Liste der Stellungnahmen, zu denen keine Vertreterin bzw. 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
Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF)
Arbeitsgruppe Epidemiologische Methoden der DGEpi, GMDS & DGSMF für die Deutsche Gesellschaft für Epidemiologie (DGEpi e.V.)
co.value®
Deutschen Gesellschaft für Gesundheitsökonomie e. V. (dggö)
Deutschen Gesellschaft für Orthopädie und Unfallchirurgie (DGOU), der Deutschen Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC) und der Deutschen Gesellschaft für Unfallchirurgie (DGU)
Deutschen Gesellschaft für Pharmazeutische Medizin e. V. (DGPharMed)
GlaxoSmithKline GmbH & Co. KG (GSK)
Institute of Public Health, Medical Decision Making and Health Technology Assessment, UM-IT Tirol
Maimon Research LLC
MSD Sharp & Dohme GmbH
Novo Nordisk Pharma GmbH
Sanofi-Aventis Deutschland GmbH

3.3 Tagesordnung der wissenschaftlichen Erörterung

	Begrüßung und Einleitung
TOP 1	Anwendungsbegleitende Datenerhebungen
TOP 2	Kein Zusatznutzen auf der Basis von KNBs
TOP 3	Besonderheiten von Orphan Drugs in KNBs
TOP 4	80 %-Regel
TOP 5	Umgang mit variablen Beobachtungszeiten
TOP 6	Signifikanz und multiples Testen
TOP 7	Verschiedenes

3.4 Protokoll der wissenschaftlichen Erörterung

Datum: 28.04.2023, 12:00 bis 14:10 Uhr

Ort: KOMED im MediaPark GmbH,
Im Mediapark 7, 50670 Köln

Moderation: Thomas Kaiser

3.4.1 Begrüßung und Einleitung

Moderator Thomas Kaiser: Es freut mich sehr, dass wir heute hier in diesem schönen Raum zusammen sind. Aufgrund der Menge der Stellungnahmen, die dann natürlich auch zu einer Menge von Personen führt, haben wir uns dazu entschieden, das in diesem Raum zu machen, einfach weil das besser möglich ist, wir ein bisschen mehr Nähe haben. Wundern Sie sich nicht darüber, dass die einen vorne und die anderen hinten sitzen, das ist alleine dem Alphabet geschuldet innerhalb Ihrer Organisationen. Ob das dazu führt, dass sich einzelne Firmen oder Organisationen zukünftig umbenennen werden, lasse ich mal dahingestellt. Aber nichts desto trotz, vielen Dank auch für das Interesse an dem Methodenpapier, das Sie ja auch durch Ihre Stellungnahmen und Ihre Anwesenheit zeigen. Es ist uns ja auch klar, dass das ein Methoden- und ein Verfahrensinteresse ist, weil viele von Ihnen davon ja mittel- oder unmittelbar betroffen sind.

Wir haben in dem Methodenpapier verschiedene Punkte aufgegriffen. Wir haben zum Beispiel Dinge ändern müssen, ändern sollen, die sich durch neue Aufgaben ergeben haben, anmeldungsbegleitende Datenerhebung zum Beispiel. Wir haben neue wissenschaftliche Erkenntnisse, die sich zum großen Teil auch aus unseren eigenen Bewertungen ergeben haben. Da spreche ich mal unterschiedliche Beobachtungsdauern an. Und natürlich haben wir uns auch - dazu haben einige von Ihnen auch Stellung genommen - jetzt noch mal neu der Kosten-Nutzen-Bewertung gewidmet, einfach um dieses Verfahren gangbarer zu machen. Deswegen sind natürlich Ihre Stellungnahmen auch entsprechend vielfältig und reichhaltig.

Bevor wir starten, möchte ich ein paar Vorbemerkungen machen. Diese Vorbemerkungen kommen Ihnen bekannt vor, wenn Sie schon mal an einer Erörterung des IQWiG entweder zum Methodenpapier oder zu anderen Themen teilgenommen haben.

Der Grundsatz dieser Erörterung ist, dass wir offene Fragen besprechen möchten. Daraus hat sich dann ja auch die Ihnen übermittelte Tagesordnung ergeben. Das bedeutet auch, Sie müssen Ihre Stellungnahmen, die wir sehr sorgfältig gelesen haben, nicht noch mal bei Ihren Beiträgen im Detail wiederholen. Natürlich steht es Ihnen frei, wenn wir bestimmte Punkte diskutieren, dass Sie darauf dann auch noch mal eingehen, da, wo Ihnen das wichtig ist.

Ein anderer wichtiger Punkt ist, dass wir in der Anhörung zum Methodenpapier und damit auch in dieser Erörterung ausschließlich diejenigen Punkte diskutieren, die sich geändert haben im Vergleich zur vorhergehenden Version. Wir werden nicht noch mal neu Dinge aufgreifen, die eigentlich schon manifest sind, zum Beispiel in der Version 6.1.

Wichtig ist für Sie auch, dass die Erörterung aufgezeichnet wird. Wir haben einen Protokollanten. Dieser Protokollant wird durch eine Audioaufnahme unterstützt. Also, alle Ihre Redebeiträge werden aufgenommen. Sie werden auch dann als Wortprotokoll mit Ihrem Namen veröffentlicht. Also, Ihr Redebeitrag wird Ihnen persönlich zugeordnet werden

können. Sie sind darüber informiert worden, dass Sie mit Ihrer Teilnahme dem zustimmen. Jetzt ist gerade die letzte Gelegenheit, die Zustimmung zurückzuziehen durch Verlassen des Raumes. Das sehe ich aber an keiner Stelle. Insofern haben wir das noch mal bestätigt.

Und weil wir das eben aufzeichnen, möchte ich Sie bitten, auch um dem Protokollanten das zu erleichtern, dass Sie jedes Mal, wenn Sie einen Redebeitrag machen, noch mal vorher Ihren Namen nennen, auch wenn Sie schon zum zweiten oder dritten Mal dran sind, damit wir das auf jeden Fall gut zuordnen können.

Gut, wunderbar. Nachdem wir diese Dinge, denke ich, geklärt haben, können wir auch schon in die Tagesordnung einsteigen. Unter TOP 7 wird es auf jeden Fall Möglichkeit geben, Punkte von Ihrer Seite noch mal anzusprechen. Aber von unserer Seite gibt es eben jetzt diese sechs Punkte.

Wir steigen ein mit

3.4.2 Tagesordnungspunkt 1: Anwendungsbegleitende Datenerhebung

Moderator Thomas Kaiser: Das ist etwas Neues im Gesetz seit 2019, § 35a Abs. 3b SGB V. Das wird von Volker Vervölgyi vom Ressort Arzneimittelbewertung eingeführt.

Volker Vervölgyi: Vielen Dank, Thomas. Auch schönen guten Morgen von mir an dieser Stelle.

Ich möchte mit dem spezifischen Punkt anfangen, nämlich der systematischen Confounder-Identifikation, die ja für die AbD eine wesentliche Bedeutung hat. Mehrere Stellungnehmende haben kritisiert, dass das Verfahren, das Pufulete et al. letztes Jahr im Journal of Clinical Epidemiology vorgeschlagen haben, das einzige Verfahren sei, das in der AbD oder insgesamt akzeptiert werden würde. Da möchte ich zunächst mal sagen, dass es da ein Missverständnis gibt, weil weder die Formulierung im Methodenpapier noch in den AbD-Konzepten, die wir bisher veröffentlicht haben, so gemeint ist, dass es zwangsläufig genauso gemacht werden solle, müsse, wie das bei Pufulete et al. beschrieben ist. Und zwar sind verschiedene Verfahren denkbar, die Confounder systematisch und vollständig zu identifizieren. Es muss die Voraussetzung erfüllt sein, dass alle relevanten Confounder identifiziert werden einschließlich wichtiger Wechselwirkungen. Solange diese Voraussetzungen erfüllt sind, sind natürlich mehrere Wege denkbar.

Zum anderen ist es aber so - das haben wir auch selber gesehen, als wir uns mit dem Thema auseinandergesetzt haben -, dass insgesamt der Themenbereich „systematische Confounder-Identifikationen“ noch nicht hinreichend erforscht ist. Aber das ist wichtig und auch essenziell insbesondere für die AbD, wo wir prospektiv Daten erheben wollen, einfach um zu gewährleisten, dass die Daten so generiert werden und ausgewertet werden können, dass auch nachher Schlussfolgerungen daraus gezogen werden können. Wir haben auch in den

vergangenen Verfahren zu Dossierbewertungen gesehen, dass genau dies einer der wesentlichen Punkte war, die dazu geführt haben, dass die Daten nicht verwertbar waren.

Unsere Frage an Sie wäre jetzt, ob Sie weitere veröffentlichte Verfahren kennen, die vorgeschlagen haben, systematisch Confounder zu identifizieren und vollständig zu identifizieren. Das ist unsere Frage.

Moderator Thomas Kaiser: Sehe ich eine Wortmeldung? Viele von Ihnen oder einige von Ihnen haben das Thema ja aufgebracht. Man kann durchaus natürlich sagen: Ist das denn das Einzige Verfahren? Die Frage ist: Welche gibt es noch? - Das scheint nicht der Fall zu sein.

Dann ist, würde ich sagen, Volker, die Frage damit beantwortet.

Haben wir noch eine Frage zum Thema „AbD“? Ralf Bender bitte.

Ralf Bender: Schönen guten Morgen auch von meiner Seite. Wenn wir Confounder-Identifikation jetzt schon verlassen, hätten wir noch einen zweiten Punkt zu AbDs.

Neben dem neuen Abschnitt 3.3.4 haben wir auch den Abschnitt 9.3.6 zur Adjustierung und multifaktoriellen Verfahren erweitert im Hinblick auf AbDs. Es gibt einige Stellungnahmen, die meinen, einen Widerspruch entdeckt zu haben zwischen dem neuen, erweiterten Abschnitt 9.3.6 und dem alten, nicht veränderten Abschnitt 9.3.8 zu indirekten Vergleichen. Bevor wir uns zu diesen Stellungnahmen, zu diesem Widerspruch äußern, hätte ich sicherheitshalber die Frage, worauf dieser Widerspruch basiert. Ganz konkret: Basiert dieser Widerspruch darauf, dass im Abschnitt 9.3.8 indirekte Vergleiche ohne Brückenkomparator, also indirekte Vergleiche bei unverbundenen Netzwerken, abgelehnt werden, dagegen im Abschnitt 9.3.6 im Zusammenhang mit AbDs nicht randomisierte Vergleiche doch akzeptiert werden? Auf den ersten Blick könnte das tatsächlich ein Widerspruch sein, weil nicht randomisierte Vergleiche zum Beispiel mit Propensity Scores ausgewertet werden können. Dieses Verfahren der Propensity Scores könnte man auch bei indirekten Vergleichen bei unverbundenen Netzwerken anwenden. Das akzeptieren wir aber nicht. Also, ist das der Widerspruch, der hier kritisiert wird? Die Frage geht insbesondere an Ecker + Ecker und an Pfizer.

Moderator Thomas Kaiser: Da hinten sehe ich eine Hand, die aus der Richtung Pfizer kommt, und dann sehe ich eine, die aus der Richtung Ecker + Ecker kommt. Dann machen wir das in der Reihenfolge. - Bitte sehr.

Sarah Böhme: Zur Frage: Es ist dieser Widerspruch gemeint. Worauf wir uns in unserer Stellungnahme konkret bezogen haben, ist: Es geht uns darum, dass es im Abschnitt 9.3.8 zu den indirekten Vergleichen sozusagen eine absolute Ablehnung ist. Dass sozusagen in den meisten Fällen aus methodischen Gründen sicherlich die Voraussetzung nicht erfüllt ist, das kann man diskutieren. Aber wir zielen auf diese absolute Ablehnung in dem Kontext, dass es

ja auch bei unverbundenen Netzwerken theoretisch möglich ist, dass man zwei Studien hat, mit denen man in beiden Fällen individuelle Patientendaten hat. Und dann stellt sich für uns in so einer Situation die Frage: Was ist sozusagen der Unterschied dann zur methodischen Diskussion bei der AbD? Also, die Fragestellung bezog sich, wie gesagt, auf die weiterhin absolute Ablehnung im Abschnitt 9.3.8 und ob man da vielleicht gegebenenfalls eine Einfügung hinzufügen könnte, die auf Sonderfälle hinweisen würde.

Moderator Thomas Kaiser: Ralf, bevor wir da weitermachen, möchte ich gerne fragen: Haben Sie dazu Ergänzungen? Sehen Sie das ähnlich?

Katharina Musiolik: Genau. Unser Punkt ging genau in dieselbe Richtung, explizit eben auch die absolute Ablehnung, die da beschrieben wird, dass es eigentlich keine Alternativen gibt gerade in Bezug auf diese nicht verbundenen Netzwerke und sozusagen welche Alternativen denn sonst gesehen werden, obwohl dann ein Propensity-Score-Matching für die AbD okay ist, aber gleichzeitig in 9.3.8 auch wieder abgelehnt wird.

Ralf Bender: Ja, da bin ich erst mal beruhigt, dass ich das richtig verstanden habe. Es gibt einen elementaren Unterschied zwischen dem Abschnitt 9.3.8 und 9.3.6. In 9.3.8 sind typische indirekte Vergleiche gemeint, die zumindest zum Teil auf aggregierten Daten basieren. Wir werden noch mal prüfen, ob wir das vielleicht auch im Methodenpapier ein bisschen genauer erklären und ergänzen müssten, dass das der Unterschied ist.

Ich erkläre es hier an dieser Stelle mal kurz - wir werden das auch in der DWA dann natürlich noch mal aufgreifen -: Der Unterschied ist, bei 9.3.8 sind indirekte Vergleiche mit zumindest zum Teil aggregierten Daten gemeint. Und mit aggregierten Daten kann man eben keine adäquate vollständige Confounder-Adjustierung hinbekommen bzw. nur mit ganz, ganz starken Annahmen, wo es einen wissenschaftlichen Konsens gibt, dass diese starken Annahmen in der Praxis eigentlich nie erfüllt sind. In 9.3.6 bei AbDs bei multifaktoriellen Verfahren gehen wir hingegen davon aus, dass vollständige individuelle Patientendaten vorliegen, wo man eben alle Möglichkeiten der Modellierung hat.

In einem Fall, dass wir nicht in der AbD sind, sondern einen indirekten Vergleich machen wollen, aber vollständige IPD-Informationen hätten, dann würde das Gleiche gelten wie in dem neuen Abschnitt 9.3.6. Dann könnte man in der Tat Propensity-Scores zum Beispiel anwenden. Aber wenn wir nur zum Teil aggregierte Daten haben, können wir keine adäquate Confounder-Adjustierung hinbekommen. Das ist der Unterschied.

Moderator Thomas Kaiser: Ralf, vielen Dank. Gibt es dazu noch Ergänzungen, Nachfragen jenseits des Punktes, dass wir uns das noch mal genauer angucken, ob wir da noch zur Abgrenzung etwas genauer formulieren müssen? - Das scheint nicht der Fall zu sein.

Dann sind wir eigentlich mit den geplanten Fragen durch. Aber weil wir jetzt so wenig Zeit verwendet haben, habe ich eine Frage an die Firma Pfizer. Sie haben in Ihrer Stellungnahme angegeben, ob es nicht einfacher wäre, statt einer AbD, wie sie jetzt geplant wird, eine minimalinterventionelle Studie durchzuführen. Meine konkrete Frage wäre: Was hindert Sie daran?

Stephan Rauchensteiner: Danke, dass Sie die Frage aufnehmen. Im Grunde würde ich sagen, uns hindert gar nichts daran. Aber es ist natürlich eventuell zu diskutieren, dass ich in den Ausführungen in dem Rapid Report zu versorgungsnahen Daten und auch der Dominanz des Themas „Register“ es so verstanden habe, dass es vom IQWiG und wahrscheinlich G-BA so gewollt wäre, dass wir vorwiegend diese idealtypischen Register heranziehen. Aber wir wollten dann natürlich darauf aufmerksam machen, dass diese minimalinterventionelle Studie eine Möglichkeit wäre, die, so, wie wir das derzeit auch sehen, eventuell sogar für uns einfacher sein könnte. Das muss natürlich geklärt werden, auch aufgrund der Tatsache, dass jetzt in diesen EMA-Guidelines, die jetzt im Januar 2023 in Kraft getreten sind, dieses Konstrukt der minimalinterventionellen Studie dort eindeutig definiert wird, sodass man vielleicht nicht mehr so viele Hindernisgründe hat, diese klinischen Prüfungen in der Art und Weise durchzuführen. Aber es muss diskutiert werden. Ich glaube, es ist nicht im IQWiG Rapid Report zu versorgungsnahen Daten erwähnt, aber ich bin mir nicht mehr sicher.

Moderator Thomas Kaiser: Vielen Dank erst mal für den Hinweis. Das kann ich dann gut klarstellen. Es ist so, dass bei der anwendungsbegleitenden Datenerhebung ja qua Gesetz randomisierte Studien, interventionelle Studien mit einer randomisierten Zuteilung ausgeschlossen sind. Es ist trotzdem so, wenn Sie sich die AbD-Konzepte anschauen, die wir durchführen - und das ist ja auch der Auftrag des Gemeinsamen Bundesausschusses und natürlich auch Wesen dieses Verfahrens -, dass man sich zunächst anschaut, ob denn nicht Studien geplant sind, wie auch immer und welche Studien auch immer, die das Ziel, das man mit der anwendungsbegleitenden Datenerhebung verfolgt, ohnehin schon adressieren und ausreichend adressieren, sodass es nicht sinnvoll und notwendig ist, eine weitere Datenerhebung aufzusetzen. Das mal vorneweg gesagt. Mal angenommen, Sie kommen in ein Verfahren. Sie haben schon eine minimalinterventionelle Studie geplant. Das ist im Register, also nicht in einem Indikationsregister, sondern im Studienregister sichtbar durch den entsprechenden Eintrag. Man sieht, was geplant ist, wie das geplant ist. Dann ist das ja ein Gegenstand des AbD-Konzepts, genau das zu beurteilen und dann potenziell zu sagen, mit dieser Studie wird das bereits erfüllt. Man kann sie halt nicht beauftragen, das ist das Problem im Rahmen der AbD. Da steht es Ihnen natürlich frei, bei den aktuell oder zukünftig laufenden Gesetzgebungsverfahren eine Änderung anzufragen, dass auch randomisierte Studien im Bereich der anwendungsbegleitenden Datenerhebung möglich sind. Und natürlich macht so eine minimalinterventionelle Studie auch Sinn, wie das im Wesen beschrieben ist, also nicht

eine sehr starke Abweichung von dem, was tatsächlich in der Versorgung ist, das heißt nicht keine Abweichung, aber nicht sehr starke Abweichung.

Und ich sage mal: Das Thema „Register“ ist derzeit das vielversprechendste aufgrund der Datenstrukturen, die man hat. So eine minimalinterventionelle Studie kann man ja auch durchführen, indem man vorhandene Registerstrukturen nutzt. Das ist ja letztendlich eine Datenplattform. Aber man kann die auch durchführen, indem man die Datenplattform für die minimalinterventionelle Studie wieder separat aufsetzt.

Um Sie da also zu bestätigen: Ja, wir denken auch, dass so etwas Sinn machen kann. Das hängt von der konkreten Fragestellung ab. Sie sind frei, so etwas zu planen. Wir werden feststellen, wenn es so etwas gibt, dass es das gibt, und in dem konkreten Fall sicherlich auch empfehlen. Wenn diese minimalinterventionelle Studie das Ziel, das man mit der AbD verfolgen möchte, tatsächlich adressiert, werden wir auch vorschlagen, dass man diese Studie heranziehen sollte und nicht eine separate AbD machen sollte.

Stephan Rauchensteiner: Danke, dass Sie das aufgenommen haben. Ich glaube, wie Sie das darstellen, ist es auch daher positiv zu bewerten, dass jeder darüber nachdenken kann. Man kann ja die minimalinterventionelle Studie innerhalb des Registers machen, oder man kann entscheiden, ob man sie außerhalb des Registers macht. Aber ich glaube, es ist grundlegend jetzt positiv zu bewerten, dass hier darüber gesprochen wurde. Danke schön.

Moderator Thomas Kaiser: Vielen Dank noch mal für diese Ergänzung. Vielleicht kann ich da dann auch auf einen Punkt noch mal hinweisen, der möglicherweise untergegangen ist bei dem einen oder der anderen, weil das sowohl in dem - oder nicht „weil“ – sondern es ist in dem Beschluss des Gemeinsamen Bundesausschusses zu den letzten anwendungsbegleitenden Datenerhebungen enthalten, dass zwar primär eine registerbasierte anwendungsbegleitende Datenerhebung sinnvoll ist mit der konkreten Benennung des Registers, wenn das aber nicht durchführbar erscheint, zum Beispiel, weil der Aufwand für die konkrete Anpassung des Registers zu hoch ist, dass es dem Unternehmer freisteht, eine eigene Studie mit einer eigenen Datenplattform aufzusetzen. Also, das ist eine Option, die Ihnen ja freistehen muss als Unternehmen.

Gut, dann wären wir soweit mit dem Thema „anwendungsbegleitende Datenerhebungen“ durch und kommen zu

3.4.3 Tagesordnungspunkt 2: Kein Zusatznutzen auf der Basis von KNBs

Moderator Thomas Kaiser: Da gebe ich an Anja Schwalm ab vom Bereich Gesundheitsökonomie.

Anja Schwalm: Guten Tag auch von mir. Ich würde zunächst einmal die Frage vorlesen.

In unserem ersten Fragenkomplex geht es um die Berücksichtigung des Zusatznutzens in der Kosten-Nutzen-Bewertung. Ich würde erst kurz etwas ausholen. In § 35b SGB V - das ist ja die gesetzliche Grundlage für die Kosten-Nutzen-Bewertung - ist festgelegt, dass diese unter Berücksichtigung des therapeutischen Zusatznutzens für die Patienten im Verhältnis zu den Kosten erfolgen soll. Es ist weiterhin darin festgelegt, dass mit dem Beschluss zur Kosten-Nutzen-Bewertung auch der Zusatznutzen sowie die Therapiekosten festgestellt werden sollen. Wir haben im Methodenpapier beschrieben, dass sich eine Kosten-Nutzen-Bewertung auf Basis eines Modells nicht als Basis für den Nachweis des Zusatznutzens eignet, da zum einen die Zielsetzung als auch die Datengrundlage sich von der von klinischen Studien unterscheiden, und dass daher das Ergebnis der Kosten-Nutzen-Bewertung das Ergebnis der vorhergehenden Nutzenbewertung nicht infrage stellen kann. In verschiedenen Stellungnahmen wurden quasi unsere Festlegungen zur Berücksichtigung des Zusatznutzens aufgegriffen. Meine Frage jetzt an Sie: Wie kann aus Ihrer Sicht die Berücksichtigung des therapeutischen Zusatznutzens erfolgen in der KNB?

Moderator Thomas Kaiser: Da wieder die Frage: Wer hat da einen Vorschlag jenseits dessen, was im Methodenpapier als Option beschrieben ist? Das Thema haben ja einige von Ihnen angesprochen. - Herr Dintsios, Markos.

Ch.-Markos Dintsios: Ich habe es nicht angesprochen, aber die Frage ging ja ans gesamte Publikum.

Es gibt ein Problem dahingehend, dass das IQWiG eine Empfehlung ausspricht nach Sozialgesetzbuch und dort ja einen quasi aggregierten Zusatznutzen über alle Endpunkte hinweg, die positiven oder negativen Effekte auch angibt. Das Problem natürlich ist, wenn man eine Kosteneffektivitätsstudie macht auf einen klinischen Endpunkt, hat man aus dieser Gesamtbewertung einen Aspekt herausgerissen. Man hat nicht die Gesamtbewertung in der gesundheitsökonomischen Evaluation. Wir haben seinerzeit - das liegt lange Jahre zurück - Vorschläge gemacht, wie man diesem quasi Manko entgegenkommen kann, wenn man die Gesamtbewertung des IQWiG, also das, was man als therapeutischen Zusatznutzen im Sozialgesetzbuchjargon meint ... Das eine ist, das, was viele Gesundheitsökonomien propagieren, der QALY. Ich mache keinen Hehl daraus: Ich bin dezidiert Gegner des QALYs. Da gibt es extreme methodische Engpässe. Das andere ist, die einzelnen Endpunkte, so, wie sie aus der jeweiligen Bewertung kamen, gewichten zu lassen und einen Nutzenaggregat zu bilden und wenn es einen Zusatznutzen gibt, dementsprechend einen Zusatznutzenaggregat und dann im Sinne einer Kosteneffektivitätsanalyse, wo die Effekte als aggregierter Nutzen angegeben werden, diese durchzuführen. Aber hinsichtlich der QALY-Kritik spreche ich nicht für mein Unternehmen, da spreche ich für mich selber.

Moderator Thomas Kaiser: Vielen Dank. Gibt es weitere Vorschläge? - Herr Rasch.

Andrej Rasch: Unsere Stellungnahme hat ja in ähnlicher Art und Weise den Punkt adressiert. Ich glaube, dass eine KNB auf dem zuvor bereits festgestellten Zusatznutzen basiert, das ist soweit klar. Das Ganze ist insofern etwas unklar geblieben, weil einerseits eine Möglichkeit für eine Modellierung skizziert wird, zugleich aber auch gesagt wird, dass eben der Zusatznutzen im Rahmen des RCT-basierten Ansatzes aus der zuvor durchgeführten Nutzenbewertung herkommen muss. Und da war für uns etwas unklar, was diese Modellierung insgesamt dann hergeben kann, welche Möglichkeiten Sie überhaupt eröffnen. Vermutlich wird dies jetzt im Rahmen des „IQWiG im Dialog“ adressiert. Insofern bin ich etwas erstaunt, dass man sozusagen hier diesen Punkt vorweggegriffen hat. Aber vielleicht können ja die Erkenntnisse dann trotzdem noch mal einfließen. Das ist der eine Punkt.

Und den zweiten Punkt können wir vielleicht später noch diskutieren. Da geht es um die Interpretation des Zusatznutzens bei den zVT-Szenarien. Aber vielleicht können wir das anschließend diskutieren.

Moderator Thomas Kaiser: Herr Wolf hat sich noch gemeldet.

Jürgen Wolf: Vielleicht jetzt nicht methodisch so spezifisch für die Diskussion bis jetzt, aber aus der Erfahrung in den G-BA-Sitzungen und in Diskussionen bei der Bewertung des Zusatznutzens neuer Medikamente. Wir stoßen ja oft an die gleichen Probleme, dass wir seltene oder niedrigfrequente molekulare Subgruppen von Patienten haben, wo keine randomisierte Studie aus verschiedensten Gründen möglich ist, nicht nur bezüglich des Aufwands, der ist ja meistens stemmbar, sondern bezüglich der Sinnhaftigkeit. Patienten wollen das nicht mehr, lehnen die Randomisierung ab. Die Standards ändern sich viel zu schnell, was die Vergleichsgruppen anbelangt. Und dann sehen wir Medikamente, die in der klinischen Realität eindeutig einen Mehrnutzen haben, zum Beispiel Ansprechraten, wo wir 40 % Ansprechraten haben, wo die Standardtherapie unter 10 % hat, zum Beispiel beim NSCLC im Rezidiv nach Standardtherapie. Und wir haben halt keine Überlebensdaten. Wir stoßen immer wieder hier an diese Mauer. Dann versucht man, es mit Patient-Reported-Outcomes zum Teil zu kompensieren. Und hier würden wir uns wünschen, dass man ... Wir haben das Problem, dass diese seltenen Subgruppen nicht aus Orphan Diseases anerkannt werden, auch von der EMA nicht, was vielleicht methodisch hier manches erleichtern würde oder von der Beurteilung. Aber wir würden uns hier ein ausgewogeneres Konzept der Beurteilung wünschen, also Registerdaten, die integrierte Betrachtung der Endpunkte, vielleicht auch die Patientensicht. Und hier haben wir den Eindruck, so ein bisschen oder regelmäßig an Mauern zu stoßen, wo wir im klinischen Eindruck und auch das, was Patienten sagen, das, was wir erleben, einfach sehen, da ist ein neues Medikament, das bringt eindeutig weniger Toxizität, bessere Lebensqualität, wir sehen ein Ansprechen, und dann gibt es halt keinen Zusatznutzen. Das ist so.

Uns sind die Probleme vollkommen klar. So, wie wir alle sozialisiert wurden in der Zeit, im Zeitalter der Chemotherapie, randomisierte kontrollierte Studien sind da der Goldstandard in so einem Setting. Aber wir haben eben das nicht mehr. Wir haben jetzt biologische Rationale, wir haben diese kleinen Gruppen. Also, es nützt uns nichts, dass wir einen Goldstandard haben, wenn der einfach nicht mehr zeitgemäß ist oder nicht mehr umsetzbar in jeder Situation. Und da war so unser Wunsch, ob man da konstruktiver eine integrativere Sicht der Dinge erarbeiten könnte. Ja, das ist vielleicht ein bisschen unpräzise, aber wir waren wirklich in vielen G-BA-Sitzungen dabei und wir haben das halt oft erlebt. Zum Teil sind ja auch Medikamente dann vom Markt genommen worden, was dramatische Situationen zum Teil im klinischen Alltag nach sich gezogen hat. Und da würden wir uns ein anderes Vorgehen wünschen und konstruktivere Diskussionen.

Herr Kaiser, wir haben es ja neulich mal andiskutiert. Ich glaube, dass man sich einigt auf prospektive Registerstudien, wo wir Ihre Qualitätsmaßstäbe so weit wie möglich verwirklichen. Das finde ich zum Beispiel einen sehr spannenden, konstruktiven Ansatz vielleicht für die nahe Zukunft.

Aber nur, weil Sie es gefragt haben, wollte ich es nicht so ganz unkommentiert lassen. Das geht halt über die reine methodische Diskussion hinaus.

Martin Sebastian: Wenn ich ergänzen darf zu diesem prospektiven Register. Ich glaube, das kriegen wir heute nicht gelöst, aber wir können gemeinsam auch eine Initiative unterstützen. Das, was wir im Moment machen in diesen seltenen Subentitäten, beispielsweise vom Lungenkarzinom. Wenn wir die immer noch als Lungenkarzinom sehen, dann haben wir ein Verständnis der Onkologie von vor fast dem letzten Jahrhundert. Wir rennen uns immer wieder den Kopf blutig damit, so, wie es Herr Wolf jetzt gesagt hat. Ich glaube, dass wir einfach mit der Art und Weise, wie wir das im Moment beurteilen, dem ganzen onkologischen Verständnis und eben auch den Patienten nicht mehr gerecht werden.

Moderator Thomas Kaiser: Vielen Dank. Das geht jetzt eigentlich so ein bisschen weg von dem Punkt Kosten-Nutzen-Bewertung. Aber durch den Begriff „Zusatznutzen“ kann ich verstehen, dass Sie jetzt auf diesen Punkt eingegangen sind.

Anja, ich werde kurz auf den Punkt eingehen und dann würde ich dich bitten, auch noch mal kurz zu den Beiträgen etwas zu sagen, was Zusatznutzen, gesetzter, kein Zusatznutzen im Falle des vorhergehenden, kein Zusatznutzen aus der Nutzenbewertung in der Kosten-Nutzen-Bewertung angeht.

Vielleicht zu dem Punkt. Insbesondere möchte ich den letzten Satz, den Sie gesagt haben, Herr Sebastian, noch mal aufgreifen, wir gingen damit nicht richtig um. Ja, wir gehen nicht richtig in der Datenerhebung damit um. Das ist ein ganz wichtiger Punkt. Worauf sich offensichtlich

fokussiert wird - und da ändert es auch nichts daran, dass man zukünftig gerne in der Versorgung prospektive Studien planen möchte ... Worauf man sich derzeit konzentriert, ist, in der Zulassung das zu erreichen, was zur Zulassung führt. Und damit haben Sie leider in diesen konkreten Situationen die Informationen zum Vergleich häufig gar nicht, teilweise rudimentär, teilweise durchaus vorhanden - es gibt auch Beispiele, wo es vorhanden ist, wo es gut gemacht worden ist -, aber eben in vielen Fällen einfach diese Fragestellung nicht adressiert, vor dem Zeitpunkt oder bis zu dem Zeitpunkt, wo diese Medikamente tatsächlich in die Versorgung kommen. Und da muss man sagen, das ist ein Problem, weil Sie diesen Vergleich dann eigentlich nicht haben. Also, wohin man sich eigentlich aufmachen muss - und da ist weniger die Frage, macht man das mit einem Register, macht man das mit bestimmten Studienprogrammen, macht man das gestuft, indem man während der Zulassung schon bestimmte Studien aufsetzt, nachdem man, sage ich mal, mit dem ersten Programm schon mal den Zulassungsantrag gestellt hat; da gibt es durchaus Beispiele, Osimertinib ist so ein Beispiel ... Wie schafft man es eigentlich, die Daten so früh in der Qualität zu generieren, dass diese Daten dann, wenn es um die Versorgung geht, denn die Nutzbewertung wird ja gemacht, wenn das Arzneimittel in die Versorgung geht, vorhanden sind. Daran muss man arbeiten und nicht an dem Weg, wie man sagt, aus meiner Sicht zumindest, dass man ja die Standards von Bewertungen runtersetzt. Weil die Frage ist nun mal da. Es gibt ein neues Arzneimittel. Wie schneidet es ab im Vergleich zu dem, was da ist? Und daran hapert es.

Jürgen Wolf: Also, mit diesem methodischen Ansatz, dass man das deutlich früher machen muss als jetzt und möglichst abgesprochen, da würde ich Ihnen 100 % zustimmen. Und da ist die Datenqualität, die von den hauptsächlich retrospektiven Real-World-Datenanalysen bis jetzt geliefert wurde, nicht ausreichend, auch wenn man das FDA-Positionspapier liest, zum Beispiel das vom letzten Jahr. Da sind wir auch total konstruktiv. Die Effektstärke, habe ich manchmal das Gefühl, geht so gar nicht in die Beurteilung ein über der ganzen Methodik. Wir haben ja auch Beispiele, wo wir 60, 70 % mehr Ansprechende in den Phase-2-Studien haben. Da sind natürlich andere Bewertungsmaßstäbe wie früher in der Chemotherapie, wo ein sechs Wochen längeres Überleben schon eine positive Studie war. Ich glaube, das ist das, was Herr Sebastian auch gemeint hat. Das, glaube ich, ist ein wichtiger Punkt.

Und man muss es schon sagen: Die G-BA-Beurteilung auf der Basis Ihrer Beratungen ist in vielen Fällen auch wirklich konträr zu allen nationalen und internationalen Leitlinien. Ich denke, das ist schon ein Problem. Also, wir sind konstruktiv. Ich finde den Ansatz, prospektive Registerdaten so früh wie möglich zu erheben, sehr gut. Ich habe auch das Gefühl, in der Pharmaindustrie ist großes Interesse daran. Aber den Punkt muss man halt doch noch mal machen, weil, wie gerade gesagt, wir holen uns halt regelmäßig eine blutige Nase und sind manchmal einfach entsetzt, wenn wir in Deutschland halt wieder diesen Sonderweg gehen und alle amerikanischen und ESMO-Guidelines sagen, das ist jetzt der Fortschritt und das ist die Therapie der Wahl für die Patienten im Gegensatz zum alten Standard. Und dann hören

wir halt, es gibt keinen Mehrnutzen. Also, ich glaube, wir müssen uns beide einfach bewegen. Das ist mein Appell: Wir sollten uns beide bewegen.

Moderator Thomas Kaiser: Vielen Dank. Dann wäre für mich vielleicht noch mal so eine abschließende Bemerkung dazu. Es gibt einen sehr interessanten Draft gerade von der FDA zu einarmigen Studien, der letzten Monat erschienen ist. Er ist nicht direkt auf einarmige Studien betitelt, sondern auf die Frage von beschleunigter Zulassung im Bereich der Onkologie. Sehr lesenswert, finde ich, wo die US-amerikanische Zulassungsbehörde FDA dieses Risiko, das man mit den einarmigen Studien in der Vergangenheit eingegangen ist, sehr klar beschreibt. Und aus meiner Sicht wird da auch ein Weg beschrieben, dass man sich eigentlich wieder davon lösen muss, weil es zu wirklich Schwierigkeiten in der Interpretation führt. Also, Draft Guideline zu onkologischen beschleunigten Zulassungen.

Jetzt habe ich dazu eine Meldung. Herr Friede.

Tim Friede: Vielen Dank, Herr Kaiser. Ich vertrete hier die Deutsche Region der Internationalen Biometrischen Gesellschaft, die GMDS und die DGK. Ich will ganz kurz noch mal die Diskussion aus methodischer Sicht kommentieren. Also, auch aus meiner Sicht, und ich denke, das ist eine Sicht, die viele Kolleginnen und Kollegen in unseren Fachgesellschaften teilen, sieht man durchaus eine Gefahr der einarmigen Studien und wünscht sich daher die kontrollierten Studien, sieht aber auch gleichermaßen, dass diese nicht alleine stehen können, so, wie Sie es beschrieben haben. Aber Sie haben auch, finde ich, einen anderen wichtigen Punkt gemacht, dass man das Wissen über die Biologie hier nicht ausklammern kann, weil das ein ganz starkes Vorwissen ist. Ich würde dafür plädieren, und ich denke, da spreche ich auch im Namen unserer Fachgesellschaften, zu gucken, inwieweit man in transparenter Form Evidenz, die außerhalb von Studien existiert, zum Beispiel im Sinne des Verständnisses der molekularen Biologie, hier auch in die Studie mit einbeziehen kann und wie man auch in gewichteter Form eben auch Informationen aus Studien in anliegenden Bereichen mit einbeziehen kann. Das wäre aus meiner Sicht, denke ich, ein Weg nach vorne, eben halt nicht auf die Kontrolle zu verzichten, sondern die drinnen zu behalten, aber auf gleicher Weise zu gucken, inwieweit man andere Informationen, die vorliegen, formalisieren kann, sodass sie mit in die Auswertung einfließen können.

Moderator Thomas Kaiser: Vielen Dank. Nach diesem Schwenk kommen wir vielleicht noch mal zurück zu dem Tagesordnungspunkt 2, zur Kosten-Nutzen-Bewertung. Anja, vielleicht möchtest du noch mal auf die beiden Beiträge zu der Frage Zusatznutzen in der KNB eingehen.

Anja Schwalm: Es hat sich ja jetzt sehr stark um den Zusatznutzen, die Zusatznutzenbewertung, die ist ja auch der essenzielle Teil der Kosten-Nutzen-Bewertung ... Aber ich würde gerne zurückgehen auf die Kosten-Nutzen-Bewertung. Ich habe mir jetzt drei Punkte aufgeschrieben.

Zum einen noch mal: Wir haben, wie gesagt, diese Prämisse festgelegt, dass das Ergebnis der Kosten-Nutzen-Bewertung das Ergebnis der Zusatznutzenbewertung oder der Nutzenbewertung nicht infrage stellen kann. Ich habe den Eindruck, die Diskussion hier zeigt das noch mal. Diese Prämisse gilt übrigens in beide Richtungen, weil das war auch eine Frage in den Stellungnahmen, das heißt sowohl wenn ein Zusatznutzen vorliegt, als auch wenn keiner vorliegt. Wir haben bewusst keine Festlegung gemacht, wie konkret die Umsetzung ist, wenn kein Zusatznutzen vorliegt, dass auch in der Kosten-Nutzen-Bewertung kein Zusatznutzen generiert wird, weil es da wohl schon sehr viele oder verschiedene Fallkonstellationen gibt. Wir haben uns jetzt bewusst dagegen entschieden, das jetzt im Methodenpapier so festzulegen.

Es waren aber noch zwei Fragen. Zum einen ging es um das Maß des Gesamtnutzens, wie aus dem Ausmaß des Zusatznutzens aus dieser Gesamtbewertung hinsichtlich verschiedener Endpunkte das Ganze in die Kosten-Nutzen-Bewertung übertragen werden kann. Zum einen ist es erst mal so, dass letztlich in unserem Auftrag festgelegt wird - und das macht der Auftraggeber, das ist der Gemeinsame Bundesausschuss -, welche Endpunkte wir zu betrachten haben in der Kosten-Nutzen-Bewertung. Und dann gibt es im Prinzip verschiedene Optionen, wie man damit umgehen kann, wenn es mehrere Endpunkte gibt. Die wurden genannt, und die haben wir aber nach meinem Verständnis auch alle im Methodenpapier dargelegt, nämlich es gibt einmal diese Möglichkeit, eine Gesamtbetrachtung über QALYs zu machen, oder es gibt die Möglichkeit, zu sagen, wenn man denn gewichten will, das auf Basis von Präferenzenerhebungen zu machen. Auch diese Option haben wir eröffnet. Und viel mehr andere Optionen sehe ich tatsächlich nicht, außer, dass man dann gar nicht gewichtet.

Und eine weitere Frage ging noch in die Richtung, was denn die Kosten-Nutzen-Bewertung eigentlich dann zusätzlich liefern kann. Das haben wir ja eigentlich auch beschrieben in der Zielstellung. Wir verstehen sie schon so, dass sie einfach zum einen im Hinblick auf die Kosten und auch noch mal andere Kostenkomponenten betrachtet und das Ganze auch über einen anderen Zeitraum noch mal betrachten kann, aber eben innerhalb eines Modells mit entsprechenden Annahmen, die letztlich die Aussagefähigkeit natürlich dementsprechend auch einschränken.

Moderator Thomas Kaiser: Vielen Dank. - Gibt es dazu jetzt noch mal Nachfragen, Konkretisierungsbedarf? - Wenn das nicht der Fall ist, machen wir mit TOP 3 weiter, ist auch noch Kosten-Nutzen-Bewertung, und hier mit dem konkreten Punkt bei den Orphan Drugs, also den Arzneimitteln für seltene Erkrankungen.

3.4.4 Tagesordnungspunkt 3: Besonderheiten von Orphan Drugs in KNBs

Anja Schwalm: Genau, Kosten-Nutzen-Bewertungen sind auch bei Orphan Drugs möglich, allerdings haben Orphan Drugs, wie Sie alle wissen, eine Sonderstellung im AMNOG. Sie

durchlaufen ja nach Marktzugang zunächst nicht das reguläre Nutzenbewertungsverfahren. Der Zusatznutzen gilt mit der Zulassung als belegt. Und es muss für dieses Verfahren auch keine zweckmäßige Vergleichstherapie festgelegt werden. Eine reguläre Nutzenbewertung gegenüber einer zVT erfolgt insbesondere erst nach Überschreiten einer Umsatzgrenze von 30 Millionen Euro pro Jahr.

In verschiedenen Stellungnahmen wurde angemerkt, dass die Besonderheiten von Orphan Drugs in unserem Methodenpapier nicht berücksichtigt wurden. Meine konkrete Frage bezieht sich jetzt auf eine dieser Besonderheiten, und zwar auf diese Fallkonstellation, dass eben noch keine reguläre Nutzenbewertung stattgefunden hat. Im Nachgang eines Schiedsspruchs kann dann ja theoretisch eine Nutzenbewertung erfolgen, das heißt, es gibt ein Zusatznutzen, aber es gibt keine zVT, die festgelegt wurde in diesem Verfahren. Die konkrete Frage ist: Wie kann eine Kosten-Nutzen-Bewertung bei Orphan Drugs mit einem fiktiven Zusatznutzen erfolgen, für die eben noch keine reguläre Nutzenbewertung stattgefunden hat und es auch noch keine zVT gibt aus dem vorherigen Verfahren? Gibt es dazu Vorschläge?

Moderator Thomas Kaiser: Ja, Ideen, nicht ganz einfach. - Herr Dintsios meldet sich.

Ch.-Markos Dintsios: Die Frage kommt auch nicht von mir, aber bei den Orphan Drugs muss man sich anschauen, ob es überhaupt einen Wettbewerb gibt oder nicht, also rein methodisch jetzt auch zulasten meiner Branche sozusagen, was Sie rein methodisch als Fragestellung hier vorschreiben. Also, ich sieze, weil ich bin im Protokoll, sonst kenne ich ja die meisten von euch. Wenn es ein Solist ist, ist es der natürliche Verlauf. Therapeutischer Nihilismus kann auch als Kontrolle gefahren werden. Natürlicher Verlauf beinhaltet Leistungsanspruchnahme, damit Ressourcenverbrauch, damit Kosten. Wenn es eine Wettbewerbsindikation ist im Orphan-Bereich, kann man auch einen Wettbewerber nehmen. Pulmonale Hypertonie, sieben Vertreter, glaube ich, oder acht. Also, machbar ist das, nur - ich will nicht pessimistisch sein - das wird keine Sau machen.

Moderator Thomas Kaiser: Ist das zufriedenstellend, Anja?

Anja Schwalm: Diese eine Option, natürlicher Verlauf, haben wir auch schon diskutiert. Die anderen, muss ich zugeben, habe ich akustisch nicht ganz verstanden. Könntest du das wiederholen?

Ralf Bender: Bitte langsam, Markos, langsam, deutlich, keine Wörter verschlucken.

Ch.-Markos Dintsios: Ralf, ich habe meinen eigenen Stil und ich kann ihn nicht ändern wegen euch, tut mir leid. Ich bin so geboren, und Fremdwörter benutze ich auch ab und zu aufgrund meiner Herkunft.

Also, wenn ein Orphan Drug in einem Medikationsgebiet zugelassen ist, in dem es Wettbewerb gibt, gibt es potenzielle Komparatoren. Nicht alle Orphan Drugs sind absolute Solisten.

Ist das jetzt klar verstanden worden, Ralf, akustisch? Danke.

Moderator Thomas Kaiser: Gut, vielen Dank.

Dann hätte ich da - und gerne können sich andere daran beteiligen -, weil das jetzt gerade von dir, Markos, gekommen ist, die Frage: Die Feststellung des - ich nenne das jetzt einfach mal so - fiktiven Zusatznutzens beruht ja auf einer Situation, wo genau diese Frage erst mal nicht beantwortet worden ist. Gibt es Wettbewerb, gibt es Komparatoren, was sind die Komparatoren? Das ist ja die Ausgangssituation. Der G-BA macht eine Feststellung, sagen wir mal, nicht quantifizierbar. Das ist ja relativ häufig, das haben wir ja empirisch festgestellt. Relativ häufig sagt der G-BA, es gibt einen nicht quantifizierbaren Zusatznutzen bei Orphan Drugs in der ersten Bewertung. Wie löst man das jetzt, dass kein Komparator definiert wurde, man aber theoretisch natürlich einen Komparator definieren kann für die Zwecke der Kosten-Nutzen-Bewertung, das ja sogar auch in § 35b SGB V steht? Der Gemeinsame Bundesausschuss benennt gegebenenfalls weitere Komparatoren. Wie löst man das denn, dass man einen fiktiven Zusatznutzen festgestellt hat gegen Fragezeichen und eine Kosten-Nutzen-Bewertung durchführt gegen ABC mit dem fiktiven Zusatznutzen aus der Bewertung gegenüber dem Fragezeichen? Ich weiß, das ist schwierig, aber deswegen die Frage, weil Sie das alle so aufgeworfen haben, also nicht alle, aber einige von Ihnen.

Gut, vielleicht kann man auch sagen, das Problem löst sich ab 2026 zur Hälfte und ab 2029 komplett, weil in der europäischen Bewertung ab 2026 die Orphan Drugs im onkologischen Bereich sicherlich auch mit Vergleichen bewertet werden auf europäischer Ebene und ab 2029 dann alle.

Jetzt doch noch mal Markos, bitte.

Ch.-Markos Dintsios: Das Problem mit dem nicht quantifizierbaren Zusatznutzen betrifft ja auch Non-Orphans. Das heißt, die problematische Lage haben wir durchaus. Wenn der G-BA so etwas entscheidet aufgrund der Datenlage, muss man natürlich bei einer Kosteneffektivitätsstudie, wo man einen reinen Endpunkt nehmen würde und die Datenlage nicht ausreicht, so ehrlich sein und sich zugestehen, dass auch keine Kosten-Effektivitäts-Relation ohne Weiteres möglich ist.

Moderator Thomas Kaiser: Gut. Gibt es dazu noch, Anja, Klärungsbedarf? Von Ihrer Seite noch mal irgendetwas dazu? - Herr Rasch.

Andrej Rasch: Ich befürchte tatsächlich, dass an dieser Stelle die Diskussion vermutlich methodisch kaum weitergeführt werden kann. Ich möchte das ungern zu einer rechtlichen Diskussion machen, ob überhaupt dann eine KNB bei einer Orphan Drug ohne zuvor bestimmte zVT durchgeführt werden dürfte. Aber das müssen wir hier tatsächlich, glaube ich, nicht diskutieren.

Moderator Thomas Kaiser: Vielen Dank. Es steht Ihnen natürlich frei, methodische Beiträge zu liefern, denn es ist ja eine Erörterung zum Methodenpapier. Bislang habe ich das nicht gehört. Das ist unsere primäre Frage. Gut, gibt es auch weiterhin nicht. Dann vielen Dank, Herr Rasch.

Dann wären wir schon bei ...

Andrej Rasch: Herr Kaiser, Entschuldigung. Ich hatte ja noch einen Punkt bezüglich der Kosten-Nutzen-Bewertung. Das hätte eher zu TOP 2 gepasst, aber TOP 3 ist immer noch KNB. Und bevor wir ganz andere Themen ansprechen, möchte ich das gerne noch ansprechen. Das ist das, was wir in unserer Stellungnahme eingebracht haben, nämlich die Interpretation des Zusatznutzens gegenüber der zVT, wenn es mehrere zVT-Alternativen gibt, also die Frage der zVT-Szenarien. Das wird ja in dem Methodenpapier so beschrieben, dass dann angenommen wird, als würde man das jetzt alles übertragen können auf die anderen zVT-Vertreter bei einer Oder-Verknüpfung. Wir haben ja - ich will jetzt nicht alles wiederholen, was in der Stellungnahme steht - hinreichende Beispiele beschrieben, warum es aus unserer Sicht problematisch ist, beispielsweise bezüglich einzelner Endpunkte von gleicher Wirksamkeit auszugehen, insbesondere weil es durchaus Beispiele gibt, wo trotz einer Oder-Verknüpfung auch ein beträchtlicher oder auch ein geringer Zusatznutzen insgesamt oder basierend auf einzelnen Endpunkten vorliegen kann. Wäre es möglich, das noch anzusprechen?

Moderator Thomas Kaiser: Ja, das haben Sie ja gerade schon angesprochen. Wir hatten dazu keine Frage, weil uns Ihre Anmerkungen klar waren. Ich meine, das ist ja kein spezifisches Thema der Kosten-Nutzen-Bewertung, sondern das ist ja die Frage der Definition der zweckmäßigen Vergleichstherapie generell. Und wenn man Ihren Weg verfolgt, würde man damit die gesamte Festlegung der zweckmäßigen Vergleichstherapie durch den G-BA infrage stellen, der in der Regel die pharmazeutischen Unternehmen folgen. Insofern: Kann man aufbringen, hat aber weitreichende Folgen und hat mit der Kosten-Nutzen-Bewertung primär nichts zu tun.

Aber da gibt es eine Wortmeldung, die das möglicherweise anders sieht. - Herr Sauerbruch.

Tilman Sauerbruch: Nein, ich wollte nur, bevor Sie von der Kosten-Nutzen-Bewertung wegkommen, auch noch eine Frage, die wir noch gestellt hatten, stellen. Inwieweit sind

Mitarbeiter vom IQWiG nach Abschluss der Kosten-Nutzen-Bewertung in die Preisverhandlungen mit eingeschaltet, und inwieweit kann es da mehr Transparenz geben?

Moderator Thomas Kaiser: Das kann ich nur insofern beantworten, als wir bei der Preisverhandlung keine Rolle haben.

Und über die Herstellung oder Nichtherstellung von Transparenz in der Preisverhandlung kann ich Ihnen nichts sagen. Das ist die Aufgabe der dort beteiligten Parteien. Dass man darüber streiten kann, ob das derzeit so sinnvoll ist, da gebe ich Ihnen recht, aber das ist nicht unsere Rolle.

Aber jetzt hat sich Markos Dintsios noch einmal gemeldet.

Ch.-Markos Dintsios: Bei den Ausführungen von Herrn Rasch ist ein kleines Verständnisproblem aufgetreten. Ich glaube, Herr Rasch meint die Darstellung aller potenziellen Komparatoren visuell über ihre Gesamtkosten, aber ohne potenzielle Nutzendifferentiale mit einzubeziehen. Ich glaube, das ist das Hauptproblem. Es gibt ja Komparatoren, die bereits bewertet wurden. Es gibt Fallkonstellationen, wo drei sogar bewertet wurden, potenzielle Komparatoren, die mit der Oder-Verknüpfung durch den G-BA als solche benannt werden können. Und die Frage ist, ob in Kenntnis dessen natürlich so eine visuelle Darstellung in die Irre führt. Weil ganz einfach: Sie lässt vorhandene Informationen vollkommen außen vor. In den Konstellationen, in denen zu den Komparatoren überhaupt nichts hinsichtlich der Evidenz vorliegt, könnte man noch so was wagen. Aber es ist immer noch als visuelles Signal zu hinterfragen. Aber in denen, in denen sich die Komparatoren, weil sie bereits bewertet wurden, hinsichtlich ihres Zusatznutzens unterscheiden, ist das eine Informationsvorenthaltung. Das ist ja restringierend. Wieso soll man diese Information visuell dem Entscheidungsträger, den visiert ja so eine Grafik an, vorenthalten wollen? Ich glaube, das hat Herr Rasch gemeint.

Moderator Thomas Kaiser: Wenn ich darauf antworten darf. Das haben wir durchaus verstanden, dass Herr Rasch das gemeint hat. Wir kennen ja auch die schriftliche Stellungnahme. Aber das, was du gerade erwähnt hast, trifft auf die Beschlüsse zur Nutzenbewertung gleichermaßen zu. Es gibt eine Aussage zum Zusatznutzen und zehn Zeilen zu den Kosten, nämlich zu den einzelnen mit „oder“ verknüpften Arzneimitteln. Das ist nicht unterschiedlich.

Gerne noch mal, aber dann müssen wir unser Geplänkel beenden.

Ch.-Markos Dintsios: Ja, noch mal, aber die Darstellung ist eine reine Kosteneffektivitätsdarstellung. Da machen wir uns nichts vor. Auf der Ordinate ist der Nutzen und auf der Abszisse sind die Kosten. Damit wird etwas suggeriert, was nicht stimmt. Denn alle diejenigen, zu denen Evidenz vorliegt und zu denen eine Zusatznutzenbewertung gemacht wurde, werden

gleichgestellt. Das ist auch wissenschaftstheoretisch ein Fehler. Das muss man klipp und klar sagen. Denn Sie stellen als IQWiG, also „Sie“ wieder als Höflichkeitsformel, wir kennen uns ja persönlich alle hier, eine Kosteneffektivität dar und reduzieren die Informationen. Und noch mal: Das ist wissenschaftstheoretisch falsch.

Stefan Lange: Es würde jetzt das Format der Veranstaltung sprengen, wenn wir über Wissenschaftstheorie sprechen. Ich wundere mich so ein bisschen. Im TOP vorher hatte ich den Eindruck, es kamen jedenfalls keine anderen Ideen dazu, dass es schon notwendig ist, dass zunächst mal der Beschluss der Nutzenbewertung gilt, auch für die Kosten-Nutzen-Bewertung. Damit ist eigentlich alles klar. Im Zweifel muss man sich an anderer Stelle noch mal an den G-BA wenden, wenn man den Eindruck hat, hier ist die zVT falsch bestimmt. Aber das ist dann nicht hier die Baustelle.

Moderator Thomas Kaiser: Jetzt meldet sich erst Herr Rasch und dann Herr Dintsios noch mal.

Andrej Rasch: Vielleicht das an einem Beispiel noch mal zu präzisieren und zu sagen: Die Festlegung der zVT dient in erster Linie der Möglichkeit des Zusatznutzennachweises gegenüber diese in der Oder-Verknüpfung aufgezählten Komparatoren. Das ist erst mal die eine Fragestellung.

Für die KNB würde ich aber dann tatsächlich die Sachlage ignorieren. Ich nenne einfach mal ein Beispiel. Ich habe Ergebnisse zu einem PASI 90 oder PASI 100 in der Psoriasis. Ich weiß, dass A gegenüber B eine klare Überlegenheit zeigt und darauf sogar ein beträchtlicher Zusatznutzen basiert. Aber unabhängig davon, welcher Zusatznutzen hieraus folgt, ich weiß, dass in diesem Endpunkt eine klare Überlegenheit da ist, die auch in der IQWiG-Nutzenbewertung zuvor aufgezeigt wurde. Es wäre aus meiner Sicht aber wirklich schräg, wenn man tatsächlich dann in diesem zVT-Szenario, in der grafischen Darstellung, einfach eine Linie ziehen würde und diese Erkenntnis ignorieren würde. Das ist wirklich dann unschlüssig.

Moderator Thomas Kaiser: Herr Dintsios, Markos, noch mal notwendig?

Ch.-Markos Dintsios: Ja, noch mal. Es ging ja nicht um die Auswahl des Komparators. Ich glaube, da bin ich missverstanden worden. Es geht um die Darlegung der Kosteneffektivität potenzieller Komparatoren auf Basis der G-BA-Empfehlung. Das war der Hinweis. Das kann man ja tabellarisch machen. Man kann es nur auf die Kosten beziehen, ist ja kein Problem. Dann macht das IQWiG eine schöne Tabelle mit den Kosten der potenziellen Komparatoren und es reicht aus. Da brauche ich keine Visualisierung, wo ich eine Effektordinale drin habe, weil dazu habe ich ja keine Daten.

Ich hoffe, es ist jetzt besser verstanden worden.

Moderator Thomas Kaiser: Ja, vielen Dank. Wie gesagt, es war auch vorher schon verstanden worden. Es gibt bislang für mich keine neuen Informationen an der Stelle. Es bleibt dabei, auch da, wo Sie eben gerade gesagt haben, Herr Rasch, na ja, bei der Nutzenbewertung, das ist ja dafür. Ihr A-versus-B-Beispiel, und A und B sind ja „oder“-verknüpfte Sachen, hat genau die gleiche Konsequenz für die Nutzenbewertung. Wenn Sie jetzt einen Vergleich von C - das ist das neue - gegenüber B machen und da einen Zusatznutzen feststellen, aber eigentlich A gegenüber B schon nachgewiesenermaßen besser war, egal, ob das IQWiG oder wer auch immer das bewertet hat, trotzdem beide in der zweckmäßigen Vergleichstherapie sind, und die Nutzenbewertung feststellt, dass der Zusatznutzen von C nicht nur gegenüber B, sondern auch gegenüber A gilt - und das ist das Verfahren - ... Und da kann ich nur noch mal sagen: Dieser zweckmäßigen Vergleichstherapie wird in 99 % der Fälle von den pharmazeutischen Unternehmen zugestimmt. In den Dossiers ist das genau die gleiche Situation. Der Zusatznutzen bezieht sich immer auf alle „oder“-verknüpften Komparatoren.

Gut, ich glaube, wir sind uns einig, dass wir uns uneinig sind. Vermutlich müssen Sie noch mal eine Stellungnahme beim G-BA abgeben an der Stelle.

Aber jetzt habe ich noch mal eine andere Meldung, und zwar von der Firma Novartis, von Frau Lambertz.

Lisa Lambertz: Wir hatten hinsichtlich der Kosten-Nutzen-Bewertung noch einige konkretisierende Fragen gestellt hinsichtlich der konkreten operativen Schritte, die da gegangen werden müssen, beispielsweise zur Kosteneffektivitätsanalyse, wenn, wie eben schon angesprochen, mehrere Endpunkte als Zielgrößen beauftragt werden. Da hat es uns an einer Konkretisierung der Methodik für diese Präferenzhebung durch Patienten gemangelt an der Stelle. Ist zu erwarten, dass da von Ihrer Seite noch Konkretisierungen folgen hinsichtlich unserer Rückfragen?

Moderator Thomas Kaiser: Wenn ich darauf antworten darf, Anja. Ich antworte deswegen, weil es nicht nur für die Kosten-Nutzen-Bewertung gilt. Sie werden sicherlich verstehen, dass eine abschließende Beurteilung nach dieser Erörterung dann in der Dokumentation dargelegt wird. Aber wir nehmen - und das meinen wir auch ernst - Ihre Anregungen ernst. Es wird sicherlich Änderungen geben. Ob das jetzt aber der konkrete Punkt zum Beispiel ist, braucht einfach die Beurteilung. Aber die Sachen sind gelesen, bekannt. Die Methoden 7 werden nicht so aussehen, wie der Methodenentwurf 7, das ist völlig klar. Und wir sind ja dankbar für hilfreiche Vorschläge, wo es unklar war. Also, das ist auf jeden Fall hilfreich.

Lisa Lambertz: Gut, vielen Dank.

Moderator Thomas Kaiser: Gut, dann verlassen wir jetzt das Feld Kosten-Nutzen-Bewertung und kommen zu der 80 %-Regel. Das macht Guido Skipka vom Ressort Biometrie.

3.4.5 Tagesordnungspunkt 4: 80 %-Regel

Guido Skipka: Kurzer Titel des TOPs, 80 %-Regel. Ich möchte Sie kurz da abholen, was wir darunter verstehen. Es gibt Situationen, wo unsere Fragestellung, die wir bewerten wollen, enger gefasst ist als die Fragestellung einer vorliegenden Studie. Es kann sein, dass die Population der Studie breiter aufgestellt ist als für die Fragestellung, die wir haben. Das kann sich aber auch auf die Charakteristika der Interventionsgruppen beziehen. Ich werde im Folgenden mal bei dem Beispiel Population bleiben, weil das für die Interventionsgruppen gleichermaßen gilt.

Wir haben bisher folgendes Vorgehen: Wenn wir eine Studie haben, die breiter aufgestellt ist als unsere Fragestellung, dann schauen wir zunächst, ob wir eine Subgruppenanalyse oder eine Analyse der Teilpopulation haben, die auf unsere Fragestellung zutrifft. Dann nehmen wir die Auswertung der Teilpopulation für unsere Bewertung. Wenn das nicht der Fall ist, dann überlegen wir, inwieweit wir das gesamte Ergebnis übertragen können. Da haben wir als Richtwert eine 80 %-Regel eingeführt - das ist schon seit Längerem so -, wo wir dann schauen: Wenn mindestens 80 % der Population auf unsere Zielpopulation zutrifft, dann übertragen wir das Studienergebnis, was sich auf die weitere Population bezieht, auf unsere Zielpopulation.

Wir hatten auch im bisherigen Vorgehen den Fall, dass wir vielleicht unter 80 % liegen. Um solche Ergebnisse dann heranzuziehen, brauchen wir geeignete Informationen, die es uns erlauben, festzustellen, dass diese Übertragung trotzdem statthaft ist.

So viel zum bisherigen Vorgehen.

Wir haben den Abschnitt 9.1.1 zu Kriterien für den Einschluss von Studien etwas umformuliert. Dazu gab es einige Stellungnahmen von Ihnen. Deswegen richte ich meine Ausführungen auch an Sie alle in Ihrer Gesamtheit. Es gab zum Teil Unklarheiten oder auch Missverständnisse von Ihrer Seite. Ich werde versuchen, die jetzt weitestgehend auszuräumen.

Wir haben gar nicht viel geändert. Das Einzige, was neu hinzugekommen ist, ist, dass wir, auch wenn die 80 %-Regel für eine Studie zutrifft, dann trotzdem schauen müssen, wenn wir gesicherte Erkenntnisse haben über eine stärkere Effektmodifikation, die genau dieses Kriterium, was die Populationsausweitung oder -einschränkung angeht, betrifft, wie wir mit dem Ergebnis umgehen. Anders ausgedrückt: Es gibt Situationen, wo wir trotz Erfüllung der 80 %-Regel die Ergebnisse nicht uneingeschränkt übertragen können. Es hat in der Vergangenheit ein, zwei Beispiele gegeben, was zu diesem Änderungsvorschlag geführt hat.

Ich hoffe, dass ich die Unklarheiten in Ihren Stellungnahmen jetzt erst mal ausräumen konnte. Ansonsten wäre jetzt meine Frage, ob Sie weiteren Erläuterungsbedarf sehen.

Moderator Thomas Kaiser: Dann gebe ich das mal an das Publikum, an die Stellungnehmenden weiter. Gibt es Klärungs-, Beratungs-, Diskussionsbedarf noch zu dem Punkt?

Guido Skipka: Ich kann, während Sie noch überlegen, kurz schon mal sagen: Ich habe mir den Text natürlich die Woche auch noch mal durchgelesen und konnte verstehen, dass es da zu Missverständnissen kommen kann. Ich kann, glaube ich, hier schon mal sagen, wir werden das auf jeden Fall mal klarer formulieren, so, wie ich es jetzt eben versucht habe, zu verdeutlichen.

Moderator Thomas Kaiser: Herr Rasch.

Andrej Rasch: Ich hoffe, dass das jetzt klarer geworden ist. Aber korrigieren Sie mich, wenn ich doch falsch liege in meinem Verständnis. Wir hätten einen Fall, wo die Studienpopulation zu 90 % die Fragestellung erfüllt, Sie aber sehen sozusagen, dass es hinsichtlich dieser restlichen 10 % eine sehr starke Effektmodifikation bei einem Endpunkt gibt. Für mich ist immer noch etwas unklar: Was ist dann die Konsequenz? Wäre die Studie dann in der Gesamtheit nicht mehr verwertbar, oder würde das hinsichtlich der Diskussion des Verzerrungspotenzials eine Rolle spielen?

Moderator Thomas Kaiser: Guido, bitte.

Guido Skipka: Das hängt von der Stärke der Effektmodifikation, von der Stärke der Effekte ab. Das kann im einfachsten Fall dazu führen, dass wir die Aussagesicherheit vielleicht herabsetzen. Im Extremfall kann das dazu führen, dass wir die Ergebnisse gar nicht berücksichtigen.

Moderator Thomas Kaiser: Wenn ich da kurz ergänzen darf, Herr Rasch. Wenn Sie sich zum Beispiel mal vorstellen, es gibt bestimmte Ereignisse, die treten bei 90 % und genau diesen 90 %, die interessieren, praktisch nicht auf, die treten aber bei 50 % der Population, die die anderen 10 %, die nicht interessieren, auf, dann haben Sie ja immerhin 5 Prozentpunkte innerhalb der Gesamtpopulation und dadurch möglicherweise einen Effekt, der ganz alleine auf diesen 10 % beruht. Ich beschreibe gerade eine extreme Effektmodifikation. Und das ist das, was Herr Skipka eben meinte. In solchen Fällen kann man die Ergebnisse nicht für die interessierende Population heranziehen.

Eine andere Situation haben Sie, wenn Sie vielleicht einen kleinen Effekt sehen in der 90%-Population, die interessiert, aber einen sehr starken in der 10%-Population. Dann haben Sie ja schon mal eine ähnliche Effektrichtung. Dann geht es primär um die Effektstärke und die Gesamtaussagesicherheit.

Das sind so Situationen, die man unterscheiden muss. Das ist der Punkt.

Sonst noch Fragen dazu?

Guido, du hast noch einen Punkt.

Guido Skipka: Ich könnte noch ergänzen. Sie brauchen keine Sorge zu haben, dass wir jetzt regelhaft, auch wenn die 80 % erfüllt sind, die Studien nicht heranziehen. Das sind Einzelfälle, die wir hatten, womit wir uns jetzt aber einen flexibleren Umgang wünschen. Aber das wird nicht der Regelfall sein. Die 80 % wird in der Regel so wie gehabt angewendet.

Moderator Thomas Kaiser: Herr Ellis.

Aneurin Ellis: Nur eine kurze Verständnisfrage, was die 80 % angeht. Wenn man jetzt zum Beispiel das nur zu 70 % erfüllt, heißt das dann jetzt, dass die Studien gar nicht mehr betrachtet werden können, oder ist das in die Richtung auch noch offen?

Guido Skipka: Danke, dass Sie da noch mal nachfragen. Wie gesagt, wir werden das in dem Abschnitt noch mal klarer formulieren. Daran hat sich nichts geändert. Wie auch im alten Entwurf können wir Ergebnisse heranziehen, auch wenn die 80 %-Regel verfehlt wird. Das hat natürlich auch Grenzen. Aber wenn wir jetzt eine 70%ige Überdeckung sehen, haben aber gesicherte Kenntnisse, dass die Übertragbarkeit nicht eingeschränkt ist - das wird auch nicht oft vorkommen, fürchte ich -, ziehen wir die Ergebnisse natürlich heran.

Moderator Thomas Kaiser: Stefan, bitte.

Stefan Lange: Ich würde nur gerne noch ergänzen, damit kein falscher Eindruck entsteht, wir könnten die Studie nicht verwerten. Natürlich können wir die Studie verwerten, idealerweise indem wir die spezifischen Informationen haben zu den 70 und 30 %. Dann gibt es gar kein Problem. Das vielleicht noch an der Stelle als Ergänzung.

Moderator Thomas Kaiser: Genau, das ist ein wichtiger Punkt, dass man im Zweifelsfall in einer Situation, wo man möglicherweise ein Dossier erstellt oder beauftragt, eher die Informationen so vorlegt, dass man beide Wege gehen könnte, die interessierende Population sich isoliert anzuschauen oder aufgrund einer Argumentation keine relevante Effektmodifikation die Gesamtpopulation darstellt. Und, wie gesagt, am besten beides.

Gut, dann wären wir, glaube ich, mit diesem Punkt „80 %“ durch. Jetzt gibt es doch noch eine Wortmeldung.

Vanessa Schaser: Ich habe noch eine Frage. Sie haben ja gemeint, dass die Frage, die Sie sich stellen oder die wir uns stellen, im Dossier vielleicht enger sein kann als die der Studienpopulation. Da gibt es ja auch, um dem Power-Verlust entgegenzuwirken, wenn wir die engere Population betrachten wollen, das von Ihnen veröffentlichte Arbeitspapier zur

Übertragbarkeit von Studienergebnissen auf die Teilpopulation. Uns hat sich die Frage gestellt, warum es nach wie vor nicht Einzug gefunden hat ins Methodenpapier oder vielleicht auch in der Praxis gar nicht wirklich Anwendung findet.

Moderator Thomas Kaiser: Dieses Arbeitspapier ist ja schon etwas älter. Und dass es keinen Eingang in die Praxis gefunden hat, auch nicht ins Methodenpapier, zeigt, dass wir hier eine Methodenarbeit gemacht haben, die uns aber nicht derart überzeugt hat, dass das, was man hier gehen könnte, tatsächlich mit einem solchen Power-Gewinn einhergeht und in Abwägung von Inkaufnahme von Unsicherheit - das ist ja auch damit verbunden -, dass wir das als einen Regelfall ansehen würden.

Jetzt aber sind wir durch mit der 80 %-Regel? - Okay, wunderbar. Vielen Dank.

Dann sind wir bei

3.4.6 Tagesordnungspunkt 5: Umgang mit variablen Beobachtungszeiten

Moderator Thomas Kaiser: Das wird Daniela Preukschat aus dem Ressort Arzneimittelbewertung machen.

Daniela Preukschat: Guten Tag auch noch mal von meiner Seite. Wir haben im Entwurf zum neuen Methodenpapier ein paar Sätze in einem Abschnitt ergänzt, den es schon länger gibt, nämlich den Abschnitt zum Umgang mit variablen Beobachtungszeiten. Wir haben dort noch mal herausgehoben, dass wir bemängeln, dass in der Regel für die Patient Reported Outcomes, die PROs, uns Daten nach Behandlungsabbruch, nach Progress fehlen und dass das aus unserer Sicht einfach zeigt, dass diese Endpunkte systematisch verkürzt erhoben werden und dass infolge die Interpretierbarkeit der PRO-Ergebnisse stark eingeschränkt ist.

Wir haben zum anderen beschrieben, dass die dauerhafte Veränderung im Fall von relevant unterschiedlichen Beobachtungszeiten in der Regel nicht für die Nutzenbewertung herangezogen werden kann, da das Grundproblem besteht, dass sie im länger beobachteten Arm einfach schwerer zu erreichen ist und es dadurch zu einer Verzerrung kommt. In diesem Zusammenhang würde ich gerne festhalten, dass wir grundsätzlich Zustimmung in den Stellungnahmen zu dieser Beobachtung sehen, dass die dauerhafte Veränderung im länger beobachteten Arm schwerer zu erreichen ist.

Daran schieße ich die erste Frage an. In einigen Stellungnahmen - ich nenne jetzt hier mal stellvertretend Pfizer und vfa - klingt es für uns so, als ob eine Diskrepanz gesehen wird zwischen unseren Ausführungen im Methodenpapier und den FAQs des G-BA, die auf der Website im Zusammenhang mit der Modulvorlage zu finden sind. Deshalb die Frage an alle und spezifisch an die beiden: Ist dies so, und - wenn ja - worin genau besteht aus Ihrer Sicht die Diskrepanz?

Moderator Thomas Kaiser: Gut, zwei stellungnehmende Organisationen, Firmen sind hier angesprochen worden. - Herr Rasch.

Andrej Rasch: Die Diskrepanz ist vielleicht gar nicht so tiefgehend, wie sie interpretiert wurde. Die ist letztlich so gemeint, dass der G-BA hier ... Sie wissen ja, dass ich mit den Vorgaben in den FAQs meine Probleme habe. Aber unabhängig davon, das ist ein pragmatischer Vorschlag, bei dem zunächst einmal die Möglichkeit gegeben wird, eine Operationalisierung vorzulegen und die auch zu begründen. Wir verstehen es ja methodisch und wollen das auch in dem Sinne nicht großartig hinterfragen. Nur, wie gesagt, es erschien so, als würde der G-BA in der Hinsicht das pragmatisch handhaben wollen und erst mal den Weg eröffnen wollen, auch bei Unterschieden in der Beobachtungsdauer die Operationalisierung zu wählen, die aber auch entsprechend zu begründen. Vielleicht läuft es aber auch im Endeffekt auf dasselbe hinaus, wenn ich das in der Nutzenbewertung tue und die Unterschiede als nicht maßgeblich eingestuft werden, dass dann auch das IQWiG das entsprechend übernimmt. Aber vielleicht sollte man das auch einfach etwas präzisieren im Methodenpapier, dass diese Möglichkeit besteht und dass es dann auch auf diese Abwägung ankommt, auf die Beurteilung der Unterschiede.

Moderator Thomas Kaiser: Frau Böhme.

Sarah Böhme: Ich würde gerne dazu ergänzen, weil ja erwähnt wurde, dass wir es in unserer Stellungnahme kommentiert haben. Wir hatten uns vielmehr darauf bezogen - in den FAQs ist es entsprechend beschrieben -, die methodischen Herausforderungen, Probleme in der Interpretation der dauerhaften Verschlechterung. Und es wird letztendlich darauf verwiesen, dass zumindest die Zeit bis zur ersten oder bis zur bestätigten Verschlechterung in diesen Konstellationen heranziehbar ist. So, wie wir das Methodenpapier verstanden haben, wird sozusagen an der einen Stelle gesagt, dass solche Daten in Situationen von gravierend unterschiedlichen Beobachtungszeiten gar nicht zu interpretieren sind. Wir wollten einfach mit unserer Stellungnahme darauf hinweisen, dass wir es als wichtig erachten, dass man sich natürlich auch die konkrete Datensituation im Kontext der Verläufe der Beobachtungen anschaut - die sind ja auch fester Bestandteil der Dossiers, werden entsprechend mit eingereicht -, dass man sozusagen auf Basis der Daten schauen muss - hier geht es auch wieder um eine kategorische Ablehnung, dass wir die nicht sehen -, wie der Verlauf bei den Patient Reported Outcomes ist. Wir kennen das ja auch im Kontext der unerwünschten Ereignisse, dass man häufig sieht, dass eine Verschlechterung direkt am Anfang ist. Und wenn ich dann beispielsweise die Zeit bis zur ersten Verschlechterung betrachte, dann kann keine Verschlechterung mehr zum späteren Zeitpunkt kommen. Das war der Punkt, auf den wir eingehen wollten. Vielleicht haben wir das Methodenpapier an der Stelle auch falsch verstanden. Aber deswegen hatten wir das noch mal kommentiert.

Moderator Thomas Kaiser: Vielen Dank. - Möchtest du dazu etwas sagen, Daniela?

Daniela Preukschat: Gerne. Vielen Dank für die Ausführungen. Vielleicht kann ich noch mal erläutern, was in den FAQs des G-BA steht. Dort steht, dass bei unterschiedlichen Beobachtungszeiten immer die erstmalige Veränderung vorzulegen ist. Das sehen wir genauso. Das vielleicht noch mal zur Klarstellung. Und die ist dann auch in der Regel interpretierbar und kann für die Nutzenbewertung herangezogen werden, wohingegen die bestätigte und die dauerhafte in diesen Situationen problematisch ist. Also, dahingehend sollen unsere Äußerungen im Methodenpapier mit den Vorstellungen des G-BA übereinstimmen, während es, wenn wir keine relevanten unterschiedlichen Beobachtungsdauern haben, grundsätzlich inhaltlich sinnvoll ist - und das beschreibt auch der G-BA so -, eine bestätigte oder dauerhafte Veränderung zu betrachten. Dort wird im Einzelfall dann geprüft, ob in der vorliegenden Datensituation diese Operationalisierungen herangezogen werden können.

Gibt es vielleicht noch Ergänzungen?

Moderator Thomas Kaiser: Ja, ich habe noch, danke Daniela, eine Ergänzung, Frau Böhme, zu Ihrem Punkt, den Sie eben gemacht haben, man sollte sich das, sage ich mal, in der Datensituation mit den Verläufen angucken. Hier liegt ja genau das Problem. Wenn Sie einen Arm schnell beenden, können Sie den Verlauf sich ja nicht mehr anschauen. Sie haben gerade so ein Beispiel gemacht. Es gibt jetzt vielleicht ganz früh Nebenwirkungen, und das ist auch kongruent damit. Aber ich sage mal, gehe ich jetzt mal auf den Punkt Lebensqualität, gehe ich mal auf Schmerzen, gehe ich auf andere Symptome, dann heißt das ja nicht, dass die alleine durch eine spezifische Nebenwirkung beeinflusst werden, sondern potenziell eben auch durch weitere Dinge, die im Krankheitsverlauf entstehen. Und da liegt genau das Problem, dass ich zwar Ihr Argument nachvollziehen kann, man sollte sich die Daten dann im Verlauf anschauen, aber das geht eben nicht, weil Sie in einem Arm genau diese Beobachtung zu einem relativ frühen Zeitpunkt beenden. Das ist das Problem, das da entsteht.

So, jetzt habe ich erst mal Frau Böhme und dann Stefan.

Sarah Böhme: Vielen Dank, Herr Kaiser. Ich weiß nicht, ob hier ein Missverständnis vorliegt mit dem, was ich meinte. Ich habe es darauf bezogen, wenn wir uns jetzt anschauen, wir haben Daten für beide Arme, also noch nicht der Zeitpunkt, wo wir keine Beobachtung mehr für einen Arm haben, aber wir haben einen Zeitraum T0 bis T und sehen, dass in diesem Zeitraum einfach in einem Behandlungsarm beispielsweise 90 % der Patienten - das ist ein seltener Fall bei PROs; das weiß ich aber - eine Verschlechterung haben, dann ist für mich, selbst wenn die Beobachtung dann endet in einem der weiteren Arme, die Zeit bis zur erstmaligen Verschlechterung definitiv relevant. Diesen Punkt hatten wir gerade geklärt. Und darum ging es auch, einfach da zu berücksichtigen, welche Analyse im Hinblick auf welche Ereignisse kann ich eventuell denn nicht mehr beobachten. Das muss ich aber berücksichtigen. Was sehe ich im Zweifel nicht, wenn ich die Beobachtung abbreche? Und einfach sozusagen da auf Basis

der Daten zu schauen, ob hier überhaupt noch eine Veränderung letztendlich auch in den Ergebnissen stattfinden kann, darauf bezogen war es. Die UEs waren nur ein Beispiel, es war gar nicht darauf bezogen.

Moderator Thomas Kaiser: Okay. Vielen Dank. - Stefan, du hast dich gemeldet.

Stefan Lange: Ja, um es vielleicht noch etwas deutlicher zu machen. Das Problem entsteht ja gar nicht, wenn einfach die Beobachtung nicht abgebrochen wird. Das ist ja leider in den Studien so angelegt, dass systematisch in bestimmten Situationen die Beobachtung abgebrochen wird aus Gründen, die mehr oder weniger nachvollziehbar sind, zum Beispiel weil in der Onkologie eine Progression eingetreten ist. Man kann ja weiter beobachten, natürlich wenn die Patientinnen und Patienten einverstanden sind. Das macht auch deshalb Sinn, weil wir inzwischen doch auch bestätigte Evidenz haben, nachhaltige Evidenz haben, dass generell das Einholen der Symptomatik bei den Patientinnen und Patienten ihr Überleben sogar verbessert. Also, es spricht überhaupt nichts dagegen, das fortzuführen. Man kann sich vielleicht dann überlegen, dass man nicht so ganz komplizierte Fragebögen nimmt mit 448 Items. Darüber kann man ja sprechen. Aber, wie gesagt, das ganze Problem würde überhaupt nicht auftreten, wenn eben Studien in dieser Weise anders konzipiert würden.

Moderator Thomas Kaiser: Gut. Da gibt es eine Meldung. Herr Friede.

Tim Friede: Ganz kurz dazu. Kurz würde ich sagen, komplexe Verläufe lassen sich selten einfach modellieren. Hier ist aus meiner Sicht ganz entscheidend, dass viel zu fahrlässig mit der Zensierung umgegangen wird. Also, die Tatsache, dass die Beobachtung beendet wird, das generell als Zensierung zu modellieren, ist eben halt fahrlässig, wenn dieser Anlass der Beendigung der Beobachtung eben dazu führt, dass sich die Risiken oder die Ereignisraten für das eigentlich zu beobachtende Ereignis ändern. Und das scheint hier häufig der Fall zu sein. Es ist ja nicht so, dass wir uns nur einfach abwenden und die Nachverfolgung nicht weiter fortsetzen, sondern es hat einen Anlass gegeben, der auch ein Ausdruck ist für einen Grundzustand der Patientin oder des Patienten usw., der sich da geändert hat und der auch dann dazu führt oder vermutlich dazu führt, dass sich die Wahrscheinlichkeit für das Auftreten des Ereignisses, um das es da geht, ändert. Insofern muss man auf jeden Fall - das haben wir auch in unserer Stellungnahme deutlich gemacht -, wenn es um solche Zusammenhänge geht, adäquate Methoden verwenden. Einfach so eine Art Kaplan-Meier-Ereigniszeitanalyse ist da einfach in vielen Fällen nicht adäquat.

Moderator Thomas Kaiser: Vielen Dank. - Ralf, möchtest du vielleicht kurz etwas zu dem Punkt sagen?

Ralf Bender: Eigentlich nur, dass ich das voll bestätigen kann, dass es natürlich enorm wichtig ist, den Zensierungsgrund hier mit zu berücksichtigen und nicht einfach zu sagen, zensiert, und

so zu tun, als ob diese Zensierung sozusagen rein zufällig entstanden ist, was ganz klar nicht der Fall ist.

Moderator Thomas Kaiser: Genau. Was ja in einer Situation, sage ich mal, wo man - Sie haben eben den Punkt Modellierung angesprochen - so tut, als ob bestimmte Informationen gleichlautend wären oder wie auch immer wären aus Situationen, wo man sie nicht beobachtet hat, dann für die Interpretation heranzieht, genau da entsteht ja die Situation. Also, das ist bei der erstmaligen Veränderung sicher ein geringeres Problem, weil da häufig der Effekt schon sehr früh aufgetreten ist. Das ist der Punkt an der Stelle. Aber das bedeutet, was Sie sagen, dass man gegebenenfalls eben auch bei einer Überarbeitung von Modulvorlagen und einfach nur ... Wir befinden uns jetzt gerade hier in diesem Spezialthema Module für die frühe Nutzenbewertung, aber das Thema ist natürlich unabhängig von der Arzneimittelbewertung ganz generell relevant. Aber man muss sich gegebenenfalls überlegen, bei den Modulvorlagen zukünftig hier auch noch mal über andere Methoden bei solchen Verfahren und bei solchen Auswertungen nachzudenken und die auch zu beschreiben. - Herr Bauer.

Robert Bauer: Ich wollte noch mal auf das Thema „Estimands“ eingehen. Also, für mich ist das im Methodenpapier eine eindeutige Festlegung auf den Treatment Policy Estimand im Moment. Ich denke, das Thema „Estimands“ ist ja hervorragend dafür geeignet, das Ganze so ein bisschen ganzheitlicher abzudecken. In der EU-HTA-Guideline ist das ein bisschen breiter gefasst. Da steht, dass verschiedene Estimands eben auch relevant sein können für die Nutzenbewertung. Und ich wollte mal hören, inwieweit beim IQWiG in die Richtung gedacht wird, ob man bei der Festlegung auf den Treatment Policy Estimand bleibt oder da eben auch noch weiter forscht im Moment.

Moderator Thomas Kaiser: Jetzt machen Sie aber das ganz große Ding auf, was mit den Methoden 7 eigentlich nichts mehr zu tun hat.

Robert Bauer: Ja. Diese systematisch verkürzten Beobachtungszeiten machen ja nur Sinn, wenn man vom Treatment Policy Estimand ausgeht, weil da will man eben gerne die Daten erheben, egal, was über die Studie hinweg passiert, ob man jetzt die Behandlung beendet hat oder nicht.

Moderator Thomas Kaiser: Das würde mich auch interessieren.

Robert Bauer: Von daher macht das ja für mich auch Sinn, wenn man das zusammensetzt. Die Frage ist nur, ob dieser Treatment Policy Estimand tatsächlich auch immer der geeignete ist, über alle Endpunkte hinweg, über alle Indikationen hinweg usw. Deshalb würde ich es gerne in dem Fall zumindest mal als Frage aufbringen.

Moderator Thomas Kaiser: Ralf.

Ralf Bender: Ja, in der Tat sehen wir den Treatment Policy Estimand als den zentralen Estimand an, weil wir denken, dass der eben die Frage der Nutzenbewertung beantwortet. Wir haben aber nirgendwo im Methodenpapier andere Estimands ausgeschlossen. Und es gibt sicherlich Situationen, wo man auch andere Estimands betrachten kann, zumindest mal als ergänzende Analysen. Das ist nirgendwo im Methodenpapier ausgeschlossen. Alles, wozu wir im Methodenpapier nicht einen expliziten Abschnitt haben, ist ja jetzt nicht verboten. Also, alles, was sinnvoll ist, kann man natürlich auch tun, und das gilt auch für die Estimands. Aber in der Tat ist es so, den Treatment Policy sehen wir als den zentralen Estimand an, der die Frage der Nutzenbewertung primär beantwortet.

Moderator Thomas Kaiser: Stefan, du hast noch eine Ergänzung?

Stefan Lange: Es geht so ein bisschen in eine ähnliche Richtung. Vielleicht nur eine etwas andere Sichtweise. Natürlich können einen verschiedene Estimands interessieren, völlig unbenommen. Aber wenn einen der Treatment-Policy-Effekt interessiert, was durchaus legitim ist, dann braucht man die vollständige Beobachtung. Da beißt die Maus keinen Faden ab. Insofern ist die vollständige Beobachtung immer relevant, damit man eben im Zweifel, wenn einen das interessiert, diesen Treatment-Policy-Effekt schätzen kann. Wenn man daneben auch noch irgendwie diesen hypothetischen Estimand - ich habe den Namen vergessen - oder diese anderen Estimands, die man eigentlich gar nicht sehen kann, gerne schätzen will, dann soll man das doch machen. Aber noch mal: Wir brauchen grundlegend immer die vollständige Beobachtung, weil in aller Regel nicht nur wir, sondern auch viele andere, übrigens auch Patientinnen und Patienten und Verordnerinnen und Verordner in der Praxis, den Treatment-Policy-Effekt für eine Entscheidung benötigen.

Moderator Thomas Kaiser: Herr Ellis, ich habe Sie gesehen. Bevor ich Sie drannehme, möchte ich kurz einen Punkt noch verstärken.

Sie haben in einer Situation, wo Sie vollständig beobachtet haben, verschiedene Möglichkeiten der Auswertung, vollständige Beobachtung bis zur Progression, was auch immer. In einer Situation, wo Sie von Anfang an unvollständig beobachten, nehmen Sie sich Auswertungsmöglichkeiten weg. Und das ist gerade das, was passiert. Deswegen muss man an die vollständige Beobachtung ran, auch wenn man für verschiedene Zwecke durchaus unterschiedliche Auswertungen für relevant halten kann. Das kann man dann mit diesen vollständig beobachteten Daten alles machen.

Moderator Thomas Kaiser: Herr Ellis.

Aneurin Ellis: Nach den Estimands vielleicht die etwas simplere Frage: Was ist dann für Sie eine vollständige Beobachtung? Also, unvollständige Beobachtung ist für mich leicht zu erkennen, aber wann ist die Beobachtung für Sie vollständig? Das kann ja dann potenziell sehr,

sehr lange gehen. Dann frage ich mich auch im Hinblick auf die Estimands, ob dann wirklich immer nach zig Jahren alles noch so relevant ist.

Und zweitens auch im Hinblick auf die unterschiedlichen Beobachtungsdauern: Diese beiden Themenkomplexe sind für mich etwas vermischt worden, als ob die unterschiedlichen Beobachtungsdauern immer dadurch verursacht werden, dass wir unvollständig erheben. Dabei kann man ja durch eine verkürzte Erhebung leicht gleiche Beobachtungsdauern hinkriegen. Und wenn man dann über Jahre hinweg geht, dann wird es möglicherweise durch die Grunderkrankung schon unterschiedlich.

Moderator Thomas Kaiser: Ja, natürlich, Sie haben völlig recht. Oder was heißt, Sie haben völlig recht? Es ist sinnvoll, zwischen verschiedenen Situationen zu unterscheiden. Unser - ich nenne das jetzt einfach mal so - Hauptproblem derzeit in der Interpretation von vergleichenden Studien in der Onkologie, was jetzt diese patientenberichteten Endpunkte angeht, ist weniger, dass, sage ich mal, nicht zehn Jahre beobachtet wurde, sondern dass die Beobachtungsdauer zwischen den beiden Therapiearmen so deutlich unterschiedlich ist, weil eben das Beobachtungsende an ein Ereignis geknüpft wird, mit allen Folgen, auch den methodischen, die eben Herr Friede beschrieben hat. Und natürlich kann man, wenn man jetzt mal die Definition vollständig macht, sagen, vollständig bis zu dem Zeitpunkt, der einen interessiert. Man könnte zum Beispiel sagen, wir machen eine Studie, die läuft insgesamt mit einer Beobachtung von drei Jahren. Aber die patientenberichteten Endpunkte, die beobachten wir nur, aber vollständig, bis zwei Jahre, weil uns der Zeitpunkt zwei Jahre interessiert. Dann kann man eben sagen, wir sind ausreichend mit unseren Daten und auch vollständig für einen Zeitraum bis zwei Jahre und für andere Endpunkte vielleicht bis drei. Und da muss man individuell eben gucken: Ist das für eine Abwägung von Vor- und Nachteilen ausreichend oder ist das nicht ausreichend? Das hängt sicherlich von der konkreten Fragestellung ab. Wenn Sie daran interessiert sind, wie es eigentlich mit den patientenberichteten Endpunkten zwei Jahre nach Therapiebeginn ist, dann können Sie die Frage genau damit beantworten. Wenn Sie aber wissen wollen, wie es mit der Lebensqualität bis zum Versterben ist, dann können Sie die Frage damit nicht unbedingt beantworten, weil Sie ja gar nicht bis zum Versterben beobachten, sondern nach zwei Jahren aufhören. Dann haben Sie das methodische Problem gelöst, aber nicht das inhaltliche.

Insofern ist der Punkt, den Sie gemacht haben, sehr wichtig gerade, also nicht, dass man den Begriff „Vollständigkeit“ falsch versteht. Es gibt vollständig in Bezug auf Methodik. Das ist der wichtigste Punkt oder das wichtigste Problem, das wir gerade haben. Aber natürlich gibt es auch Vollständigkeit in Bezug auf Inhalt: Was interessiert mich eigentlich? Was ist meine Fragestellung? - Herr Rasch.

Andrej Rasch: Hier noch ergänzend. In unserer Stellungnahme hatten wir auch ein gewisses Problem, dass hier mit dem Begriff „systematisch verkürzte Beobachtungsdauer“ ziemlich

pauschal umgegangen wird. So erschien es. Sprich, es ist immer systematisch verkürzt, wenn nicht bis zum Lebensende, egal, was mit den Daten passiert, und egal, was die Hintergrunderkrankung ist, und weil es durchaus passieren kann, dass eine weitere Beobachtung auch gar keinen Sinn ergibt, weil da auch kein Mehrwert hinsichtlich der Interpretation mehr dazukäme. Mit dieser Pauschalität hatten wir ein Problem.

Moderator Thomas Kaiser: Vielen Dank. Das nehmen wir jetzt einfach noch mal mit. Ich bin nur immer ein bisschen zurückhaltend mit dem Begriff „kein Mehrwert“. Ich glaube, dass Betroffene durchaus darauf eine andere Sichtweise haben können.

Andrej Rasch: Also, kein Mehrwert im Sinne der methodischen Interpretation hinsichtlich des Verzerrungspotenzials, das meinte ich.

Moderator Thomas Kaiser: Das verstehe ich. Also, ich glaube, wenn ich das mal so mitnehmen darf, wir müssen noch mal draufgucken, ob wir diesen methodischen und inhaltlichen Bezug, den Sie gerade aufgemacht haben, mit Vollständigkeit, systematisch verkürzt da noch mal stärker zwischen differenzieren.

Gibt es zu dem Thema jetzt noch etwas, Daniela?

Daniela Preukschat: Ich habe noch eine weitere Frage. Wir haben grundsätzlich den Eindruck, dass in vielen Stellungnahmen gesehen wird, dass diese Beobachtung über Progress über Behandlungsabbruch hinweg inhaltlich sinnvoll ist. Es wird oft angeführt - ich nenne jetzt mal stellvertretend Pfizer und Amgen hier -, dass eine Beobachtung über Progress praktisch nicht möglich sei, nicht durchführbar sei. Dabei wird unter anderem auch mit einer zu großen Belastung für die Patienten argumentiert durch diese Beobachtung über Progress. Da würde sich jetzt unsere Frage anschließen, welche Daten Ihnen vorliegen, dass eine Beobachtung über Progress jetzt zwangsläufig mit einer erhöhten Patientenbelastung einhergeht.

In diesem Zusammenhang fanden wir eine von AbbVie vorgelegte Publikation sehr interessant, von Aiyegbusi, lautet der unaussprechliche Name. Die führt aus unserer Sicht recht differenziert an, dass es natürlich immer sinnvoll ist, weiter zu untersuchen, wie man die Belastung der Patientinnen und Patienten beim Beantworten des PROs reduzieren kann, also, dass man über eine klar definierte und kommunizierte Rationale reden muss, über den Einsatz von elektronischen Erfassungsmöglichkeiten für PROs. Man muss die Erhebungszeitpunkte durchdenken und sorgfältig wählen. Aber was in diesem Artikel auch betont wird, ist, dass die grundsätzliche Befürchtung, dass Patienten zu krank sind, um einen Fragebogen auszufüllen, eigentlich nicht durch Evidenz belegt ist und dass die Zurückhaltung, die daraus entsteht gegenüber PRO-Erhebungen, eher kontraproduktiv ist, um eigentlich sinnvolle Evidenz zu generieren. Deshalb an dieser Stelle noch mal die Frage: Haben Sie andere Daten,

Untersuchungen, dass grundsätzlich so eine Beobachtung über Progress nicht durchführbar ist bzw. mit einer für die Patienten nicht zumutbaren Belastung einhergeht?

Moderator Thomas Kaiser: Vielen Dank. - Die Betonung liegt auf „Daten“, bitte nicht, wie wir es schon erlebt haben, Polemik, „Sie wollen ja noch einen Fragebogen auf dem Sterbebett“. Das hilft uns hier allen nicht weiter. Welche Daten gibt es?

Andrea Friederich: Ich habe jetzt keine Publikation. Wir haben in unserer Stellungnahme Herrn Professor Wörmann zitiert, wo das Thema bis zur Lebenszeiterhebung auch in einer Anhörung besprochen wurde. Patienten im Progress haben einen langwierigen Krankheitsweg hinter sich. Sie wechseln die Therapie. Und die Frage, die sich uns stellt, ist: Muss man das Patienten wirklich zumuten? Wir sehen es auch in der akuten lymphatischen Leukämie zum Beispiel, dass die Rücklaufquoten bei den PRO-Erhebungen gering sind, weil eben die Patienten sehr krank sind. Und für uns ist es konträr, dass man einerseits so lange wie möglich erheben will, andererseits aber gleichzeitig Rücklaufquoten von über 70 % fordert. Das ist für uns einfach ein Widerspruch.

Moderator Thomas Kaiser: Vielen Dank. Jetzt haben Sie doch den Bezug gemacht. Ich muss ehrlich sagen, ich halte die Aussage, die da in der Anhörung gefallen ist, für sehr paternalistisch, und auch für wenig zuwendend zu Patientinnen und Patienten, auch zu der Forschung, die es gibt. Ich empfehle tatsächlich, diese Arbeit von Aiyegbusi, die von AbbVie in dem Verfahren vorgelegt wurde - ich hoffe, ich habe es richtig ausgesprochen -, mal zu lesen. Es geht eben darum, das bekommt man nicht von alleine. Man muss dafür eben investieren, dass man überzeugt, gute Daten macht, auch gucken, dass man das nicht zu komplex macht, dass man das situationsangepasst macht, trotzdem validiert. Aber da gibt es Möglichkeiten. Ich glaube, es macht wirklich Sinn, sich da auf den Weg zu machen, das zu verändern.

Stefan, du hast noch einen Punkt.

Stefan Lange: Ich möchte das gerne noch mal betonen. Es gibt Evidenz dafür, dass die regelmäßige und systematische Erhebung von patientenberichteten Endpunkten das Überleben der Betroffenen verlängert. Jetzt muss man ganz klar sagen: Wenn man das unterlässt, kann man es fast als einen Kunstfehler bezeichnen, um es mal auf einen anderen Punkt zu bringen.

Moderator Thomas Kaiser: Gut, dann wären wir, glaube ich, mit dem Punkt, Daniela, so weit. Noch Fragen? - Frau Denzer-Lippmann.

Melanie Denzer-Lippmann: Es geht darum, dass ich gerne eine Konkretisierung hätte, gerade jetzt aufgrund der PROs-Erhebungen. Wird es vielleicht dann auch in Zukunft mal eine konkretere Zahl geben, was Sie nach Progress-Erhebungen akzeptieren? Also, bis zum Tod, das ist schon relativ lange und manchmal auch wirklich, je nachdem, was das für eine Situation

ist, einfach nicht umsetzbar. Wenn man jetzt wirklich als PU gewillt ist, mehr Daten zu generieren, auch natürlich nach dem Progress, wird es da irgendwie eine angemessene Anzahl an Post-Progress-Erhebungen geben? Bis jetzt sehe ich da noch nichts unbedingt, was sich sehr klar abzeichnet, außer eben bis zum Tod, was natürlich dann schon irgendwie teilweise einfach nicht umsetzbar ist.

Moderator Thomas Kaiser: Sie sprechen gerade noch mal den Punkt an, den Herr Ellis eben aufgemacht hat, nämlich die Differenzierung zwischen einem methodischen Problem wegen - das nenne ich jetzt mal so - systematisch verkürzter Beobachtungsdauer in einem Arm und der Situation, dass Sie nur einen Teil - und das ist jetzt inhaltlich - der Gesamtstudiendauer beobachten, da aber eigentlich kein methodisches Problem haben, weil Sie - hoffentlich zumindest - bis zu einem bestimmten Zeitpunkt erheben, aber nicht, denn dann haben Sie das methodische Problem einfach nur verschoben, drei Erhebungen nach Progress. Dann haben Sie das methodische Problem nur verschoben. Also, das hängt ganz davon ab, welche Fragestellung Sie haben und welche Fragestellung Sie beantworten möchten. Sie entscheiden quasi mit Ihrer Erhebung in der Studie, welche Fragestellung Sie beantworten können. Wenn Sie nicht bis zum Lebensende erheben - und noch mal: bitte diese Polemik mit „auf dem Sterbebett“ weglassen -, dann werden Sie über die Lebensqualität der Patientinnen und Patienten bis zum Lebensende keine Aussage treffen können. Sie werden aber potenziell, abhängig davon, wie lang Ihre Beobachtungsdauer ist, in Bezug auf das mittlere Überleben etwas sagen können für den Beginn der Behandlung bis nahezu zum Versterben bzw. zum mittleren Versterben. Aber das hängt einfach von Ihrer Frage ab.

Der wichtige Punkt - ich kann es nur noch mal betonen -, der derzeit das Problem ist, ist die unterschiedliche Beobachtungsdauer. Der primäre Punkt ist nicht, dass nicht bis zum Versterben erhoben wird. Es wäre hilfreich, zumindest erst mal diese systematisch unterschiedlichen Beobachtungsdauern zu beenden.

Gut. Gibt es dazu noch Fragen? - Wenn das nicht der Fall ist, wären wir schon beim letzten TOP vor „Verschiedenes“.

3.4.7 Tagesordnungspunkt 6: Signifikanz und multiples Testen

Moderator Thomas Kaiser: Ralf.

Ralf Bender: Ich habe wieder nur eine kleine Verständnisfrage. Ich glaube zumindest, dass sie klein ist. Wir sind jetzt im Abschnitt 9.3.2, zur statistischen Signifikanz. Diesen Abschnitt hatten wir ja gekürzt und etwas umformuliert. Es gibt mehrere Stellungnahmen, die diese Kürzungen durchaus begrüßen. Aber es gibt auch ganz viele Stellungnahmen - im Grunde genommen fast alle Stellungnahmen aus der pharmazeutischen Industrie haben diesen Punkt aufgebracht - ... Deshalb geht meine Frage auch an alle Vertreterinnen und Vertreter aus der pharmazeutischen Industrie. Es wurde dort sehr betont, dass doch für ergänzende Analysen

und Sensitivitätsanalysen keine Adjustierung für multiples Testen gemacht werden solle. Offensichtlich gibt es hier Befürchtungen, dass das IQWiG so etwas plant. Meine Frage ist, ob diese Befürchtungen auf dem folgenden Satz aus diesem neuen Abschnitt beruhen - ich lese den mal vor -:

„Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an, z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.“

Interpretiert man diesen Satz so, dass wir zukünftig für ergänzende Analysen und Sensitivitätsanalysen die Anwendung von Verfahren zur Adjustierung für multiples Testen planen?

Moderator Thomas Kaiser: Konkrete Frage. Konkrete Antwort?

Ralf Bender: Oder ist es etwas anderes, wie man darauf kommt?

Moderator Thomas Kaiser: Herr Rasch.

Andrej Rasch: Ja, die konkrete Befürchtung - verzeihen Sie mir, wenn wir das falsch verstanden haben - wäre in der Tat, dass man das so interpretieren könnte - bleiben wir doch mal bei meinem Lieblingsendpunkt PASI -, wenn ich mehrere Operationalisierungen zum Endpunkt PASI vorlege, dass es dann zu einer Adjustierung hinsichtlich des multiplen Testens führen könnte. Ist das ausgeschlossen, ja oder nein?

Moderator Thomas Kaiser: Gut. - Ralf.

Ralf Bender: Wir planen nicht, für Sensitivitätsanalysen und ergänzende Analysen Verfahren für multiples Testen einzusetzen. Diese Aussage ist auch gar nicht neu. Die stand auch schon vorher im Methodenpapier. Da war sie etwas anders formuliert. Vorher stand da:

„Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an. Im Rahmen der Nutzenbewertung (siehe Abschnitt 3.1) versucht das Institut, den Fehler 1. Art bezüglich der Aussagen zu jedem einzelnen Nutzenaspekt getrennt zu kontrollieren.“

Der ist umformuliert worden. Das hat man offensichtlich nicht richtig verstanden. Wir werden jetzt überlegen, in welcher Art und Weise wir Missinterpretationen eher ausschließen können. Gemeint ist, wenn wir verschiedene Operationalisierungen zu einem Endpunkt haben, die man als gleichwertig betrachtet, dann wäre es nicht richtig, wenn man zum Beispiel acht verschiedene Operationalisierungen des gleichen Endpunkts hat, und bei einem dieser acht gibt es eine Signifikanz, dann zu sagen, okay, jetzt haben wir den Nutzen nachgewiesen. Das

wäre nicht korrekt. In einem solchen Fall müsste man für multiples Testen adjustieren. Das kommt in der Praxis aber eigentlich so gut wie nie vor. Nicht gemeint ist damit, man hat eine primäre Analyse, die zeigt irgendetwas, und man macht dazu für Robustheit ergänzende Analysen, Sensitivitätsanalysen. Dann ist natürlich eine Adjustierung für multiples Testen nicht erforderlich. Wir werden überlegen, inwieweit wir diesen Satz umformulieren, damit dieses Missverständnis beiseite geräumt wird.

Moderator Thomas Kaiser: Gut. Ralf, vielen Dank. - Gibt es dazu noch Klärungsbedarf, Fragen?
- Wenn das nicht der Fall ist, dann wären wir schon bei

3.4.8 Tagesordnungspunkt 7: Verschiedenes

Ein paar Punkte sind von Ihnen ja schon weitreichend - bis in die EU-Regulierung und in andere Stellen - gebracht worden. Aber gibt es noch andere Punkte, die Sie aus Ihren Stellungnahmen hier diskutieren wollen? Wie gesagt, wir haben alles gelesen, aber vielleicht gibt es ja noch einen Punkt. - Herr Sauerbruch.

Tilman Sauerbruch: Auf Seite 59 gehen Sie darauf ein, dass Studienergebnisse, wenn die Altersgruppe außerhalb der Zulassung ist, nicht berücksichtigt werden. Es kann aber sein, dass diese Medikamente ja gerade bei den Alten dann auch angewendet werden sollen oder werden. Und da wäre es doch sehr wichtig, wenn man vielleicht doch die Information hätte.

Moderator Thomas Kaiser: Vielen Dank. Dieser Satz ist nicht so gemeint, dass, wenn in einer Studie ältere Menschen nicht untersucht worden sind, wir dann sagen, grundsätzlich und die Grenze endet bei 68 und deswegen ist der Nutzen nur auf unter 68-Jährige zu übertragen, sondern dieser Punkt bezieht sich auf die Situation, dass es eben zum Beispiel Zulassungen im Jugendlichen- oder im Kindesbereich gibt, aber dass eben solche Patientinnen und Patienten komplett ausgeschlossen worden sind und dass - und das ist eben je nach Situation und Indikation, Kinder sind keine kleinen Erwachsenen - diese Ergebnisse nicht einfach so übertragbar sind. Der konkrete Fall, den Sie beschrieben haben, da gibt es ja in der Regel auch gar keine Zulassungseinschränkung. Also, Sie haben zwar vielleicht eine Studie durchgeführt, die ein mittleres Alter von 64 hat und maximal bis 70 eingeschlossen hat, aber die Zulassung schränkt das nicht auf 70-Jährige ein. Möglicherweise wird beschrieben, relativ wenig Daten, aber es gibt keine Zulassungseinschränkung. Also, ich glaube, dass deswegen der Punkt, den Sie machen, für Ihren konkreten Fall, den Sie geschildert haben, hier in der Formulierung keine Auswirkung hat.

Gibt es weitere Fragen? - Bitte sehr.

Katharina Musiolik: Ich wollte noch mal nachfragen, ob es noch in Planung ist, dass gegebenenfalls zum Hedges' g bzw. zur allgemeinen standardisierten Mittelwertsdifferenz, die ja auch angegeben werden soll theoretisch regelhaft bei kontinuierlichen Analysen,

insbesondere Pro- und Verlaufsanalysen, zumindest laut Modulvorlage, die ja nun etwas schwierig zu berechnen ist bzw. nicht so einfach darzustellen ist in Bezug auf Analysen, die mit MMRM durchgeführt werden sollen, ob da Ihrerseits gegebenenfalls noch mal eine Konkretisierung etc. oder ein Literaturhinweis im Methodenpapier geplant ist.

Moderator Thomas Kaiser: Wenn ich darauf antworten darf. Die Modulvorlagen liegen in der Hand des Gemeinsamen Bundesausschusses.

Katharina Musiolik: Ja, genau. Nicht in Bezug auf die Modulvorlage, sondern im Methodenpapier wird ja nun eben auch in Bezug auf die 15-%-Grenze und Responder-Analysen unter anderem auch die Möglichkeit dargestellt, kontinuierliche Analysen mit darzustellen und unter anderem da eben auch eine Angabe zu machen zur klinischen Bewertbarkeit, wo regelhaft normalerweise ja häufig der Hedges' g herangezogen wird, was eben bei MMRMs schwierig ist. Ist da in Zukunft gegebenenfalls noch mal von Ihrer Seite aus geplant, eine Konkretisierung oder eine Hilfestellung zu geben?

Moderator Thomas Kaiser: Ich muss gerade sagen, ich weiß nicht ganz genau, wo Ihre Frage hinzielt bzw. wo Sie den Konkretisierungsbedarf sehen. Können Sie vielleicht einfach mal ein Beispiel machen? Also, es gibt irgendwie ein Dossier oder was auch immer und da ist das und das gemacht worden, und dieses und jenes war für Sie schwierig. Also, ich habe tatsächlich noch nicht verstanden, wo Ihr Konkretisierungsbedarf liegt, mal jenseits der Frage, weil das möchten wir bitte trennen, was in Modulvorlagen gefordert ist.

Katharina Musiolik: Ja, auf jeden Fall, sehr gerne. Wie gesagt, es ist so, dass kontinuierliche Auswertungen nun eben ja häufig auf Grundlage von MMRMs durchgeführt werden und die Allgemeinen Methoden bei kontinuierlichen Auswertungen für die Relevanzbewertung eine standardisierte Mittelwertsdifferenz fordern. Aber eben gleichzeitig hatten Sie auch in einem Rapid Report mal die Aussage gemacht, dass bei MMRMs eben die gepoolte Standardabweichung schwierig bzw. nicht adäquat berechnet werden kann und deswegen nicht entsprechend herangezogen werden kann. Und der Hedges' g ist, wie gesagt, normalerweise eben das, was man als standardisierte Mittelwertsdifferenz bzw. als klinischen Effekt, sage ich mal, im Dossier dann eben darstellen würde, um zu versuchen, die klinische Relevanz darzustellen, zu bewerten, wenn man eben nicht über diese Responder-Analysen geht, auch wenn die natürlich vorzuziehen sind. Unseres Erachtens - das hatten wir auch in der Stellungnahme geschrieben - ist es eben da an der Stelle für uns nicht eindeutig klar, in welche Richtung das gehen soll, gerade weil der Rapid Report eben was anderes sagt. Ich meine, es war ein Rapid Report. Und da haben wir uns halt eben die Frage gestellt, und das ist zumindest ein Punkt, den man hier noch mal anbringen könnte, in die Richtung.

Moderator Thomas Kaiser: Vielen Dank. Den Rapid Report, auf den Sie jetzt referenzieren, haben Sie auch in Ihren Stellungnahmen zitiert. Richtig?

Katharina Musiolik: Ja, haben wir.

Moderator Thomas Kaiser: Gut, dann gucken wir uns das einfach noch mal genauer an. Aber, Guido, du hast dich ohnehin gemeldet.

Guido Skipka: Danke für die Frage. Die treibt uns auch um. Also Hedges' g ist erst mal nur für - ich nenne das mal so - t-Test-Situationen, also einen einfachen Zweigruppenvergleich definiert. Ein MMRM ist ein deutlich komplexeres Modell, und da ist nicht ganz klar, wie man da am besten einen standardisierten Effekt ausrechnet. Es gibt verschiedene Möglichkeiten. Man kann dann Annahmen treffen. Das ist meines Wissens noch nicht geklärt. Wir akzeptieren bis jetzt mehrere Möglichkeiten. Wenn wir die nachvollziehen können aus den Unterlagen, dann akzeptieren wir das. Aber vielleicht ein Aufruf an Herrn Friede für die biometrische Gemeinschaft. Das ist tatsächlich ein Problem, was man mal angehen könnte, wie man aus komplexeren Analysen ein Analogon zu Hedges' g ausrechnen kann.

Moderator Thomas Kaiser: Herr Friede, wollen Sie sich positionieren?

Tim Friede: Ich könnte Ihnen sagen, wie ich das mache. Also, ich nehme einfach die Mittelwertdifferenz aus dem Repeated-Measures-Modell und teile das dann zur Veranschaulichung einfach durch die gepoolte Standardabweichung, weil ich das nicht als Inferenz sehe, sondern einfach nur, um das ein bisschen einzuordnen mit anderen Dingen. Verglichen mit den anderen Problemen, die wir vorhin besprochen haben, ist das, glaube ich, ein kleineres, aber es ist ein interessantes. Das kann man sich sicherlich mal angucken.

Moderator Thomas Kaiser: Ralf, du hast dich auch noch gemeldet.

Ralf Bender: Nur eine kleine Ergänzung. Es gibt Vorträge davon von unserem Kollegen Lars Beckmann auf der GMDS und dem Biometrischen Kolloquium. Diese Vorträge kann man verwenden als Hilfestellung, welche Methoden möglich sind. Aber eine ganz klare Antwort gibt es halt im Moment nicht: Das ist jetzt der Goldstandard und das musst du immer machen. - Den gibt es aktuell leider nicht.

Moderator Thomas Kaiser: Außer den Goldstandard nach Friede, aber den haben wir ja eben gehört.

Vielen Dank. - Dann noch mal Herr Sauerbruch.

Tilman Sauerbruch: Das ist noch mal ein anderes Thema, auch auf Seite 19. Da geht es um die Frage der Anwendung von Arzneimitteln außerhalb der Fachinformationen. Ich meine, Drug repurposing spielt eine große Rolle. Das haben wir jetzt bei COVID-19 gesehen. Wie will das IQWiG mit diesen Daten umgehen?

Moderator Thomas Kaiser: Wir haben immer eine Fragestellung, die wir vom Gemeinsamen Bundesausschuss bekommen. In der Regel bekommen wir keine Fragestellung der Anwendung eines Arzneimittels, zumindest nicht eins, das primärer Bewertungsgegenstand ist außerhalb der Zulassung. Es kann schon sein, dass der Vergleichsmaßstab außerhalb der Zulassung ist. Das ist ja häufig in der Onkologie der Fall. Lassen wir das BSG-Urteil, das es gerade gegeben hat, mal außen vor, wie das da an der Stelle noch weitergehen wird. Aber angenommen, der Gemeinsame Bundesausschuss beauftragt uns mit einer Bewertung in genau dieser Konstellation, die Sie beschrieben haben, es gibt ein zugelassenes Arzneimittel, es gibt Studien der Anwendung dieses Arzneimittels außerhalb der Zulassung, zum Beispiel für eine pandemische Situation, und der G-BA bewertet uns damit, ja, dann ist das die Fragestellung, und wir gucken wir uns selbstverständlich die Studien auch an. Klar.

Gibt es weitere Fragen? - Wenn das nicht der Fall ist, dann danke ich ganz herzlich dafür, dass Sie so zahlreich gekommen sind, auch wirklich für die sehr wichtigen informativen, in allergrößten Teilen verständlichen, nur in wenigen Fällen polemischen - einen Punkt habe ich eben erwähnt; das war aber auch nur ein Zitat von einer anderen Person aus einem Anhörungsverfahren - Stellungnahmen. Wir werden uns das in Ruhe anschauen, was jetzt auch noch mal die Erörterung ergeben hat. Wir haben uns natürlich viele Dinge angeschaut.

Vielleicht für Sie zur zeitlichen Perspektive: Das wird noch einige Zeit in Anspruch nehmen. Wir gehen, glaube ich, davon aus, dass das so im vierten Quartal dieses Jahr fertiggestellt werden wird.

Gut. Dann wünsche ich Ihnen eine schöne Heimfahrt. Vielen Dank und tschüss.

Anhang A Dokumentation der Stellungnahmen

Inhaltsverzeichnis

	Seite
A.1 Stellungnahmen von Organisationen, Institutionen und Firmen.....	A 3
A.1.1 AbbVie Deutschland GmbH & Co. KG.....	A 3
A.1.2 Amgen GmbH	A 17
A.1.3 AMS - Advanced Medical Services GmbH.....	A 31
A.1.4 Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF)	A 39
A.1.5 Arbeitsgruppe Epidemiologische Methoden der DGEpi, GMDS & DGSMP für die Deutsche Gesellschaft für Epidemiologie (DGEpi e. V.)	A 46
A.1.6 Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ).....	A 50
A.1.7 Bayer Vital GmbH	A 56
A.1.8 Boehringer Ingelheim Pharma GmbH & Co. KG	A 73
A.1.9 Bristol-Myers Squibb GmbH & Co. KGaA	A 77
A.1.10 Bundesverband Medizintechnologie (BVMed) e. V.	A 88
A.1.11 Clinical Research platform Into molecular testing, treatment and outcome of non-Small cell lung carcinoma Patients (CRISP), Nationales Netzwerk Genomische Medizin, Lungenkrebs (nNGM), Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (DGHO)	A 94
A.1.12 co.value	A 98
A.1.13 Daiichi Sankyo Deutschland GmbH.....	A 106
A.1.14 Deutsche Gesellschaft für Gesundheitsökonomie e. V. (dggö).....	A 111
A.1.15 Deutsche Gesellschaft für Innere Medizin e. V. (DGIM).....	A 119
A.1.16 Deutsche Gesellschaft für Kardiologie – Herz- und Kreislaufforschung e. V. (DGK).....	A 127
A.1.17 Deutsche Gesellschaft für Orthopädie und Unfallchirurgie (DGOU), Deutsche Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC), Deutsche Gesellschaft für Unfallchirurgie (DGU).....	A 129
A.1.18 Deutsche Gesellschaft für Pharmazeutische Medizin e. V. (DGPharmMed) ...	A 132
A.1.19 Deutsche Krankenhausgesellschaft e. V. (DKG).....	A 136
A.1.20 Deutsches Netzwerk Versorgungsforschung e. V. (DNVF).....	A 147
A.1.21 Deutsche Region der Internationalen Biometrischen Gesellschaft (IBS-DR), Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V.	A 152
A.1.22 Ecker + Ecker GmbH	A 163
A.1.23 GlaxoSmithKline GmbH & Co. KG (GSK)	A 174

A.1.24 Institute of Public Health, Medical Decision Making and Health Technology Assessment, UM-IT Tirol	A 177
A.1.25 Ipsen Pharma GmbH	A 190
A.1.26 Janssen-Cilag GmbH	A 195
A.1.27 Kassenzahnärztliche Bundesvereinigung (KZBV)	A 206
A.1.28 Maimon Research LLC	A 210
A.1.29 Merck Healthcare Germany GmbH	A 230
A.1.30 MSD Sharp & Dohme GmbH	A 239
A.1.31 Novartis Pharma GmbH	A 244
A.1.32 Novo Nordisk Pharma GmbH	A 254
A.1.33 Pfizer Pharma GmbH	A 261
A.1.34 Roche Pharma AG	A 301
A.1.35 Sanofi-Aventis Deutschland GmbH	A 315
A.1.36 SmartStep Consulting GmbH	A 320
A.1.37 UCB Pharma GmbH	A 325
A.1.38 Verband Forschender Arzneimittelhersteller e. V. (vfa)	A 330
A.2 Stellungnahmen von Privatpersonen	A 350
A.2.1 Krones, Tanja	A 350

A.1 Stellungnahmen von Organisationen, Institutionen und Firmen

A.1.1 AbbVie Deutschland GmbH & Co. KG

Autorinnen und Autoren

- Heuck, Alexander
- Gothe, Henrike

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Dr. Alexander Heuck
Dr. Henrike Gothe
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: AbbVie Deutschland GmbH & Co. KG
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
3.3.4 „Konzepte für anwendungsbegleitende Datenerhebungen gemäß §35a Abs. 3b SGB V (AbD-Konzepte), S.66-68	<u>Anmerkung:</u> Der Abschnitt 3.3.4 zum Konzept für anwendungsbegleitende Datenerhebungen (AbD) gemäß §35a Abs. 3b SGB V wurde neu eingeführt. AbbVie begrüßt grundsätzlich, dass die vom IQWiG im AbD-Konzept behandelten Themen im Methodenpapier verankert wurden. Mit Blick auf die Anforderungen an die AbD besteht aus Sicht von AbbVie dennoch Konkretisierungs- bzw. Korrekturbedarf. Das Methodenpapier beschränkt sich auf die Beschreibung der Konzeptentwicklung. Die anzuwendenden Methoden und Qualitätsstandards werden nicht beschrieben, sondern verweisen in vielen Fällen auf den Rapid Report A19-43 (1), der ohne Stellungnahmeverfahren erstellt wurde. Die AbD hat eine große Relevanz für das Verfahren der Nutzenbewertung sowie ggf. für die Arzneimittelversorgung der Patientinnen und Patienten aufgrund der möglichen Beschränkung der Versorgungsbefugnis. Aus diesen Gründen sollten Methodik und Qualitätsstandards der AbD in das Methodenpapier aufgenommen und dem Prozess der Stellungnahme unterzogen werden. Dabei können die anzuwendenden Methoden und Qualitätsstandards auf Basis bisheriger Erfahrungen und im Hinblick auf eine erhöhte Praktikabilität im Versorgungsalltag überarbeitet werden.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	In Hinblick auf den Abschnitt zur Prüfung, ob bereits Datenplattformen existieren, die als Datenquelle für die AbD dienen können, möchte AbbVie auf die sehr hohen methodischen Anforderungen hinweisen, beispielsweise in Bezug auf die Datenqualität. Diese werden in der Realität wahrscheinlich von kaum einem der aktuellen Register in Deutschland erfüllt und es ist fraglich, ob sie in diesem Umfang erfüllt werden können. Das primäre Ziel vieler Register in Deutschland ist die Dokumentation der Diagnose und Behandlung eines bestimmten Krankheitsbildes im klinischen Alltag und nicht die vergleichende Datenerhebung zum Zwecke der Nutzenbewertung. Es ist folglich zu erwarten, dass die im Rapid Report A19-43 beschriebenen Anforderungen an die Datenqualität wie beispielsweise die Verwendung exakter Datumsangaben, die detaillierten Angaben zur medikamentösen Therapie [inkl. Dosis, Dosisänderung und Datumsangaben], die Erhebung aller relevanten Confounder, regelmäßige Schulungen sowie Source Data Verification oder interne/externe Audits (1) ohne umfangreiche Anpassungen der Registerstruktur sichergestellt werden können. Zudem verhindern die hohen Anforderungen die Nutzung retrospektiver Daten oder schränken diese zumindest stark ein. Diese Annahme wird durch die veröffentlichten AbD-Konzepte für Onasemnogen Abovirvec, Brexucabtagen Autoleucel, Risdiplam und Fedratinib gestützt (2-5). <i>Beschreibung des Umfangs der AbD:</i> Der Begriff des „akzeptablen Zeitraums“ sollte genauer erläutert werden. Ziel der AbD ist es die Evidenzlage für die Bewertung des Zusatznutzens eines Arzneimittels im Rahmen des AMNOG Prozesses zu verbessern. Um dieses Ziel nicht aus den Augen zu verlieren wäre es hilfreich, anzugeben, in welchem Zeitraum die AbD die geplanten Endpunkte erheben kann. Diese Information ist wichtig, um Aussagen zur Angemessenheit der Studiendauer – auch mit Blick auf den Zeitraum bis zum Verlust der Exklusivität oder die sich möglicherweise ändernde Versorgungslage (wie Änderung der zweckmäßigen Vergleichstherapie) – treffen zu können. Im abschließenden <i>Fazit</i> sollte die Einschätzung, ob die geplante AbD grundsätzlich realisierbar ist, weiter konkretisiert werden. Das IQWiG stellt an die AbD aufgrund des nicht-randomisierten Studiendesigns hohe Anforderungen an die Qualität und den Umfang der Datenerhebung. Unter anderem ist aufgrund der Anforderungen an die Effektstärke eine große Fallzahl

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	notwendig, um ausreichend große Effekte beobachten und daraus eine Aussage z.B. zum Nutzen eines Arzneimittels treffen zu können. Bei Indikationen bzw. Arzneimitteln, für die eine AbD gefordert werden kann, handelt es sich jedoch häufig um Indikationen mit einer begrenzten Anzahl an Patientinnen und Patienten, die für eine Datenerhebung zur Verfügung stehen. Zusätzlich kann die AbD mit weiteren Studien um Teilnehmerinnen und Teilnehmer konkurrieren, die aus den Auflagen der Zulassungsbehörde (Europäische Arzneimittel-Agentur) resultieren. Dies ist beispielsweise bei den bisherigen AbD-Verfahren für Onasemnogen Apeparvovec, Brexucabtagen Autoleucel und Risdiplam zu beobachten (2-4). Es ist zu erwarten, dass die AbD aufgrund der hohen Anforderungen als Folge methodischer Limitationen (u.a. Erwartungen an die Effektgrößen, Stichprobengröße und Dauer der Erhebung) mögliche Unterschiede nur bei wenigen Endpunkten und unter Inkaufnahme einer langen Studiendauer darstellen kann. Das Fazit sollte auf Basis der vorliegenden Erkenntnisse eine detaillierte Einschätzung enthalten, für welche Endpunkte die AbD in der beschriebenen Form Aussagen zur Nutzenbewertung erzielen kann und für welche nicht. Darüber hinaus sollte das Fazit auch detailliert auf den ggf. notwendigen Änderungs- und Erweiterungsbedarf eingehen, damit das jeweilige Register als primäre Datenquelle für die Registerstudie im Sinne der AbD dienen kann. Dies beinhaltet sowohl die notwendige Erweiterung der Dokumentation als auch die ggf. nötige Einbindung weiterer Zentren, um die erforderliche Rekrutierung sicherzustellen. AbbVie betrachtet beide Aspekte als sehr relevante Faktoren, um zu beurteilen, ob die geplante AbD grundsätzlich realisierbar ist. <u>Vorgeschlagene Änderung:</u> Die Methodik und die Qualitätsstandards zur Durchführung der AbD sollten in das Methodenpapier aufgenommen werden. Das abschließende Fazit sollte eine Einschätzung dazu enthalten, ob die vorgeschlagene AbD innerhalb des für den Zweck der Nutzenbewertung angemessenen Zeitraums realisierbar ist, inkl.: • Einer Zusammenfassung der Endpunkte, für welche auf Basis der vorliegenden Erkenntnisse und des Studiendesigns Unterschiede erwartet werden und für welche nicht.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	• Einer Beschreibung des Änderungs- und Erweiterungsbedarf der primären Datenquelle
9.1.1 „Kriterien für den Einschluss von Studien“, S.159	<u>Anmerkung:</u> Die Änderungen in Abschnitt 9.1.1 betreffen Studien, in denen nicht alle Patientinnen und Patienten die Einschlusskriterien bezüglich Population und Intervention erfüllen. Über die Berücksichtigung einer Studie soll nun sowohl auf Basis des quantitativen Anteils an Patientinnen und Patienten als auch unter Berücksichtigung einer potenziellen Effektmodifikation durch das betreffende Einschlusskriterium entschieden werden, falls geeignete Informationen vorliegen. AbbVie stimmt damit überein, dass Studien in denen mindestens 80% aller Patientinnen und Patienten die Einschlusskriterien erfüllen als für die Nutzenbewertung geeignet angesehen werden können. Eine zusätzliche Betrachtung der Effektmodifikation durch ein nicht erfülltes Einschlusskriterium ist bei mindestens 80% Erfüllung aller Einschlusskriterien nicht notwendig und nicht sachgerecht, da die Nicht-Zielpopulation nur wenige Patientinnen und Patienten umfasst. Die Stärke der Effektmodifikation kann somit nicht hinreichend sicher beurteilt werden und kann darüber hinaus irreführend sein (6, 7). Bei Erfüllung der Einschlusskriterien bei einem Großteil der prädefinierten Studienpopulation ist die Berücksichtigung der Gesamtstudie für die Nutzenbewertung aus Sicht von AbbVie somit sachgerecht. Falls weniger als 80% der Patientinnen und Patienten die Einschlusskriterien erfüllen, ist AbbVie der Ansicht, dass die Betrachtung der Gesamtstudie trotz Abweichungen bei den Einschlusskriterien als hinreichend sicher und geeignet angesehen werden kann, wenn die Stärke der Effektmodifikation für die zu betrachtenden Endpunkte nicht maßgeblich ist. Die Berücksichtigung von Studien wird zudem befürwortet, wenn trotz Abweichungen bei den Einschlusskriterien der medizinische Versorgungskontext hinreichend sicher plausibel durch die Gesamtstudie abgedeckt wird. Dies hat den Vorteil, dass die Studie wie prädefiniert ausgewertet wird und die statistische Power gemäß der ursprünglichen Studienplanung erhalten bleibt. Bei größeren Abweichungen der Studienpopulation gegenüber der Zielpopulation kann die von IQWiG vorgeschlagene Methodik der Anhebungsregel „AHR₁₅“ angebracht sein. Das Verfahren zur Übertragbarkeit von Studienergebnissen auf Teilpopulationen wurde vom IQWiG untersucht und bei Einhaltung einer

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	definierten Testprozedur als geeignet eingestuft (8). Leider ist eine Akzeptanz dieses Verfahrens in der Praxis kaum gegeben. <u>Vorgeschlagene Änderung:</u> AbbVie ist der Ansicht, dass bei mindestens 80% Erfüllung der Einschlusskriterien eine zusätzliche Bewertung von Effektmodifikationen durch Einschlusskriterien nicht notwendig und nicht sachgerecht ist. Zudem wird vorgeschlagen, dass die Methodik der Anhebungsregel „AHR₁₅“ als geeignetes Verfahren zur Übertragbarkeit von Studienergebnissen auf Teilpopulationen im Methodenpapier gewürdigt wird.
9.3.2 „Statistische Signifikanztests und Konfidenzintervalle“, S.170	<u>Anmerkung:</u> Im Abschnitt 9.3.2 wird mit dem Entwurf der Version 7.0 des Methodenpapiers das Vorliegen mehrerer Operationalisierungen für einen Endpunkt exemplarisch als Möglichkeit aufgeführt, bei welcher die Methoden der Adjustierung für multiples Testen Anwendung finden könnte. Eine präferiert anzuwendende Methode wird hierbei nicht definiert. Aus Sicht von AbbVie sollte auch weiterhin keine generelle Adjustierung für multiples Testen im Rahmen der frühen Nutzenbewertung Anwendung finden. Je nach Wahl der Methode könnten unterschiedliche Bewertungen des Zusatznutzens für den entsprechenden Endpunkt entstehen, was unter anderem die Vergleichbarkeit zwischen einzelnen Verfahren gefährdet. Zudem wird bei einzelnen Methoden, wie dem hierarchischen Testen, eine Präspezifizierung der Testsequenz benötigt (9, 10). Folglich hängt die statistische Signifikanz einzelner Operationalisierungen von den Ergebnissen anderer ab. Inwiefern eine solche Sequenz im Rahmen der frühen Nutzenbewertung für alle durch die Modulvorlage erforderlichen Endpunkte präspezifiziert werden kann oder soll, bleibt fraglich. Zusätzlich sind neben den im statistischen Analyseplan (SAP) definierten Endpunkt-Operationalisierungen oftmals weitere für die frühe Nutzenbewertung relevante Zusatzanalysen unabdingbar, um den (Zusatz-)Nutzen eines Medikaments bewerten zu können. Daraus ergibt sich, dass solche Analysen post hoc den AMNOG-Anforderungen entsprechend durchgeführt werden. Es wird daher mehr als eine Operationalisierung in dem Nutzendossier dargestellt, um neben den AMNOG-spezifischen und geforderten Analysen auch die präspezifizierten Analysen des SAP transparent darzustellen. AbbVie begrüßt dies auch weiterhin entsprechend zu würdigen und auf Endpunktebene das bisherige Vorgehen einer getrennt voneinander stattfindenden Betrachtung unterschiedlicher

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Operationalisierungen beizubehalten. Es sollten die für die Patientinnen und Patienten relevanten Operationalisierungen zur Ableitung des Zusatznutzens herangezogen werden. Weitere Operationalisierungen dienen an dieser Stelle der Robustheit der Ergebnisse und werden im Allgemeinen ergänzend dargestellt. Ein multipler Hypothesentest im ursprünglich definierten Sinn liegt folglich nicht vor. <u>Vorgeschlagene Änderung:</u> In Nutzenbewertungen ist es aufgrund der Anforderungen an die Ableitung eines Zusatznutzens basierend auf für die Patientinnen und Patienten relevanten Ergebnissen unausweichlich, neben den präspezifizierten Analysen weitere Operationalisierungen darzustellen. Diese sollen den für die Patientinnen und Patienten spürbaren und relevanten Nutzen des neuen Medikaments abbilden. Es sollte weiterhin das qualitative Vorgehen bei der Formulierung der Gesamtaussage zum Zusatznutzen auch für die Bewertung eines durch verschiedene Operationalisierungen untersuchten Endpunktes Anwendung finden. Folglich ist keine quantitative Bewertung, sondern vielmehr eine unabhängige Betrachtung der dargestellten Operationalisierungen empfehlenswert. Insbesondere eine Adjustierung für multiples Testen sollte im Rahmen der Nutzenbewertung gemäß dem bisherigen Vorgehen generell keine Anwendung finden. Dies sollte aus dem Abschnitt entsprechend ersichtlich werden.
9.3.12 „Umgang mit variablen Beobachtungszeiten“, S.190-191	<u>Anmerkung:</u> Der Abschnitt 9.3.12 wurde durch den Umgang mit variablen Beobachtungszeiten bei der Erhebung patientenberichteter Endpunkte (PRO) erweitert. Dabei wird insbesondere die verkürzte Beobachtung von PRO sowie die Operationalisierung der „dauerhaften Veränderung“ bei unterschiedlichen Beobachtungszeiten kritisiert. Weiterhin verlangt das IQWiG die Auswertung auf Basis des <i>Treatment Policy Estimands</i>. Zudem wird durch das IQWiG angemerkt, dass die Ergebnisse der Endpunkte mit „systematisch verkürzten“ Daten die Gesamtaussage zum Zusatznutzen nicht dominieren können. Methodisch ist die Kritik an der „dauerhaften“ Veränderung bei unterschiedlichen Beobachtungszeiten grundsätzlich nachvollziehbar. Trotzdem kann je nach Indikationsgebiet und Fragestellung die Auswertung einer nachhaltigen Veränderung inhaltlich sinnvoll sein. Zudem hat auch die Auswertung der „Zeit bis zur erstmaligen Veränderung“ von PRO methodische und konzeptionelle Nachteile. Zum einen wird nur ein Bruchteil der verfügbaren Daten in der Analyse angemessen berücksichtigt, da Werte nach der erstmaligen Veränderung nicht betrachtet

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	werden. Zum anderen kann nicht festgestellt werden, ob eine Veränderung nachhaltig für Patientinnen und Patienten spürbar und somit besonders relevant ist. Dies ist vor allem problematisch, wenn sich eine erste Veränderung nur kurzfristig einstellt, jedoch im weiteren Verlauf der Behandlung die gegenteilige Veränderung zu beobachten ist. Es sollte daher im Einzelfall unter Berücksichtigung der Fragestellung und Indikation die angemessene Operationalisierung gewählt werden können und ggf. durch Sensitivitätsanalysen und eine inhaltliche Einordnung ergänzt werden. Patientenberichtete Zielgrößen sind von besonderem Stellenwert für die Nutzenbewertung von Arzneimitteln, da sie, wie in § 5 Abs. 7 der AM-NutzenV beschrieben, die vom Patienten spürbare Linderung der Erkrankung erfassen. Daher sollte PRO auch weiterhin ein hoher Stellenwert in der Gesamtbewertung eingeräumt werden. Die Kritik des IQWiG zu „systematisch verkürzten“ Beobachtungszeiten und der damit verbundenen potenziell untergeordneten Bedeutung von PRO in der Nutzenbewertung ist für AbbVie nicht nachvollziehbar und beeinträchtigt die Bedeutung von PRO zum Nachteil der Patientinnen und Patienten. Die frühe Nutzenbewertung von Arzneimitteln sollte die in die Therapiesequenz eingebettete medizinische Intervention bewerten und weniger eine Abfolge an Interventionen im Sinne eines sequenziellen, medizinisch-therapeutischen Behandlungspfades mit unterschiedlichen aufeinanderfolgenden Interventionen, die jeweils verschiedene Behandlungseffekte hervorbringen. Es bleibt unklar, zu welchem Ausmaß bei einer langfristigen Erhebung von PRO über verschiedene therapeutische Interventionen hinweg der eingetretene Effekt auf die ursprünglich zu betrachtende Intervention zurückzuführen ist. Zudem bleibt offen, ob eine Beobachtung beispielsweise bis zum Tod aus logistischen, technischen, rechtlichen und ethischen Aspekten der Studienplanung und -durchführung möglich, sinnvoll und/oder praktikabel ist. Die Erhebung von PRO sollte nicht zum Selbstzweck werden, sondern im Verhältnis zur Sinnhaftigkeit und Interpretierbarkeit sowie unter Berücksichtigung der Patientenbelastung stehen (11, 12). Zuletzt steht die Forderung nach verlängerten Erhebungszeiträumen potentiell im Zielkonflikt zu den hohen qualitativ-methodischen Kriterien für die Auswertbarkeit der Analysen, insbesondere bezogen auf die Rücklaufquoten, die bei fortlaufender Erhebungszeit mit einer Abnahme einhergehen (13). Eine verminderte Relevanz für Effekte aus patientenberichteten Erhebungen ist auf Basis der vom IQWiG angeführten Argumentation zu „systematisch verkürzten Beobachtungszeiten“ somit nicht nachvollziehbar. Der Wert

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	dieser inhärent patientenrelevanten Endpunkte würde dadurch herabgestuft. <u>Vorgeschlagene Änderung:</u> Aus Sicht von AbbVie bedarf die Auswahl der Operationalisierung von PRO bei unterschiedlichen Beobachtungszeiten einer individuellen Abwägung, die methodische als auch inhaltliche Aspekte berücksichtigt. Eine pauschale Festlegung auf die „Zeit bis zur erstmaligen Veränderung“ als einzig relevante Operationalisierung ist nicht sachgerecht. Stattdessen sollte im Einzelfall unter Berücksichtigung der Fragestellung, Indikation, sowie einer inhaltlichen Einordnung entschieden werden. Der Stellenwert von patientenberichteten Zielgrößen darf zudem nicht durch eine inadäquate Kritik „systematisch verkürzter“ Beobachtungszeiten eingeschränkt werden. Eine Beobachtung nach Behandlungsende kann je nach Indikationsgebiet und Therapie wichtig sein, jedoch muss die Dauer im Verhältnis zur Sinnhaftigkeit, Realisierbarkeit und ethischen Aspekten stehen.
4 „Kosten-Nutzen Bewertung medizinischer Interventionen“, S.83ff.	<u>Anmerkung:</u> Im Entwurf des Methodenpapiers stellt das IQWiG einen grundlegend überarbeiteten Abschnitt zur Gesundheitsökonomie vor. Insgesamt dienen die vom IQWiG vorgeschlagenen Änderungen im Bereich der Kosten-Nutzen-Bewertung (KNB) zwar einem wissenschaftlichen Interesse, sie haben jedoch keine praktische Relevanz in der Umsetzung. Denn seit 2011 hat sich in Deutschland das auf der Bewertung des patientenrelevanten Wertes (Nutzen) basierte Verfahren der frühen Nutzenbewertung etabliert. Das Instrument der KNB bringt im Gegensatz dazu vor allem im Bereich der Ethik kontroverse und sensible Fragen mit sich. Vor dem Hintergrund verbleibender ethischer, methodischer und wissenschaftlicher Fragen erachtet es AbbVie als zielführender, bereits in der Anwendung befindende Fragestellungen der klinischen Nutzenbewertung, etwa die Surrogatendpunktvalidierung oder die Bewertung von Responseschwellen, weiterzuentwickeln. Die überarbeitete Version des Kapitels 4 lehnt sich vermehrt an internationalen methodischen Standards der Gesundheitsökonomie an. Eine wesentliche Änderung betrifft das in früheren Versionen beschriebene Konzept der Effizienzgrenze. Die (potentielle) Anwendung einer linearen Extrapolation einer indikationsspezifischen Effizienzgrenze war mit methodischen Problemen verbunden und in der Vergangenheit methodischer Kritik von Experten(-gruppen) unterworfen (14-16). Der neue Entwurf der Allgemeinen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Methoden zeigt eine nachvollziehbare Abkehr vom Effizienzgrenzenkonzept für KNB von Arzneimitteln gemäß § 35b SGB V (Kapitel. 4.2). Das Konzept kann gem. Kapitel 4.15 aber weiterhin für gesundheitsökonomische Evaluationen außerhalb § 35b SGB V angewendet werden. Die nun vorgeschlagenen Vorgehensweisen zeigen trotz dieser Weiterentwicklung weitere methodische und technische Defizite auf, die nachfolgend exemplarisch erläutert werden. Das IQWiG führt die KNB nach der in Auftrag (z.B. des G-BA) gegebenen definierten Fragestellung aus. Der Auftraggeber legt dabei alle relevanten Inhalte, wie beispielsweise die Population, die Wahl der Intervention und des Komparators bzw. der Komparatoren, die Wahl der Zielgrößen, den Zeithorizont o.Ä. fest. Eine fixe Vorgabe aller Parameter der KNB durch den Auftraggeber ist dabei kritisch zu sehen: Auch wenn eine Benennung der Fragestellung und die Definition einer Reihe methodischer Mindeststandards (wie bei der Nutzenbewertung) notwendig sind, um Transparenz und Vergleichbarkeit zu gewährleisten, bleibt das methodische Vorgehen zur KNB von der jeweiligen Frage- bzw. Zielstellung, Indikation etc. abhängig und ist daher im Einzelfall zu begründen. Das Methodenpapier enthält bei der Beschreibung des Referenzfalles sowie den Ausführungen zur Basis- und Szenarioanalyse einige konkrete Festlegungen auf (Prozent)werte oder Spannen, welche unter Umständen zielgerichteter auf den Einzelfall, die Indikation und die spezielle Fragestellung adjustiert werden sollten. So wird gemäß Kapitel 4.3.5 bei Szenarioanalysen regelhaft ein Zeithorizont von 5 Jahren betrachtet. Jedoch sollte auch für Szenarioanalysen gelten, dass „der Zeithorizont der Analyse allen relevanten Auswirkungen auf nachfolgende Kosten- und Ergebnisgrößen gerecht wird“ (17-19). Eine Festlegung ist in diesem Kontext als zumindest unzureichend wissenschaftlich begründet anzusehen und die Wahl des Zeithorizonts sollte aus Sicht von AbbVie begründet für die spezifische Indikation und im Rahmen einer Einzelfallbetrachtung erfolgen. Die Wahl des Diskontierungssatzes mit 3% für Kosten und Effekte im Basisfall liegt zwar innerhalb einer Spanne, welche von internationalen Agenturen regelhaft angewandt wird, weicht jedoch von den deutschen Empfehlungen zur gesundheitsökonomischen Evaluation ab (18, 20). Auf S.85 und S.88 wird eine erneute Nutzenbewertung für die Erstellung einer KNB beschrieben. Hierbei bleibt weitestgehend unklar, unter welchen Umständen diese durchgeführt werden soll

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	und welche Interventionen als „weitere Komparatoren“ zu verstehen sind. Aus Sicht von AbbVie ist in diesem Kontext auch die gesetzliche Grundlage unklar, auf der eine neue Nutzenbewertung spezifisch für KNB mit „Vergleichstherapien, zusätzlich zur gewählten zweckmäßigen Vergleichstherapie“ durchgeführt werden kann. Aus Sicht von AbbVie sollten daher die Abschnitte zur Nutzenbewertung weiterer Komparatoren aus dem Methodenpapier entfernt werden (siehe S.85 und S.88). In Version 6.1 des Methodenpapiers befand sich unter Kapitel 1.4.3 ein separater Abschnitt mit einer Würdigung des Quality-adjusted life years (QALY) - Konzeptes als Maß des Gesamtnutzens, um darauf basierend allokativen und distributiven Entscheidungen für Gesundheitsleistungen abzuleiten. Zudem enthielt die vorhergehende Version des Methodenpapiers ebenfalls einen Hinweis, der bei Verwendung des QALY-Konzeptes auch ethische, rechtliche und kulturellen Aspekte beleuchtete (21). In Version 7.0 findet das QALY-Konzept nun als Maß des Gesamtnutzens für die Kosten-Nutzwert Analyse weitestgehend unkritisch Eingang in den Referenzfall einer KNB. Hinsichtlich der Interpretation von Ergebnissen gesundheitsökonomischer Analysen sei an dieser Stelle nur am Rande auf die Notwendigkeit verwiesen, diese „(...) unter besonderen Umständen, etwa für die „Medikation seltener Erkrankungen“ (...), einer weiteren ethischen und sozialen Prüfung zu unterziehen“ (22). <u>Vorgeschlagene Änderung:</u> - Einfügen eines allgemeinen Abschnitts zur kritischen Diskussion des QALY-Konzeptes (vormals Kap. 1.4.3). - Festlegung und Begründung der Parameter für die KNB unter individueller Betrachtung der Fragestellung und Indikation [z.B. keine Vorgabe eines festgelegten Zeithorizontes für Szenarienanalysen (S.85 und S.90)]. - Einfügen einer Diskussion bzw. wissenschaftlicher Rationale für die Wahl des Diskontierungssatzes im Basisfall unter Berücksichtigung internationaler und nationaler Empfehlungen (S.85). - Entfernen der Formulierung, dass ggf. eine erneute Nutzenbewertung unter Einbeziehung weiterer Komparatoren erforderlich ist (S. 85 und S. 88).

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. Insitut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). IQWiG-Berichte – Nr. 863, Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V, Rapid Report, Auftrag: A19-43, Version: 1.1, Stand: 13.05.2020. Verfügbar unter: https://www.iqwig.de/download/a19-43_versorgungsnaher-daten-zum-zwecke-der-nutzenbewertung-rapid-report_v1-1.pdf. [Zugriffsdatum: 13.01.2023].
2. Insitut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). IQWiG-Berichte – Nr. 977, Konzept für eine anwendungsbegleitende Datenerhebung – Onasemnogen-Abeparvovec, Rapid Report, Auftrag: A20-61, Version: 1.0, Stand: 01.10.2020. Verfügbar unter: https://www.iqwig.de/download/a20-61_anwendungsbegleitende-datenerhebung-onasemnogen-abeparvovec-rapid-report_v1-0.pdf. [Zugriffsdatum: 12.01.2023].
3. Insitut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). IQWiG-Berichte – Nr. 1323, Konzept für eine anwendungsbegleitende Datenerhebung – Brexucabtagen autoleucel, Rapid Report, Auftrag: A21-130, Version: 1.0, Stand: 31.03.2022. Verfügbar unter: https://www.iqwig.de/download/a21-130_anwendungsbegleitende-datenerhebung-brexucabtagen-autoleucel-rapid-report_v1-0.pdf. [Zugriffsdatum: 12.01.2023].
4. Insitut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). IQWiG-Berichte – Nr. 1294, Konzept für eine anwendungsbegleitende Datenerhebung – Risdiplam, Rapid Report, Auftrag: A21-131, Version: 1.0, Stand: 15.02.2022. Verfügbar unter: https://www.iqwig.de/download/a21-131_anwendungsbegleitende-datenerhebung-risdiplam-rapid-report_v1-0.pdf. [Zugriffsdatum: 12.01.2023].
5. Insitut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). IQWiG-Berichte – Nr. 1384, Konzept für eine anwendungsbegleitende Datenerhebung – Fedratinib, Rapid Report, Auftrag: A21-142, Version: 1.0, Stand: 29.06.2022. Verfügbar unter: https://www.iqwig.de/download/a21-142_anwendungsbegleitende-datenerhebung-fedratinib-rapid-report_v1-0.pdf. [Zugriffsdatum: 12.01.2023].
6. Wang S-J, Hung HMJ. A Regulatory Perspective on Essential Considerations in Design and Analysis of Subgroups When Correctly Classified. Journal of Biopharmaceutical Statistics. 2014;24(1):19-41.
7. Brookes ST, Whitely E, et al. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. J Clin Epidemiol. 2004;57(3):229-36.
8. Insitut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). IQWiG-Berichte – Nr. 638, Untersuchung der statistischen Eigenschaften von Verfahren zur Übertragbarkeit von Studienergebnissen auf Teilpopulationen, Arbeitspapier, Auftrag: GA18-01, Version: 1.0, Stand: 20.06.2018. Verfügbar unter: https://www.iqwig.de/download/ga18-01_uebertragbarkeit-von-studienergebnissen-auf-teilpopulationen-arbeitspapier_v1-0.pdf. [Zugriffsdatum: 04.01.2023].

9. Gou J, Xi D. Hierarchical Testing of a Primary and a Secondary Endpoint in a Group Sequential Design With Different Information Times. *Statistics in Biopharmaceutical Research*. 2019;11(4):398-406.
10. Alosch M, Bretz F, et al. Advanced multiplicity adjustment methods in clinical trials. *Statistics in Medicine*. 2014;33(4):693-713.
11. Aiyegbusi OL, Roydhouse J, et al. Key considerations to reduce or address respondent burden in patient-reported outcome (PRO) data collection. *Nature Communications*. 2022;13(1):6026.
12. European Medicines Agency (EMA). Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man, The use of patient-reported outcome (PRO) measures in oncology studies 2016. Verfügbar unter: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf. [Zugriffsdatum: 04.01.2023].
13. Neve OM, van Benthem PPG, et al. Response rate of patient reported outcomes: the delivery method matters. *BMC Medical Research Methodology*. 2021;21(1):220.
14. Insitut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden 4.2, Version 1.0 vom 22.04.2015 2015. Verfügbar unter: https://www.iqwig.de/methoden/dwa_methoden_4_2_pdf.pdf. [Zugriffsdatum: 04.01.2023].
15. Greiner W, Kuhlmann A, et al. Ökonomische Beurteilung des Effizienzgrenzenkonzeptes. *Gesundheitsökonomie & Qualitätsmanagement*. 2010;15(05):241-50.
16. Schöffski O, Schumann A, et al. Das Schwellenwertkonzept: Theorie sowie Umsetzung beim IQWiG und anderen Institutionen. In: Schöffski O, Graf von der Schulenburg JM, editors. *Gesundheitsökonomische Evaluationen*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 111-53.
17. Schöffski O, Schulenburg JM. *Gesundheitsökonomische Evaluationen*: Springer Berlin Heidelberg; 2013.
18. Graf von der Schulenburg JM, Greiner W, et al. Deutsche Empfehlungen zur gesundheitsökonomischen Evaluation - dritte und aktualisierte Fassung des Hannoveraner Konsens. *Gesundheitsökonomie & Qualitätsmanagement*. 2007;12(05):285-90.
19. Drummond MF, Sculpher MJ, et al. *Methods for the Economic Evaluation of Health Care Programmes*: Oxford University Press; 2005.
20. Attema AE, Brouwer WBF, et al. Discounting in Economic Evaluations. *PharmacoEconomics*. 2018;36(7):745-58.
21. Insitut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden, Version 6.1 vom 24.01.2022. Verfügbar unter: <https://www.iqwig.de/methoden/allgemeine-methoden-v6-1.pdf>. [Zugriffsdatum: 06.01.2022].
22. Damm K, v. d. Schulenburg JMG. 4 Gesundheitsökonomische Evaluation und Verteilungsgerechtigkeit. In: Schöffski O, Graf von der Schulenburg JM, editors. *Gesundheitsökonomische Evaluationen*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 501-20.

|

A.1.2 Amgen GmbH

Autorinnen und Autoren

- Friederich, Andrea
- Stein, Alexandra
- Floßmann, Carmen
- Schrom, Eva
- Bartsch, Robert
- Camdere, Hatice

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Friederich, Andrea; Director Health Economics & Outcomes Research
Stein, Alexandra; Health Economics & Outcomes Research Manager
Dr. Floßmann, Carmen; Health Economics & Outcomes Research Senior Manager
Dr. Schrom, Eva; Health Economics & Outcomes Research Manager
Bartsch, Robert; Health Economics & Outcomes Research Manager
Camdere, Hatice; Executive Director Value, Access & Policy
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Amgen GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Amgen begrüßt, dass das IQWiG im Entwurf für das Methodenpapier in der Version 7.0 die evidenzbasierte Gesundheitsversorgung (Evidence-based Health Care [EbHC]) und die evidenzbasierte Medizin (Evidence-based Medicine [EbM]) als grundlegende Prinzipien für das deutsche Gesundheitssystem hervorhebt.

Die EbM fordert eine medizinische Versorgung **auf Grundlage der zu dem Zeitpunkt besten zur Verfügung stehenden Evidenz**, wobei randomisierte kontrollierte Studien (randomized controlled trials [RCTs]) die höchste Aussagefähigkeit innehaben. Sie bietet zudem Instrumente, um Unsicherheit einzuschätzen, wenn die bestverfügbare Evidenz aufgrund von Einschränkungen durch interne oder externe Faktoren mit Unsicherheiten behaftet ist. So ermöglicht sie auch in diesen Fällen evidenzbasierte Entscheidungen zu treffen.

Dies ist insbesondere im Kontext der Bewertung von Studienergebnissen zu bspw. multimorbiden Patientinnen und Patienten oder seltenen Erkrankungen mit sehr kleinen Patientenpopulationen von Relevanz (Abrahamyan et al. 2016, Sherian et al. 2016). In derartigen Situationen ist eine RCT nicht immer möglich und / oder ethisch durchführbar, sodass oft Studien mit niedrigerem Evidenzlevel, wie beispielsweise nicht-vergleichende Studien, ggf. mit geringer Teilnehmerzahl, die **bestverfügbare** Evidenz darstellen. Um auch in diesen Fällen eine evidenzbasierte Gesundheitsversorgung für alle Patientinnen und Patienten zu gewährleisten, sollten derartige Studien grundsätzlich in Betracht gezogen werden und als Datengrundlage für die frühe Nutzenbewertung nach § 35a SGB V dienen.

Auf weitere Änderungen im Methodenpapier wird nachfolgend Stellung genommen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 1.3.3 (S. 7)	<u>Anmerkung:</u> Das IQWiG definiert als gesundheitliche Frage zur Beurteilung des Nutzens bzw. Schadens einer medizinischen Maßnahme die Reduktion „ <i>beeinträchtigende[r] Nebenwirkungen</i> “. Allerdings ist die Beeinträchtigung von Patientinnen und Patienten subjektiv und heterogen. Zur besseren Beurteilung von Nebenwirkungen und ihrer Beeinträchtigung für Patientinnen und Patienten unterstützen patientenberichtete Erhebungsinstrumente die Beurteilung, wie

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	bspw. der Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) (Basch et al. 2014, Dueck et al. 2015). <u>Vorgeschlagene Änderung:</u> Der Begriff „beeinträchtigende Nebenwirkung“ sollte detaillierter erläutert werden – bspw. worauf bezieht sich die Beeinträchtigung (tägliche Arbeiten, Einfluss auf das soziale Umfeld etc.)?
Abschnitt 1.4 (S. 10 f.)	<u>Anmerkung:</u> <i>„Ziel der KNB ist es, ökonomische Informationen als Zusatz zur Nutzenbewertung insbesondere für die Preisverhandlungen im Sinne einer Informationssynthese zusammenzufassen. [...] Ziel [...] ist nicht die Ableitung von Aussagen über den Zusatznutzen [...], [denn es] kann nicht die Aussagesicherheit erreicht werden, die für einen Beleg, Hinweis oder Anhaltspunkt notwendig ist.“</i> Die KNB ist nicht Teil der Erstattungsbetragsverhandlungen und der entsprechenden Rahmenvereinbarung nach § 130b Abs. 9 SGB V. Diese kann basierend auf § 35b SGB V als Orientierung bei der Preisfindung dienen, wenn die entsprechenden Voraussetzungen erfüllt sind (G-BA 2023). Die Aussage bzw. Forderung des IQWiG, die KNB als Teil der Preisverhandlungen zu definieren, sollte bzw. kann nicht Bestandteil eines Methodenpapiers sein, da die Entscheidung nicht durch das IQWiG getroffen werden kann. Die KNB muss differenziert betrachtet und diskutiert werden, da bspw. nicht alle im IQWiG Methodenpapier dargestellten Methoden dem gesundheitsökonomischen Goldstandard entsprechen – wie beispielsweise bereits durch die Deutsche Gesellschaft für Gesundheitsökonomie e. V. (dggö) in vergangenen Verfahren dargestellt wurde (dggö 2014). Mit der AMNOG Regulierung wurde bereits ein Modell etabliert, welches einen effizienten Ressourceneinsatz sicherstellt. <u>Vorgeschlagene Änderung:</u> Der Satz zum Ziel der KNB sollte wie folgt angepasst werden (Änderungen hervorgehoben): „Ziel der KNB ist es, nach abgeschlossener früher Nutzenbewertung bei Arzneimitteln mit neuen Wirkstoffen als ökonomische Orientierung (entsprechend § 35b SGB V) als Zusatz zur Nutzenbewertung zu dienen. “
Abschnitt 2.1.3.2 (S. 20 f.)	<u>Anmerkung:</u> Das IQWiG erläutert und stellt in Abbildung 4 dar, dass bei der Erarbeitung des Konzepts externe Sachverständige als auch Betroffene bzw. Patientenorganisationen und für die Informationsbeschaffung Betreiber bereits bestehender Datenplattformen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	eingebunden werden. Für die Einschätzung und Durchführung einer AbD ist es allerdings essentiell, dass auch der betroffene pharmazeutische Unternehmer bzw. die betroffenen pharmazeutischen Unternehmen eingebunden werden. <u>Vorgeschlagene Änderung:</u> Ergänzung von Abbildung 4 inklusive des Textes (Änderungen hervorgehoben): „Kontaktaufnahme mit Betreibern potenziell für die AbD geeigneter Datenplattformen (z. B. Indikationsregister, Produktregister) sowie pharmazeutischer Unternehmen.“
Abschnitt 4.1 (S. 83 f.)	<u>Anmerkung:</u> <i>„Ziel der KNB ist es, ökonomische Informationen als Zusatz zur Nutzenbewertung insbesondere für die Preisverhandlungen im Sinne einer Informationssynthese zusammenzufassen.“</i> Wie bereits im vorangegangenen Abschnitt erläutert, ist die KNB nicht regulärer Teil der Erstattungsbetragsverhandlungen und der entsprechenden Rahmenvereinbarung nach § 130b Abs. 9 SGB V. Diese kann – wie durch das IQWiG unter den Punkten 1 bis 3 dargestellt – basierend auf § 35b SGB V als Orientierung dienen, wenn die entsprechenden Voraussetzungen erfüllt sind (G-BA 2023). Die Aussage bzw. Forderung des IQWiG, die KNB als Teil der Preisverhandlungen zu definieren, sollte bzw. kann nicht Bestandteil eines Methodenpapiers sein, da die Entscheidung zur Durchführung einer KNB und deren Berücksichtigung in Erstattungsbetragsverhandlungen nicht durch das IQWiG getroffen werden kann. <u>Vorgeschlagene Änderung:</u> Löschung bzw. Änderung des Satzes zum Ziel der KNB zur Preisverhandlung, da dies missinterpretiert und als obligat gesehen werden kann. Die KNB ist nicht regulärer Bestandteil der Erstattungsbetragsverhandlungen.
Abschnitt 4.3.2 (S. 88)	<u>Anmerkung:</u> <i>„Dies kann insbesondere auch eine KNB von Orphan Drugs betreffen, für die im Rahmen des Auftrags für eine KNB ggf. ein Komparator durch den G-BA erst festzulegen ist.“</i> Die Besonderheiten von Orphan Drugs werden in der KNB nicht berücksichtigt. Eine Festlegung eines Komparators durch den G-BA für die KNB widerspricht der Intention des § 35a SGB V. Zudem zeigt sich im internationalen Vergleich, dass die KNB Besonderheiten von seltenen Erkrankungen nicht abdecken kann, wodurch die Versorgung von Patientinnen und Patienten gefährdet

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	wird. So musste bspw. in UK der Cancer Drug Fund (NHS 2016) eingerichtet werden, um die adäquate Versorgung sicher zu stellen. Die Besonderheiten von Orphan Drugs und seltenen Erkrankungen müssen auch in einer möglichen KNB berücksichtigt werden. Aufgrund der Seltenheit der Erkrankung gibt es oftmals keine alternative oder vergleichbare Therapie bzw. Therapiemöglichkeit für Patientinnen und Patienten. Einen Komparator hier rein zum Zweck der KNB zu definieren widerspricht der Bedeutung von Orphan Drugs und wird der besonderen Therapiesituation bei seltenen Erkrankungen nicht gerecht. Damit würde sich bei Orphan Drugs der Fokus weg von den Patientinnen und Patienten und deren prekärer Therapiesituation zu einer primären Kostenanalyse verschieben. Basierend auf dem SGB V und der Verfahrensordnung des G-BA, wird in der frühen Nutzenbewertung keine zweckmäßige Vergleichstherapie für Orphan Drugs festgelegt. Diese Regularien mittels der KNB zu umgehen, ist stark zu hinterfragen. Es würden damit völlig kontraproduktive Anreize in der Entwicklung von Orphan Drugs gesetzt. Es stellt sich zudem die Frage, ob die Festlegung eines Komparators durch den G-BA vom IQWiG gefordert werden kann. <u>Vorgeschlagene Änderung:</u> Löschung des Satzes zur Festlegung eines Komparators bei einer KNB von Orphan Drugs: „Dies kann insbesondere auch eine KNB von Orphan Drugs betreffen, für die im Rahmen des Auftrags für eine KNB ggf. ein Komparator durch den G-BA erst festzulegen ist“.
Abschnitt 4.3.3 (S. 88 f.)	<u>Anmerkung:</u> „Das QALY-Konzept stellt dennoch die derzeit am umfassendsten angewandte Methodik dar, die explizit und mit einem theoretisch begründeten Konstrukt die Bewertung von Therapien in gesundheitsökonomischen Evaluationen unter Berücksichtigung von Präferenzen ermöglicht.“ Ein QALY-basierter Bewertungsansatz wird im Rahmen der KNB aufgrund methodischer, sozialrechtlicher und ethischer Aspekte (Pettitt et al. 2016) als kritisch für Deutschland angesehen. Probleme, die mit dem QALY-Konzept einhergehen sind vielseitig – wie auch durch Rand und Kesselheim evaluiert und vom IQWiG kurz zusammengefasst wurde (Rand et al. 2021). Auch wenn das QALY-Konzept derzeit die „umfassendste angewandte Methodik“ darstellt, muss hinterfragt werden, ob dieses auch auf Deutschland angewendet werden kann bzw. sollte. Denn Basis für die KNB ist die individuelle bzw. gesellschaftliche Zahlungsbereitschaft (Schlander 2010).

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Auch berücksichtigt das QALY-Konzept nicht die Schwere oder Häufigkeit einer Erkrankung, wodurch nicht alle Patientinnen und Patienten Zugang zu innovativen Arzneimitteln erhalten – insbesondere bei seltenen Erkrankungen (Schlander et al. 2014) und / oder onkologischen Erkrankungen kann dies mit gravierenden Folgen für die Betroffenen einhergehen (Richardson et al. 2018, Schlander et al. 2021, Schlander et al. 2022). Des Weiteren werden Unterschiede im sozioökonomischen Status im Modell nicht berücksichtigt, die sich allerdings nachweislich auf die Überlebenszeit bei Krebserkrankungen in Europa auswirkt (Vaccarella et al. 2022). Darüber hinaus existieren methodische Probleme, wie die mangelnde Berücksichtigung des abnehmenden Grenznutzens sowie die Annahme konstanter Trade-Offs zwischen Lebensjahren und Lebensqualität. <u>Vorgeschlagene Änderungen:</u> Die Vor- bzw. Nachteile des QALY-Konzepts und auch eine mögliche Umsetzbarkeit in Deutschland (ohne vorliegenden Grenzwert, gesellschaftlichem Diskurs bzw. Zahlungsbereitschaft) sollte im Methodenpapier 7.0 detaillierter diskutiert und dargestellt werden.
Abschnitt 4.7.2 „Bewertung der Ressourcen“ (S. 96)	<u>Anmerkung:</u> <i>„Für die Bewertung der Kostenkomponenten [...] je nach Kostenbereich, mit regulierten Preisen (z. B. Diagnosis-related Groups für den stationären Sektor) oder verhandelten Preisen (Arzneimittelpreise).“</i> Bei der hier dargestellten Kostenbetrachtung bleibt unberücksichtigt, dass nicht alle Arzneimittel über einen Erstattungsbetrag verfügen. Arzneimittel, die vor Inkrafttreten des AMNOG zugelassen und in Deutschland im Markt verfügbar sind verfügen nicht über einen „verhandelten Preis“. Dies ist für die Kostenberechnung der ZVT, welche sowohl Erstattungsbeträge als auch „nicht-verhandelte“ Preise umfassen kann (inkl. Preise für Generika bzw. Biosimilars) entscheidend. Auch kommt es bei der Festlegung des Preises durch die Schiedsstelle – in welchem Fall eine KNB durchgeführt werden könnte – nicht zu einem „verhandelten Preis“. <u>Vorgeschlagene Änderung:</u> Der Satz sollte entsprechend angepasst werden (Änderungen hervorgehoben): „... oder beispielsweise verhandelten Preisen (Arzneimittelpreise).“

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 4.9 (S. 102)	<u>Anmerkung:</u> <i>„Extrapolation [...] Für diese Extrapolation sind mindestens die beiden folgenden Szenarien zu betrachten:</i> ▪ <i>[...]</i> ▪ <i>ein optimistischeres Szenario, [...] Effekt/Unterschied zwischen den betrachteten Therapien über das Ende der klinischen Studie hinaus anhält und danach abfällt oder endet.“</i> In diesem Szenario bleibt allerdings unberücksichtigt, dass der Effekt zwischen den betrachteten Therapien auch über das Ende der klinischen Studie hinaus verstärkt werden kann. <u>Vorgeschlagene Änderung:</u> Der Satz sollte entsprechend angepasst werden (Änderungen hervorgehoben): „...verstärkt, anhält und danach abfällt oder endet.“
Abschnitt 9.1.1 (S. 159)	<u>Anmerkung:</u> <i>„Falls geeignete Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium (Population bzw. Interventionen) vorliegen, muss über die Berücksichtigung der Studie anhand der Stärke der Effektmodifikation und dem Anteil der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen bzw. dem Grad der Abweichung der Interventionen entschieden werden.“</i> Es ist unklar wie „geeignete Informationen“ definiert wird, um darauf basierend über die Berücksichtigung von Studie(n) zu entscheiden. Es bleibt offen, wie die Ergebnissicherheit der Effektmodifikation, bspw. ausreichende Größe der jeweiligen Subgruppe, Anzahl der Subgruppen die eine Effektmodifikation bestätigen, die Studiengültigkeit und -qualität, gemessen und definiert wird. Die offene Definition bietet daher zu großen Spielraum für eine ergebnisgesteuerte Berücksichtigung von Studien. Auch bleibt unklar, ob die Effektmodifikation klinisch begründet sein muss. Häufig handelt es sich in der Nutzenbewertung um zufällige Effekte in einem einzigen Endpunkt ohne klinische Begründung (Fletcher et al. 2014, Leverkus et al. 2015, Rasch et al. 2015). Auch weichen Subgruppen per se häufig von der klinischen Praxis ab (AWMF 2016), weshalb eine mögliche Effektmodifikation basierend darauf keine Aussagen zulässt. Es ist daher fraglich Studien aufgrund einer theoretischen Effektmodifikation ein- bzw. auszuschließen.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Des Weiteren berücksichtigt der Entwurf des Methodenpapiers das IQWiG-Arbeitspapier mit Auftrag GA18-01 „Untersuchung der statistischen Eigenschaften von Verfahren zur Übertragbarkeit von Studienergebnissen auf Teilpopulationen“ (IQWiG 2018) nicht. Denn die Untersuchung von Teilpopulation reduziert die Power der Studie, um den Behandlungseffekt zu zeigen. Diese Problematik bleibt im Methodenpapier unberücksichtigt. <u>Vorgeschlagene Änderung:</u> Der Abschnitt zur Effektmodifikation bzw. Anteil der Patientinnen und Patienten zur Berücksichtigung der Studie(n) sollte entfernt werden, um eine ergebnisgesteuerte (Nicht-) Berücksichtigung zu vermeiden. Eine mögliche Vorgehensweise zum Power-Verlust bei der Untersuchung von Teilpopulationen sollte ergänzt werden.
Abschnitt 9.3.6 (S. 175 ff.)	<u>Anmerkung:</u> Das IQWiG erachtet es als „<i>erforderlich, vorab alle relevanten Confounder (inklusive wichtiger Wechselwirkungen) mithilfe eines systematischen Ansatzes zu identifizieren [587], vollständig zu erheben und bei der Datenanalyse in angemessener Form im Modell zu berücksichtigen [393]. [...] Im Rahmen einer AbD spielt zur Adjustierung für Confounder insbesondere die Propensity-Score-Methode eine Rolle</i>“. Die vorgeschlagene und einzig akzeptierte Confounder-Identifikation nach Pufulete et al. (Pufulete et al. 2022) kann zu Verzerrungen und möglichen Fehlinterpretationen von Confoundern führen. Gleichzeitig ist die Einschränkung auf eine einzige Methode, dem Propensity-Score-Ansatz, nicht begründet bzw. nachvollziehbar. Denn die Confounder-Adjustierung hängt von verschiedenen Faktoren ab, u. a. auch von der jeweiligen Versorgungssituation und Indikation. <u>Vorgeschlagene Änderung:</u> Eine Einschränkung der möglichen Confounder-Adjustierung auf die Propensity-Score-Methode ist nicht adäquat und steht im Gegensatz zum Rapid Report A19-43 (IQWiG 2020). Zudem sollten die Umsetzbarkeit der Prüfung von Confoundern berücksichtigt werden.
Abschnitt 9.3.12 (S. 191)	<u>Anmerkung:</u> „Für PROs ist in einer solchen Situation durch die systematisch verkürzte Datenerhebung die Interpretierbarkeit stark eingeschränkt. Nicht sinnvoll interpretierbar sind bei sich deutlich zwischen den Behandlungsarmen unterscheidenden

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>Beobachtungsdauern sogenannte „dauerhafte“ Veränderungen“ [...] nicht sachgerecht, von einer „dauerhaften“ Veränderung zu sprechen, da die Beobachtungszeit nur einen verkürzten und ggf. sehr kleinen Anteil der gesamten Studiendauer abdeckt. Solche Auswertungen werden im Falle einer verkürzten oder sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauer daher in der Regel nicht für die Nutzenbewertung herangezogen.“</i> In der Regel ist eine Erhebung von PROs über den gesamten Studienverlauf bzw. nach Progression der Erkrankung und somit oft über Jahre hinweg praktisch nicht durchführbar. Die Belastung für die Patientinnen und Patienten ist enorm, womit auch die EMA eine klare Empfehlung für eine „wohl dosierte“ Erhebung der PROs ausgesprochen hat. In den EMA Leitlinien wird empfohlen, dass eine Erhebung nur an Zeitpunkten durchgeführt werden soll, wenn eine klare und hypothesen-getriebene Rationale dafür besteht und mit hohen Rücklaufquoten zu rechnen ist (EMA 2016). Dies wurde u. a. auch durch Professor Wörmann (G-BA 2022) für den deutschen Versorgungskontext bestätigt: <i>„Den Punkt ‚lebenslange Beobachtung‘ verstehe ich aus methodischer Sicht, für Kliniker ist das ziemlich fern der Realität. Gerade beim Ösophaguskarzinom sind das schwerstkranken, auch sterbende Patienten. Die leben nicht für eine Studie, sie haben andere Prioritäten [...]“</i>. Durch diese hohe Belastung von Patientinnen und Patienten, insbesondere bei onkologischen Erkrankungen, würde es zudem zu stark reduzierten Rücklaufquoten kommen. Aus wissenschaftlicher Sicht (Basch et al. 2012) und bei den bereits hohen Ansprüchen an die erforderlichen Rücklaufquoten von mehr als 70 % ist fraglich, ob die PRO-Ergebnisse einer langjährigen Erhebung überhaupt berücksichtigt würden bzw. berücksichtigt werden könnten und somit für die Nutzenbewertung geeignet sind. Daher ist die Forderung nach einer langjährigen Erhebung nicht nachvollziehbar und es würde zum „research waste“ kommen (Rivera et al. 2019). Gleichzeitig würde bei insbesondere onkologischen Erkrankungen eine langwierige PRO-Erhebung dazu führen, dass Ergebnisse höchst verzerrt und nicht interpretierbar für den zu bewertenden Wirkstoff wären, sondern lediglich für die nachfolgenden Therapien. Des Weiteren ist es auch ethisch äußerst fragwürdig, insbesondere in der Onkologie, Patientinnen und Patienten mit der PRO-Erhebung über den Progress hinaus zu belasten. Alle Studiendesigns und Erhebungszeitpunkte werden sowohl mit den Zulassungsbehörden als auch zuständigen Ethikkommissionen abgestimmt. Eine „systematische Verkürzung“ der Beobachtungszeit wäre hierdurch nicht umsetzbar. Zudem

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	empfeht die EMA in ihren allgemeinen Prinzipien für klinische Studien und die Erhebung von PRO, dass, „[...] <i>good science applies and objectives need to be justified alongside realistic expectations. [...]</i> “ (EMA 2016). Dies wird durch die Forderung einer langjährigen Nachbeobachtung über die Behandlung des zu bewertenden Arzneimittels hinaus in Frage gestellt. <u>Vorgeschlagene Änderung:</u> Der Wortlaut „systematisch verkürzte Beobachtungsdauer“ sollte gelöscht werden. Ethisch und methodisch ist die Belastung der Patientinnen und Patienten durch die vom IQWiG geforderte langjährige PRO-Erhebung nicht begründbar.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. Abrahamyan, L., Feldmann, B. M., Tomlinson, G., Faughnan, M. E., Johnson, S. R., Diamond, I. R. and Gupta, S. 2016. Alternative designs for clinical trials in rare diseases. *American Journal of Med Gen Part C* 172C: 313-331.
2. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF) 2016. Positionspapier zur Definition und Bewertung von Subgruppen im Verfahren der frühen Nutzenbewertung von Arzneimitteln nach §35a SGB V und aufgrund des Arzneimittelmarktneuordnungsgesetzes (AMNOG) von 2010 der Ad hoc Kommission "Frühe Nutzenbewertung" der AWMF. Verfügbar unter: https://www.awmf.org/fileadmin/user_upload/dateien/nutzenbewertungen/Positionspapier_Fruehe_Nutzenbewertung_Subgruppen_2016-04.pdf, abgerufen am: 20.01.2023.
3. Basch, E., Abernethy, A. P., Mullins, D., Reeve, B. B., Smith, M. L., Coons, S. J., Sloan, J., Wenzel, K., Chauhan, C., Eppard, W., Frank, E. S., Lipscomb, J., Raymond, S. A., Spencer, M. and Tunis, S. 2012. Recommendations for Incorporating Patient-Reported Outcomes into clinical comparative effectiveness research in adult oncology. *J. of Clin. Oncol.* 30; 4249-4255.
4. Basch, E., Reeve, B. B., Mitchell, S. A., Clauser, S. B., Minasian, L. M., Dueck, A. C., Mendoza, T. R., Hay, J., Atkinson, T. M., Abernethy, A. P., Bruner, D. W., Cleeland, C. S., Sloan, J. A., Chilukuri, R., Baumgartner, P., Denicoff, A. M., Germain, D. S., O'Mara, A. M., Chen, A., Kelaghan, J., Bennett, A. V., Sit, L., Rogak, L. J., Barz, A., Paul, D. B. and Schrag, D. 2014. Development of the national cancer institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *J Natl Cancer Inst* 106(9).
5. Deutsche Gesellschaft für Gesundheitsökonomie e.V. (dggö) 2014. Stellungnahme der Deutschen Gesellschaft für Gesundheitsökonomie zu den inhaltlichen Aktualisierungen und Ergänzungen im Entwurf der Version 4.2 Allgemeine Methoden. Verfügbar unter: https://file.dggoe.de/2014-07-29_dgg%c3%b6_Stellungnahme_IQWiG_Allgemeine_Methoden_4_2.pdf, abgerufen am: 20.01.2023.
6. Dueck, A. C., Mendoza, T. R., Mitchell, S. A., Reeve, B. B., Castro, K. M., Rogak, L. J., Atkinson, T. M., Bennett, A. V. and Denicoff, A. M. 2015. Validity and Reliability of the U.S. National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTAE). *JAMA Oncol* 1(8): 1051-1059.
7. European Medicines Agency (EMA) 2016. Appendix 2 to the guidelines on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies. Verfügbar unter: <https://www.ema.europa.eu/en/appendix-2-guideline-evaluation-anticancer-medicinal-products-man-use-patient-reported-outcome-pro>, abgerufen am: 17.01.2023.
8. Fletcher, C., Chuang-Stein, C., Paget, M.-A., Reid, C. and Hawkins, N. 2014. Subgroup analyses in cost-effectiveness analyses to support health technology assessments. *Pharmaceut. Statist.* 13: 256-274.

9. Gemeinsamer Bundesausschuss (G-BA) 2022. Mündliche Anhörung gemäß § 35 a Abs. 3 Satz 2 SGB V Verfahrensordnung des Gemeinsamen Bundesausschusses hier: Nivolumab (D-823 + D-822). Verfügbar unter: https://www.g-ba.de/downloads/91-1031-831/2022-09-05_Wortprotokoll_Nivolumab_D-822.pdf, abgerufen am: 20.01.2023.
10. Gemeinsamer Bundesausschuss (G-BA) 2023. Kosten-Nutzen-Bewertung von Arzneimitteln nach § 35b SGB V. Verfügbar unter: <https://www.g-ba.de/themen/arzneimittel/arzneimittel-richtlinie-anlagen/kosten-nutzen/#:~:text=Kosten-%E2%80%8BNutzen-Bewertung%20von%20Arznei%C2%ADmitteln%20mit%20neuen%20Wirkstoffen.%20Die%20Kosten-%E2%80%8BNutzen-Bewertung,nur%20auf%20Antrag%20und%20nur%20unter%20folgenden%20>, abgerufen am: 17.01.2023.
11. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2018. IQWiG-Berichte - Nr. 638 Untersuchung der statistischen Eigenschaften von Verfahren zur Übertragbarkeit von Studienergebnissen auf Teilpopulationen. Arbeitspapier. Auftrag: GA18-01, Version: 1.0; Stand: 20.06.2018. Verfügbar unter: https://www.iqwig.de/download/ga18-01_uebertragbarkeit-von-studienergebnissen-auf-teilpopulationen_arbeitspapier_v1-0.pdf, abgerufen am: 30.01.2023.
12. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2020. IQWiG-Berichte-Nr. 863. Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V. Rapid Report. Auftrag: A19-43; Version: 1.1; Stand: 13.05.2020. Verfügbar unter: https://www.iqwig.de/download/a19-43_versorgungsnaher-daten-zum-zwecke-der-nutzenbewertung_rapid-report_v1-1.pdf, abgerufen am: 30.01.2023.
13. Leverkus, F. and Chuang-Stein, C. 2015. Implementation of AMNOG: an industry perspective. Biometrical Journal 58: 76-88.
14. National Health Service (NHS) England 2016. Appraisal and Funding of Cancer Drugs from July 2016 (including the new Cancer Drugs Fund). A new deal for patients, taxpayers and industry. Verfügbar unter: <https://www.england.nhs.uk/wp-content/uploads/2013/04/cdf-sop.pdf>, abgerufen am: 20.01.2023.
15. Pettitt, D. A., Raza, S., Naughton, B., Roscoe, A., Ramakrishnan, A., Ali, A., Davies, B., Dopson, S., Hollander, G., Smith, J. A. and Brindley, D. A. 2016. The limitations of QALY: A literature review. J Stem Cell Res & Ther 6(4).
16. Pufulete, M., Mahadevan, K., Johnson, T. W., Pithara, C., Redwood, S., Benedetto, U., Higgins, J. P. T. and Reeves, B. C. 2022. Confounders and co-interventions identified in non-randomized studies of interventions. Journal of Clinical Epidemiology 148: 115-123.
17. Rand, L. Z. and Kesselheim, A. S. 2021. Controversy over using Quality-Adjusted Life-Years in Cost-Effectiveness Analyses: A systematic Literature Review. Health Affairs 40: 9: 1402-1410.
18. Rasch, A. and Dintsios, C. M. 2015. Subgruppen in der frühen Nutzenbewertung von Arzneimitteln: eine methodische Bestandsaufnahme. Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ) 109: 69-78.
19. Richardson, J. and Schlander, M. 2018. Health technology assessment (HTA) and economic evaluation: efficiency or fairness first. Journal of Market Access & Health Policy 7:1.

20. Rivera, S. C., Kyte, D. G., Olalekan, L. A., L., S. A., McMullan, C. and Calvert, M. J. 2019. The impact of patient-reported outcome (PRO) data from clinical trials: a systematic review and critical analysis. *Health and Quality of Life Outcomes* 17:156.
21. Schlander, M. 2010. Measures of efficiency in healthcare: QALMs about QALYs? *Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ)* 104: 209-226.
22. Schlander, M., Garattini, S., Holm, S., Kolominsky-Rabas, P., Nord, E., Persson, U., Postma, M., Richardson, J., Simoons, S., Sola Morales, O., Tolley, K. and Toumi, M. 2014. Incremental cost per Quality-adjusted Life Year gained? The need for alternative methods to evaluate medical interventions for ultra-rare disorders. *J. Comp. Eff. Res.* 3(4): 399-422.
23. Schlander, M., Hernandez-Villafuerte, K., Cheng, C.-Y., Mestre-Ferrandiz, J. and Baumann, M. 2021. How much does it cost to research and develop a new drug? A systematic review and assessment. *PharmacoEconomics* 39: 1243-1269.
24. Schlander, M. and Richardson, J. 2022. QALYs in health resource usage decisions. *Health Affairs* 41(4): 609-610.
25. Sherian, J. and Julian, G. 2016. Achievements and limitations of Evidence-based Medicine. *Journal of American College of Cardiology* 68(2): 204-2013.
26. Vaccarella, S., Georges, D., Bray, F., Ginsburg, O., Charvat, H., Mertikainen, P., Bronnum-Hansen, H., Deboosere, P., Bopp, M., Leinsalu, M., Artnik, B., Lorenzoni, V., De Vries, E., Marmot, M., Vineis, P., Mackenbach, J. and Nusselder, W. 2022. Socioeconomic inequalities in cancer mortality between and within countries in Europe: A population-based study. *The Lancet Regional Health - Europe*.

A.1.3 AMS - Advanced Medical Services GmbH

Autorinnen und Autoren

- Hücherig, Stephanie
- Wüstner, Stefanie
- Hogger, Sara
- Brade, Thomas
- Richter, Swantje

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Dr. Stephanie Hücherig, Teamlead and Senior Medical Writer
Dr. Stefanie Wüstner, Head of Business and Scientific Innovation
Sara Hogger, Senior Biostatistician
Dr. Thomas Brade, Teamlead and Senior Medical Writer
Swantje Richter, Teamlead and Senior Biostatistician
Im Namen der AMS Advanced Medical Services GmbH
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen) ☒ im Namen folgender Institution / Organisation: AMS - Advanced Medical Services GmbH ☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>
Seit dem Start der frühen Nutzenbewertung gemäß Arzneimittelmarktneuordnungsgesetz (AMNOG) zählt AMS Advanced Medical Services GmbH zu den Marktführern bei der Unterstützung von pharmazeutischen Unternehmen bei G-BA-Beratungen, Nutzendossiers, Stellungnahmeverfahren und Preisverhandlungen. Aufgrund der umfassenden Expertise und vielfältigen praktischen Erfahrung möchten wir verschiedene Aspekte des Entwurfs der Allgemeinen Methoden Version 7.0 adressieren, die unserer Meinung nach einer Überarbeitung bedürfen. Insbesondere möchten wir zu folgenden Themen Stellung nehmen: - Recherche in bibliografischen Datenbanken, in Studienregistern und in Leitliniendatenbanken - Kriterien für den Einschluss von Studien in die Nutzenbewertung - Adjustierung und multifaktorielle Verfahren - Umgang mit fehlenden Daten - Umgang mit variablen Beobachtungszeiten, insbesondere bei patientenberichteten Endpunkten (Patient Reported Outcomes, PRO)
Allgemein möchten wir hervorheben, dass im Entwurf der Allgemeinen Methoden Version 7.0 einige Passagen klarer und verständlicher formuliert wurden, was die Interpretation und Umsetzung in der Praxis erleichtert. Insbesondere im neu strukturierten Kapitel 1.3.4 zur Bedeutung der Ergebnissicherheit werden die Begriffe der internen und externen Validität sowie der Präzision prägnant und übersichtlich eingeführt, das IQWiG konzentriert sich hier auf die wesentlichen Aspekte. Der neue Abschnitt 3.2.1 zur den Konzepten für die anwendungsbegleitende Datenerhebung (AbD) fasst den Rapid Report A19-43 kurz und prägnant zusammen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung
	<i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 8.1.1 Recherche in bibliografischen Datenbanken D) Entwicklung von Suchstrategien	<u>Anmerkung:</u> Dass IQWiG führt aus, dass, falls vorhanden, validierte Studienfilter (z. B. für randomisierte kontrollierte Studien [Randomized Controlled Trials, RCT] und systematische Übersichten) oder validierte Klassifikatoren aus dem maschinellen Lernen (z. B. RCT Classifier) Anwendung finden. Bei anderen Studientypen oder Fragestellungen sei im Einzelfall

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
(S. 148; PDF-S. 164)	zu prüfen, ob validierte Studienfilter vorliegen und diese zuverlässig anzuwenden sind. <u>Vorgeschlagene Änderung:</u> Eine detaillierte Beschreibung der Anforderungen an die erwähnte Prüfung und ggf. Validierung von Studienfiltern und Klassifikatoren wäre wünschenswert, um die vom IQWiG zu Grunde gelegten Anforderungen an diesen Prozess transparent und nachvollziehbar zu machen.
Abschnitt 8.1.2 Suche in Studienregistern (S. 150; PDF-S. 166)	<u>Anmerkung:</u> Bei der Beschreibung des Clinical Trials Information Systems (CTIS) ist die URL des EU-CTR angegeben statt die des CTIS. (CTIS: https://www.clinicaltrialsregister.eu/ctr-search/search) <u>Vorgeschlagene Änderung:</u> (CTIS: https://euclinicaltrials.eu/)
Abschnitt 8.4 Recherche nach Leitlinien (S. 156; PDF-S. 172)	<u>Anmerkung:</u> Die Angabe der berücksichtigten Leitliniendatenbanken unterscheidet sich zwischen der umfassenden Leitlinienrecherche und der fokussierten Leitlinienrecherche. Es ist unklar, ob die zusätzlichen genannten Datenbanken bei der fokussierten Recherche ebenfalls fester Bestandteil der umfassenden Leitlinienrecherche sind. Insbesondere wird bei der fokussierten Informationsbeschaffung nach zentralen Leitlinien eine Suche in MEDLINE durchgeführt. Bei der umfassenden Recherche wird MEDLINE hingegen nicht genannt. Es ist unklar, ob die umfassende Recherche ebenfalls in MEDLINE durchgeführt wird. Im Gegensatz zu einer fokussierten Recherche ist unter anderem die Vollständigkeit ein wichtiger Anspruch bei der umfassenden Recherche. Somit erscheint der Einbezug von MEDLINE für eine umfassende Recherche wesentlicher und könnte ggf. bei der fokussierten Recherche entfallen. <u>Vorgeschlagene Änderung:</u> Konkretisierung des Vorgehens bei der umfassenden Leitlinienrecherche und Abgrenzung gegenüber der fokussierten Recherche.
Abschnitt 9.1.1 Kriterien für den Einschluss von Studien	<u>Anmerkung:</u> Das IQWiG nimmt eine Umstrukturierung des Kapitels zu den Kriterien für den Einschluss von Studien vor:

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
(S. 159, PDF-S. 175)	In den Allgemeinen Methoden 6.1 wird die „80%-Regel“ bezüglich der Population sowie bezüglich der Interventions- und Vergleichsgruppe in unterschiedlichen Absätzen erläutert. Im Entwurf für die Allgemeinen Methoden 7.0 werden die Einschlusskriterien bezüglich der Population sowie bezüglich der Interventions- und Vergleichsgruppe mit einem „bzw.“ verbunden und parallel abgehandelt. Aus unserer Sicht führt dies zu Unklarheiten bezüglich der Auslegung der „80%-Regel“: Das „bzw.“ suggeriert, dass entweder das eine oder das andere Einschlusskriterium zu mind. 80% erfüllt sein muss. Es ist aber davon auszugehen, dass tatsächlich beide Kriterien erfüllt sein sollen. <u>Vorgeschlagene Änderung:</u> Ersetzung von „bzw.“ durch „und“ in den ersten beiden Sätzen des zweiten Absatzes.
Abschnitt 9.3.6 Adjustierung und multifaktorielle Verfahren (S. 176, PDF-S. 192)	<u>Anmerkung</u> Die Identifizierung aller relevanter Confounder ist für die adäquate Adjustierung von AbD von zentraler Bedeutung und sollte systematisch erfolgen, um die Vollständigkeit sicherzustellen. Im IQWiG-Report A19-43 wird dazu die wissenschaftliche Literatur in Kombination mit einer Beurteilung durch Fach-Experten genannt. <u>Vorgeschlagene Änderung:</u> Konkretisierung der Methodik, nach der die Confounder-Identifizierung durchgeführt wird und ggf. Angabe weiterer Quellen.
Abschnitt 9.3.6 Adjustierung und multifaktorielle Verfahren (S. 176, PDF-S. 192) Abschnitt 9.3.11 Umgang mit unvollständigen Daten (S. 189, PDF-S. 205) Abschnitt 9.3.13	<u>Anmerkung:</u> In Abschnitt 9.3.6 (Adjustierung und multifaktorielle Verfahren) werden die wichtigsten Aspekte, die im Rahmen einer statistischen Modellbildung beschrieben werden sollen, um die Punkte „Umgang mit fehlenden Daten“ und „Umgang mit unplausiblen Daten und Ausreißern“ erweitert. Diesem Thema widmet das IQWiG in seinen Allgemeinen Methoden das Kapitel 9.3.11 (Umgang mit unvollständigen Daten), das sich in erster Linie mit dem Problem der Follow-up-Verluste in klinischen Studien auseinandersetzt. Hier schreibt das Institut, dass Ergebnisse nicht berücksichtigt werden können, falls der Umfang fehlender Werte zu hoch ist oder ungeeignete Ersetzungsstrategien angewendet wurden. Des Weiteren wird das Thema Imputation im Kapitel 9.3.13 (Darstellung von Verzerrungsarten) aufgegriffen. Hier stellt das IQWiG dar, dass fehlende Werte, die nicht durch einen Zufallsmechanismus zustande kommen, eine Verzerrung der

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Darstellung von Verzerrungsarten (S. 191, PDF-S. 207)	Ergebnisse verursachen können. Gemäß IQWiG sind bei Bedarf statistische Verfahren, darunter Imputationsverfahren, nötig, die eine mögliche Verzerrung berücksichtigen. Gemäß dem Institut gilt kein Imputationsverfahren als allgemein akzeptiert. Aus dem Methodenpapier geht nicht hervor, anhand welcher Kriterien das Institut die Eignung einer Imputationsmethode beurteilt und, ausgehend davon, welche Imputationsmethoden in welchen Fällen als geeignet angesehen werden, um fehlende Werte adäquat zu ersetzen. Im Rahmen von klinischen Studien werden beispielsweise häufig multiple Imputationen (ggf. in Verbindung mit Markow-Chain-Monte-Carlo [MCMC]-Methoden oder dem Multivariate Imputation by Chained Equations [MICE]-Algorithmus) durchgeführt; auch bayessche Methoden finden Anwendung [1, 2, 3]. Dem Umgang mit fehlenden oder unplausiblen Daten und Ausreißern kommt insbesondere im Rahmen der AbD besondere Bedeutung zu. Bei Indikationsregistern, welche die Datengrundlage für die AbD bilden, sind üblicherweise fehlende Werte in deutlich größerem Ausmaß vorhanden als bei klinischen, interventionellen, Studien, die in einem äußerst kontrollierten Rahmen durchgeführt werden [4]. Die oben genannten Kapitel geben nur bedingt Auskunft darüber, wie das Institut den Aspekt der fehlenden Werte und Ersetzungsmethoden im Rahmen der AbD bewertet. <u>Vorgeschlagene Änderung:</u> - Listung und Erläuterung der Kriterien, anhand derer die Eignung einer Imputationsmethode beurteilt wird. - Erläuterung, welche Algorithmen zur Imputation als adäquat beurteilt werden. - Quantifizierung des Schwellenwertes, ab dem der Umfang fehlender Werte „zu hoch“ ist. - Bei den oben genannten Aspekten ist die unterschiedliche Natur von beobachtenden Daten/Registerdaten und klinischen, interventionellen Daten zu beachten.
Abschnitt 9.3.6 Adjustierung und multifaktorielle Verfahren (S. 177, PDF-S. 193)	<u>Anmerkung:</u> Im Rahmen der AbD stellt das IQWiG insbesondere die Propensity-Score-Methode zur Adjustierung von Confoundern heraus. Hinsichtlich der Ansätze zur Überprüfung der zentralen Annahmen – Positivität, Überlappung und Balanciertheit – verweist das Institut auf eine Publikation von Austin (2011) [5]. Austin beschreibt in dieser Publikation konkret die Überprüfung der Balanciertheit. Hinsichtlich der Positivität und der Überlappung gibt es kein empfohlenes Vorgehen.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Insgesamt geht aus dem Methodenpapier nicht klar hervor, welche Kriterien und Schwellenwerte das Institut anlegt, um die oben genannten Annahmen zu überprüfen. <u>Vorgeschlagene Änderung:</u> Darlegung der Kriterien, um die Annahmen der Positivität, Überlappung und Balanciertheit zu überprüfen und ggf. Angabe weiterer Quellen.
Abschnitt 9.3.12 Umgang mit variablen Beobachtungszeiten (S. 191, PDF- S. 207)	<u>Anmerkung:</u> Das IQWiG gibt an, dass Überlebenszeitanalysen von PRO bei verkürzten und/oder unterschiedlichen Beobachtungszeiten durch Abbruch der Beobachtungen, beispielsweise durch Progression, nur eingeschränkt interpretiert werden können. Weiter führt das Institut aus, dass eine dauerhafte Veränderung in dem länger beobachteten Behandlungsarm aufgrund der dafür notwendigen ununterbrochenen Über- bzw. Unterschreitung des Schwellenwerts potenziell schwerer zu erreichen sei. Insbesondere dauerhafte Veränderungen würden in solchen Fällen nicht für die Nutzenbewertung herangezogen. Diese Überlegungen beziehen allerdings verschiedene relevante Aspekte nicht ein: - Wird eine dauerhafte Verbesserung oder eine dauerhafte Verschlechterung betrachtet? - Handelt sich bei dem länger beobachteten Behandlungsarm um den Arm des zu bewertenden Arzneimittels (zbAM) oder um den Arm der zweckmäßigen Vergleichstherapie (zVT)? - Werden fehlende Werte, die vereinzelt innerhalb der Beobachtungszeit auftraten, unter validen Annahmen adäquat ersetzt? - Werden Follow-up-Verluste unter validen Annahmen adäquat ersetzt? - Liegen weitere Analysen bzw. Sensitivitätsanalysen vor, die die Einordnung des Effekts möglich machen (z. B. Analysen zur erstmaligen Veränderung oder zur bestätigten Veränderung, Sensitivitätsanalysen mit gesonderter Berücksichtigung von Patienten, die zum letzten Erhebungszeitpunkt einmalig den Schwellenwert erreichten)? Unter Berücksichtigung der oben genannten Aspekte kann der Effekt einer dauerhaften Veränderung unter Umständen trotzdem zur Nutzenbewertung herangezogen werden. Liegt beispielsweise ein statistisch signifikanter Effekt zugunsten des zbAM vor, der potenziell zuungunsten des zbAM verzerrt ist, kann dies als konservative Schätzung betrachtet werden. <u>Vorgeschlagene Änderung:</u>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Berücksichtigung der oben genannten Aspekte beim Vorgehen der Bewertung von PRO, insbesondere bei dauerhaften Veränderungen.
Literaturverzeichnis Nr. 335 und 609 (S. 234, PDF-S. 250 und S. 253, PDF- S. 269)	<u>Anmerkung:</u> Die angegebenen Links führen nicht auf den richtigen DOI-Eintrag.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*. 2012;367(14):1355-60.
2. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj*. 2009;338.
3. SAS S. 14.1 User's Guide The MI Procedure. SAS Institute. 2015.
4. Mack C, Su Z, Westreich D. Managing missing data in patient registries: addendum to registries for evaluating patient outcomes: a user's guide. 2018.
5. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*. 2011;46(3):399-424.

A.1.4 Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF)

Autorinnen und Autoren

- Nothacker, Monika
- Kopp, Ina
- Schliephake, Henning
- Treede, Rolf-Detlef

Stellungnahme

Stellungnahme der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fach- gesellschaften (AWMF) zum Entwurf des IQWiG „Allgemeine Methoden - Entwurf für Version 7.0“ vom 28.02.2023

Berlin, 28.02.2023: Die AWMF wurde am 26.01.2023 um eine Stellungnahme zu dem oben genannten Entwurf gebeten. Die AWMF hat ihrerseits ihre Mitgliedsfachgesellschaften gebeten, bei Bedarf eine eigene Stellungnahme zu verfassen. Die bis einschließlich 26.02.2023 bei der AWMF eingegangenen 3 Stellungnahmen von insgesamt 6 Fachgesellschaften sind der Stellungnahme der AWMF als Anlage beigefügt (s. Anlage 1). Diese äußern sich auch zu weiteren Aspekten, die wir neben der vorliegenden Stellungnahme der AWMF ebenfalls zu berücksichtigen bitten.

Die AWMF begrüßt -wie auch schon in früheren Stellungnahmen betont- die regelmäßige Überarbeitung der „Allgemeinen Methoden“ des IQWiG, da hierdurch aktuellen methodischen Entwicklungen auf nationaler und internationaler Ebene Rechnung getragen werden kann.

Die für den aktuellen, umfassend überarbeiteten Entwurf des Methodenpapiers gesetzte Frist zur Abgabe einer Stellungnahme ist allerdings mit 4 Wochen für ein so umfassendes (266 Seiten) und bedeutendes Werk zu kurz. Dies spiegelt sich wider in einer vergleichsweise geringen Beteiligung der Wissenschaftlichen Medizinischen Fachgesellschaften an der Möglichkeit zur Kommentierung. Wir bitten, künftig eine Frist zur Stellungnahme - wie bereits früher angemahnt - von mindestens 6 Wochen einzuräumen.

Zudem regen wir an, in künftigen Methodenberichten das Ziel der internationalen Vereinheitlichung der Methoden und dabei insbesondere den Grad der Ausrichtung des IQWiG an der Methodik der GRADE (Grading of Recommendations Assessment, Development and Evaluation)-Arbeitsgruppe sowie die Aktivitäten des IQWiG in Bezug auf das Schaffen eines digitalen Evidenz-Ökosystems¹ zu adressieren.

¹ Siehe Graphik: MAGIC (Making GRADE the irresistible Choice) Evidence Eco Foundation [Webseite] Abrufbar unter:
<https://magicvidence.org/evidence-ecosystem> [letzter Aufruf 27.02.2023]

Stellungnahme zu spezifischen Aspekten

Kapitel 1.3. Evidenzbasierte Medizin

Die AWMF begrüßt die neue Strukturierung dieses Kapitels mit Einordnung der Arbeit des Instituts in die evidenzbasierte Gesundheitsversorgung und klaren Erläuterungen. In Bezug auf die grundlegenden Strategien der EbM (Abschnitt 1.3.3, S.7) wird das Bekenntnis zur Identifikation und Zusammenfassung „aller bezüglich der Qualität ihres Designs und ihrer Durchführung angemessenen Studien“ deutlicher hervorgehoben. Dabei ist es wichtig, Angemessenheit im Sinne der „bestverfügbaren“ Evidenz zu verstehen. Es besteht eine zunehmende Notwendigkeit, die Methodendiskussion zu erweitern- nicht nur um diagnostische und qualitative Fragestellungen, sondern auch in Bezug auf Registerstudien als eine besondere Form von Beobachtungsstudien oder sogar randomisierte Studien inklusive der kritischen Bewertung von digitalen Studiendesigns². Der Umgang mit digitalen -auch „in vitro trials“- bezeichneten Studien auf Grundlage elektronischer Patientendaten sollte in einer nächsten Version des Methodenreports adressiert werden.

Kapitel 1.4, Gesundheitsökonomie

Im Methodenreport ist das Ziel der Kosten-Nutzenbewertung festgehalten: *„ökonomische Informationen als Zusatz zur Nutzenbewertung insbesondere für die Preisverhandlungen im Sinne einer Informationssynthese zusammenzufassen.“* Die explizite Empfehlung eines angemessenen Preises steht jedoch nicht im Fokus. Hier wäre eine größere Transparenz des Dialogs zwischen IQWiG als Informationsgeber und G-BA als Entscheidungsträger wünschenswert, um Erstattungsentscheidungen nachvollziehbarer zu machen (siehe auch Stellungnahmen der DGIM, GMDS und IBS-DER).

Kapitel 2, Produkte des Instituts

Die AWMF begrüßt die Möglichkeit der Stellungnahme und die durch das Institut etablierte Würdigung derselben. Zu kritisieren sind Ausnahmeregelungen (Rapid Reports, Potenzialbewertungen) welche aufgrund von Zeitdruck auf das Institut bestehen. Insgesamt ist aus Sicht der AWMF für alle Produkte des IQWiG die strukturierte Einbeziehung des Sachverstands von Vertretenden der wissenschaftlichen medizinischen Fachgesellschaften sowie von Vertretenden der Patientenorganisationen unverzichtbar. Wir verweisen dazu auf die grundsätzliche Stellungnahme der AWMF³ und die aktuelle Stellungnahme der DGIM.

2 Inan OT, Tenaerts P, Prindville SA et al. Digitizing Clinical Trials. npj Digital Medicine (2020) 3:101 ; <https://doi.org/10.1038/s41746-020-0302-y>. Verfügbar: <https://www.nature.com/articles/s41746-020-0302-y>

020-0302-y. Verfügbar: <https://www.nature.com/articles/s41746-020-0302-y>

3 Stellungnahme der AWMF zum Entwurf des Methodenpapiers Version 4 des IQWiG vom 08.03.2011 zum Aspekt: Produktspezifische

Verfahrensabläufe. Verfügbar: https://www.awmf.org/fileadmin/user_upload/dateien/stellungnahmen/2011/2.pdf

Vor dem Hintergrund der europaweiten Vereinheitlichung von Health Technology Assessments (HTA)⁴ wäre zu prüfen, wie sich die Produkte des IQWiG hier einordnen lassen. Der AWMF ist bewusst, dass durch die Europäisierung Stellungnahmemöglichkeiten zu Produkten des IQWiG im Bereich HTA weiter eingeschränkt werden könnten. Die Umbenennung des Produkts „HTA-Bericht“ in ThemenCheck-Bericht als nationales Produkt wird daher sehr begrüßt. Die AWMF wünscht sich darüber hinaus, in engem Dialog mit dem IQWiG den Einbezug der Expertise ihrer Mitgliedsfachgesellschaften im europäischen Kontext zu etablieren.

Kapitel 2.1.11, Evidenzberichte für Leitlinien

Das IQWiG hat die Prozesse für das neue Produkt „Evidenzberichte“ in seiner Methodenbeschreibung konkretisiert und ist dabei der Anregung der AWMF gefolgt, den notwendigen, klinisch-methodischen Austausch zwischen Leitlinienkoordinierenden und Mitarbeitenden des IQWiG von Beginn an festzuschreiben (Seite 34, Abbildung 12). Allerdings ist diese Kooperation für jeden Arbeitsschritt festzuschreiben. Wir weisen nochmals darauf hin, dass die Rücksprache mit den Leitlinienkoordinierenden vor der abschließenden Erstellung eines Evidenzberichts aus unserer Sicht regelhaft und nicht optional zu erfolgen hat und für die Evaluation des jeweiligen Prozesses ein Abschlussgespräch in die Abbildung aufgenommen werden sollte⁵. Wir möchten seitens der AWMF nochmals betonen, dass das Digitale-Versorgung-Gesetz von Themen für Leitlinien spricht, die seitens der AWMF dem BMG vorgeschlagen werden können und nicht von explizit ausgearbeiteten Fragestellungen. Der Prozess der Konkretisierung der Fragestellung gehört demzufolge – wie auch in der Abbildung korrekt dargestellt – bereits zu den Eingangsaufgaben des IQWiG für jeden Evidenzbericht. Für die kollegiale Zusammenarbeit mit den Mitarbeitenden der zuständigen Abteilung Versorgung, durch die diese Konkretisierungen zusammen mit den Leitlinienkoordinierenden bislang verwirklicht werden konnten, möchten wir uns an dieser Stelle bedanken.

Kapitel 2.2.3, Gewährleistung der fachlichen Unabhängigkeit

Mit völligem Unverständnis nimmt die AWMF wahr, dass das IQWiG in seinem aktuellen Methodenreport die international in den letzten Jahren vereinheitlichten Terminologien und Prozesse zur Darlegung von Interessen und zum Umgang mit Interessenkonflikten offenbar ignoriert und stattdessen eine neue Terminologie („Darlegung von Beziehungen“) einführen will.

4 Regulation on Technology Assessment. Verfügbar: https://health.ec.europa.eu/health-technology-assessment/regulation-health-technology-assessment_en

5 Stellungnahme der AWMF zum Entwurf des Methodenpapiers Version 6.1 des IQWiG vom 23.08.2021. Verfügbar: https://www.awmf.org/fileadmin/user_upload/dateien/stellungnahmen/2021/20211004__AWMF_Stellungnahme_Allg._Methoden_6.1f.pdf

Wir verweisen dazu auf internationale Standards und deren Umsetzung Deutschland⁶. Die AWMF hat dazu eine Plattform zur digitalen Verwaltung von Interessenerklärungen etabliert⁷. Zusätzlicher bürokratischer Aufwand ist nicht vermittelbar. Gern stehen wir seitens der AWMF dazu beratend zur Verfügung.

Kapitel 3, Nutzenbewertung medizinischer Interventionen

In diesem Abschnitt vermisst die AWMF ein explizites Bekenntnis zu international anerkannten Methoden. Wir halten an unserem Änderungsvorschlag für das Methodenpapier 6.0 fest und schlagen folgende Ergänzung nochmals vor:

„Das IQWiG verwendet zur Nutzenbewertung und der Einschätzung der Stärke der Ergebnis(un)sicherheit den internationale Standard in der evidenzbasierten Medizin, wie er von der GRADE Gruppe erarbeitet wurde. Als erforderlich gesehene Abweichungen oder zusätzliche Regelungen werden begründet“.

Dieses Vorgehen würde auch der sich immer internationaler gestaltenden Evidenzbewertung Rechnung tragen. Eine Aktualisierung der Zitate für und ein klares Bekenntnis zur Anwendung der GRADE Methodik in Bezug auf verschiedene Fragestellungen (Diagnostik, Therapie, Prognose) halten wir für unabdingbar^{2,4,8}. Zudem sollte die Berücksichtigung indirekter Evidenz unter bestimmten Voraussetzungen adressiert werden. Dies gilt ebenso für Patient*innen mit seltenen Erkrankungen.

In Bezug auf Abschnitt 3.1.1 „Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens“ verweisen wir auf die Stellungnahme der GMDS und IBS-Dr und möchten die künftige Nutzung von Core-Outcome-Sets unterstreichen.

Es ist künftig damit zu rechnen, dass Nutzenbewertungen bzw. HTA Berichte primär auf europäischer Ebene erstellt werden. Dies ist sinnvoll im Sinne der Vermeidung von Doppelarbeit. Dazu sollten explizite Ausführungen zu einer ggf. erforderlichen Methodenharmonisierung oder bereits bestehenden Methodischen Übereinstimmung der europäischen HTA Institute z.B. aufgrund gemeinsamer Erarbeitungen im europäischen HTA Netzwerk EUnetHTA ausgeführt werden.

Während die Übertragbarkeit von Leitlinienempfehlungen im Abschnitt 5.2 kurz thematisiert wird, fehlen Ausführungen zur Übertragbarkeit für die genannten Produkte. Diese sollten ergänzt werden.

6 Guidelines International Network: Principles for Disclosure of Interests and Management of Conflicts in Guidelines, verfügbar:

<https://www.acpjournals.org/doi/10.7326/m14-1885>

7 Siehe <https://interessenerklaerung-online.awmf.org/>

8 Stellungnahme der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)

zum Entwurf des IQWiG „Allgemeine Methoden -Entwurf für Version 6.0“ vom 05.12.2019. Verfügbar:

https://www.awmf.org/fileadmin/user_upload/dateien/stellungnahmen/2020/20200131__AWMF_Stellungnahme_Allg._Methoden_6.0.pdf

Kapitel 5.2.2 Leitliniensynopsen

Die Definition evidenzbasierter Leitlinien ist dünn, wir empfehlen einen Nachweis der Evidenzbasierung in Form von Evidenztabellen und/oder konzisen kritischen Evidenz-Beschreibungen im Hintergrundtext der betreffenden Leitlinie.

Kapitel 7 Evidenzbasierte Gesundheitsinformationen

Es wird begrüßt, dass AMSTAR II Instrument nun auch offiziell Einzug in die Methoden des IQWiG gefunden hat.

Kapitel 9.1.3 Rangordnung verschiedener Studienarten / Evidenzgrade

Hier sind Entwicklungen der GRADE-Bewertung für nicht-randomisierte Studien und deren gleiche Ausgangssetzung in Bezug auf Aussagesicherheit pro Endpunkt, wenn mit dem Bewertungsinstrument Robins-I bewertet wird, nicht berücksichtigt⁹. Das IQWiG sollte in seinem Methodenpapier eine Einschätzung dazu abgeben.

Für Rückfragen stehen wir jederzeit gerne zur Verfügung.

Dr. med. Monika Nothacker, MPH [REDACTED]

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

Prof. Dr. med. Dr. med. dent. Henning Schliephake [REDACTED]

Prof. Dr. med. Rolf-Detlef Treede [REDACTED]

⁹ Schünemann HJ, Cuello C, Akl EA, Mustafa RA, et al, GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. J Clin Epidemiol. 2019 Jul;111:105-114. doi: 10.1016/j.jclinepi.2018.01.012. Epub 2018 Feb 9. PMID: 29432858; PMCID: PMC6692166. <https://pubmed.ncbi.nlm.nih.gov/29432858/>

Anlage 1:

Stellungnahme der Deutschen Gesellschaft für Innere Medizin (DGIM)

Stellungnahme der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) und der Deutschen Region der Internationalen Biometrischen Gesellschaft (IBS-DR)

Gemeinsame Stellungnahme der Deutschen Gesellschaft für Orthopädie und Unfallchirurgie (DGOU), der Deutschen Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC) und der Deutschen Gesellschaft für Unfallchirurgie (DGU)

A.1.5 Arbeitsgruppe Epidemiologische Methoden der DGEpi, GMDS & DGSMP für die Deutsche Gesellschaft für Epidemiologie (DGEpi e. V.)

Autorinnen und Autoren

- Hardt, Juliane
- Rübsamen, Nicole

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Juliane Hardt
Dr. Nicole Rübsamen
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Arbeitsgruppe Epidemiologische Methoden der DGEpi, GMDS & DGSMP für die Deutsche Gesellschaft für Epidemiologie (DGEpi e.V.)
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Die AG Epidemiologische Methoden der DGEpi, GMDS und DGSMP begrüßt folgende Ergänzungen im Entwurf für Version 7.0 der Allgemeinen Methoden:

- Erweiterung der IQWiG-Produkte um die anwendungsbegleitenden Datenerhebungen und die Prüfung von Registerdaten
- das Eingehen auf kausale Modelle und Propensity-Score-Analysen zur Adjustierung in den Analysemethoden
- die klarere Beschreibung und Darstellung von Designs diagnostisch-therapeutischer Studien

und die insgesamt verbesserte Lesbarkeit des Methodenpapiers.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
z. B. Abschnitt 3.4 (S.16)	<u>Anmerkung:</u> <u>Vorgeschlagene Änderung:</u>
	<u>Anmerkung:</u> <u>Vorgeschlagene Änderung:</u>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

A.1.6 Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ)

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ) Wissenschaftlicher Fachausschuss der Bundesärztekammer [REDACTED], [REDACTED] www.akdae.de
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: AkdÄ, Wissenschaftlicher Fachausschuss der Bundesärztekammer
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>
Aus Sicht der Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ) gibt es am Entwurf zu den Allgemeinen Methoden 7.0 des IQWiG nur wenige Änderungen, die grundlegend zu kritisieren wären. Die dargestellten spezifischen Anmerkungen betreffen eher Präzisierungen oder Ergänzungen, die hilfreich erscheinen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 2.1.3.2 S. 20 „Bei der Erarbeitung des AbD-Konzepts ...“	<u>Anmerkung:</u> Die Informationsbeschaffung über bestehende Datenplattformen und Indikationsregister erscheint sehr allgemein dargestellt. Es wäre z. B. die Frage, ob nur nationale oder auch internationale betrachtet werden, und wenn auch letztere, ob z. B. nur europäische, westeuropäische etc. In dem Rahmen wäre es nötig, die Versorgungsbedingungen näher zu definieren, die für eine Verwendung der Datenplattformen und Indikationsregister zu fordern sind. Genaugenommen wäre diese Präzisierung auch für die nationalen Datenplattformen und Indikationsregister nötig. Auch wären die Qualitätsanforderungen an ein Register an dieser Stelle schon sinnvoll. <u>Vorgeschlagene Änderung:</u> Präzisierung der Versorgungsbedingungen, die für die Verwendung der Quellen erfüllt sein müssen. Wesentliche Anforderungen an die wissenschaftliche Qualität eines Registers benennen.
Abschnitt 2.1.3.3 S. 20/21 „Recherche nach Indikationsregistern ...“	<u>Anmerkung:</u> Die Recherche nach Indikationsregistern wird nicht weiter präzisiert und bleibt so etwas im Unklaren. <u>Vorgeschlagene Änderung:</u> Darstellung der Systematik, nach der die Recherche erfolgen soll, z. B. Abfrage entsprechender nationaler Fachgesellschaften, auch internationaler (s. o.) und ggfs. welcher. Gibt es Register für diese Datenbanken und/oder Indikationsregister?

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 3.2.1, S. 55 „Umgang mit Reporting Bias ...“	<u>Anmerkung:</u> Abschnitt 3.2.1 ist nicht nur umbenannt, sondern erheblich verkürzt und nur noch sehr allgemein gehalten. Der Grund hierfür bleibt unklar. Weder inhaltlich noch vom konkreten Vorgehen sollte sich nichts geändert haben – wenn doch, wird es nicht beschrieben. <u>Vorgeschlagene Änderung:</u> Inhaltlich und textlich Beibehaltung des Abschnitts 3.2.1 wie in den Methoden Version 6.1 (Umbenennung wäre kein Problem) – oder Begründung für die Kürzung und ggfs. Darstellung zukünftiger Vorgehensweisen.
Abschnitt 3.3.4, S. 67, 2. Punkt: „Prüfung, ob ... in bereits geplanten oder laufenden Datenerhebungen zum jeweiligen Arzneimittel ...“	<u>Anmerkung:</u> Es ist nicht klar, was hier mit „in geplanten oder laufenden Datenerhebungen“ gemeint ist. Nach dem Gesamtkontext zu urteilen, müssten Studien gemeint sein. Wenn ja, sollte klarer beschrieben werden, welche Form von Studien angesprochen sind, und ob hier z. B. auch Registerstudien gemeint sind. Zudem erscheint eine Feststellung, dass die Fragestellungen in geplanten oder laufenden Datenerhebungen schon adressiert sind, unzureichend, da die Datenerhebungen auch geeignet sein müssen. <u>Vorgeschlagene Änderung:</u> Lediglich Präzisierung dieses Punktes.
Abschnitt 3.3.4, S. 67, 3. Punkt: „... neben der studienindividuellen Datenerhebung ...“	<u>Anmerkung:</u> Es erscheint unklar, was mit „studienindividueller Datenerhebung“ gemeint ist. Studien, die nicht Registerstudien sind (vom Kontext zu vermuten)? Welche (nicht randomisierten) Studienformen sind hier gemeint? <u>Vorgeschlagene Änderung:</u> Studientypen klarer benennen.
Abschnitt 3.3.4, S. 67, 3. Punkt: „... potenziell geeignete Register identifiziert“ ...“	<u>Anmerkung:</u> Wie werden „potenziell geeignete Register identifiziert“? Siehe auch oben. Hier sollten wissenschaftliche Kriterien genannt werden, die Registerdaten als geeignet beschreiben (z. B. vordefinierte feste Zeitpunkte der Datenerhebung, Qualität und Vollständigkeit der Daten, Daten, die patientenrelevante Endpunkte erfassen, Fragebögen zur Lebensqualität etc.). <u>Vorgeschlagene Änderung:</u> Präzisierung, ob neben nationalen auch internationale Register etc. infrage kommen.
Abschnitt 3.3.4, S. 68, 2. Punkt: „... innerhalb eines akzeptablen	<u>Anmerkung:</u> Wesentlich erscheint, dass der Zeitraum nicht nur akzeptabel sein sollte, sondern dass die Daten für das Arzneimittel und die zweckmäßige Vergleichstherapie (ZVT) auch zu ausreichend

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Zeitraums ... rekrutiert und beobachtet ...“	gleichen Zeiten erhoben werden, um quasi „historische Vergleiche“ zu verhindern. Vor allem muss die Methodik der Datenerhebung vorgegeben werden und vergleichbar sein (z. B. gleiche Standardtherapie oder „best supportive care“).
Abschnitt 7.4, S. 133/134: Zu „Auswahl der Evidenz“	<u>Anmerkung:</u> Die Verwendung von AMSTAR2 anstatt des Oxman-Guyatt-Index für die methodische Qualität systematischer Übersichten ist aus Sicht der AkdÄ nachvollziehbar. <u>Vorgeschlagene Änderung:</u> Es sollte ein Mindestwert für den AMSTAR2 definiert werden und es sollte eine maximale Anzahl für die qualitativ besten Übersichten definiert werden, die berücksichtigt werden – z. B. die drei besten.
Abschnitt 8.2.1, S. 155 „...Dann wird eine oder werden ggf. mehrere...werden. Bei diesem...von der verwendeten Basis- SÜ...“	<u>Anmerkung:</u> Beschränkung auf maximal drei Basis-systematische Übersichten (Basis-SÜ) erscheint sinnvoll. Im Folgesatz müsste es „von den verwendeten Basis-SÜ“ heißen. <u>Vorgeschlagene Änderung:</u> Wie angemerkt.
Abschnitt 8.2.2, S. 155	<u>Anmerkung:</u> Im neuen Abschnitt 8.2.2 werden die früheren Abschnitte 8.2.2 bis 8.2.5 stark verkürzt zusammengefasst. <u>Vorgeschlagene Änderung:</u> Keine. Zusammenfassung und Verkürzung erscheint nachvollziehbar, da sich für fokussierte Informationsbeschaffungen die Fragestellungen situativ ergeben und Variabilität gefragt ist.
Abschnitt 9.1.1, S. 159 „Einschluss von Studien“	<u>Anmerkung:</u> Dieser Abschnitt erscheint lediglich präzisiert ohne wesentliche inhaltliche Änderungen. <u>Vorgeschlagene Änderung:</u> Keine.
Abschnitt 9.3.2, S. 170 „Signifikanztests und Konfidenzintervalle“	<u>Anmerkung:</u> Der Abschnitt ist jetzt zwar knapper formuliert als in der Version 6.1, erscheint aber doch ausreichend präzise. Der Satz „Im Rahmen von Bewertungen auf Basis bereits vorhandener Daten können die Ergebnisse statistischer Signifikanztests allerdings keine streng konfirmatorische Hypothesenprüfungen darstellen.“ erscheint in dieser Form neu, kann aber nur bekräftigt werden, zumal der Punkt bei der frühen Nutzenbewertung ein gängiges Problem darstellt.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Vorgeschlagene Änderung:</u> Keine.
Abschnitt 9.3.6, S. 176/177 „Adjustierung und multifaktorielle Verfahren“	<u>Anmerkung:</u> In den Passagen zur adäquaten Adjustierung für „Confounder“ bei anwendungsbegleitenden Datenerhebungen (AbD) erscheint (wie schon oben angemerkt) der Aspekt inadäquat berücksichtigt, dass die Datenerhebung zum Arzneimittel und zur ZVT evtl. nicht zeitgleich erfolgt. Die bevorzugte Verwendung des „Propensity-Scores“ zur Adjustierung für „Confounder“ ist zu begrüßen. Es erscheint aber unsicher, ob damit auch alle „Confounder“ erfasst werden, die durch zeitlich ggfs. unterschiedliche Erfassung der Daten entstehen könnten (z. B. Basisversorgung im Zeitverlauf). <u>Vorgeschlagene Änderung:</u> Möglicherweise sollte auch eine ausreichende Überlappung hinsichtlich der Zeit der Datenerhebung für das Arzneimittel und die ZVT gefordert werden.
Abschnitt 9.3.12 S. 190/191 „Umgang mit unvollständigen Daten“	<u>Anmerkung:</u> Abschnitt 9.3.12 entspricht dem früheren Abschnitt 9.3.11. Ergänzt ist in dem Abschnitt vor allem eine ausführliche Darstellung der Probleme, die sich durch eine Verkürzung der Erhebung von unerwünschten Ereignissen (UE) und Patient-reported Outcomes (PRO) ergeben, z. B. Begrenzung der Nachbeobachtung nach Behandlungsabbruch oder -wechsel auf 30 Tage (häufig oder fast die Regel in onkologischen Studien). Die Forderung vergleichbarer Beobachtungszeiten für alle (erfassbaren) Endpunkte ist berechtigt und nachdrücklich zu begrüßen (und wurde seitens der AkdÄ auch in etlichen AMNOG-Verfahren betont). <u>Vorgeschlagene Änderung:</u> Keine. Völlige Zustimmung zu den Aussagen – auch, dass ggfs. Aussagen zu betreffenden Endpunkten bei Missachtung der Forderung nicht verwendbar sind.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

A.1.7 Bayer Vital GmbH

Autorinnen und Autoren

- Dintsios, Charalabos-Markos
- Caruso, Simone

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Dr.Dr. Charalabos-Markos Dintsios, Head Strategic Market Access Intelligence, Bayer Vital GmbH
Dr. Simone Caruso, Strategic Market Access Intelligence, Bayer Vital GmbH
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Bayer Vital GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 1.3	1.3 Evidenzbasierte Medizin Zur allgemeinen Definition: Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) arbeitet seit 2004 auf Basis der Prinzipien der Evidenzbasierten Medizin, um den Entscheidungsträgern im Gesundheitswesen (und für Patienten) wertvolle Informationen für die Beurteilung wissenschaftlicher Evidenz zur Verfügung zu stellen. Obwohl das überarbeitete Kapitel 1.3 zwischen der praktischen EBM und der evidenzbasierten Gesundheitsversorgung zu unterscheiden versucht, so ist diese Darstellung in unseren Augen nicht sachgerecht: im ursprünglichen Sinne ist die EBM explizit auf die Entscheidung hinsichtlich eines Individuums ausgerichtet, und soll ausdrücklich neben dem aktuellen externen Wissensstand gleichberechtigt die Erfahrung des Arztes sowie die individuellen Bedürfnisse und Werte des Patienten in Achtung nehmen. Da die Berichte des Institutes lediglich eine Entscheidungsgrundlage für die Auftraggeber darstellen, sollte auch das Methodenpapier bei der Darstellung der Arbeitsweise deutlich darauf verweisen, dass es sich der Methodik der EBM bedient, jedoch nur einen der drei gleichwertigen Bausteine (externe Evidenz) für die Entscheidungsträger liefern kann. Eine Entscheidung nur auf Basis der Evidenz zu treffen wäre in der ursprünglichen Bedeutung der EBM jedoch ebenso falsch, wie rein auf der Meinung des Individuums (Sackett et al. 1996). Dies sollte in dem vorliegenden Kapitel auch deutlich zum Ausdruck kommen,

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	indem die praktische EBM als die eigentliche evidenzbasierte Medizin genannt wird, deren Arbeitsweisen sich auf die Bewertung von Gruppen von Patienten zwar übertragen lassen, jedoch dem Entscheidungsträger sein Recht nicht nehmen können, jeden Einzelfall zu beurteilen. Zur bestverfügbaren Evidenz: Die stetige Weiterentwicklung der medizinischen Versorgung hat bei vielen Krankheiten bereits zu großen Fortschritten geführt, jedoch bestehen für den Patienten häufige Einschränkungen, die ihm eine normale soziale Teilhabe nicht erlauben und folglich gesellschaftliche Ressourcen binden. Der medizinische und technische Fortschritt treibt auch die Weiterentwicklung der EBM, so dass auch kleinere Effekte oder Evidenz aus versorgungsnahen Daten bei der systematischen Nutzenbewertung methodisch sachgerecht aufbereitet und entsprechend gewürdigt werden sollten.
Abschnitt 1.4	1.4 Gesundheitsökonomie Im einführenden Kapitel zur Gesundheitsökonomie wird die ‚Angemessenheit‘ mit der Kosten-Nutzen-Bewertung (KNB) assoziiert und die ‚Zumutbarkeit‘ wiederum mit der Ausgabeneinflussanalyse (BIA). Bei beiden Begriffen handelt es sich um unbestimmte Rechtsbegriffe aus dem Sozialgesetzbuch V, die durch den Entscheidungsträger, hier dem Gemeinsamen Bundesausschuss (G-BA), näher spezifiziert werden sollten, da dieser neben der Sozialgerichtsbarkeit, insofern keine begleitende Rechtsexegese durch eine Verordnung der Exekutiven erfolgt ist, die entsprechende Hermeneutik betreibt. Dies scheint auch insofern notwendig, als bis dato kein Auftrag durch den G-BA hinsichtlich einer KNB erfolgt ist und somit die Deutung der Begriffe in diesem Kontext offenbleibt. Dies bedeutet nicht im Umkehrschluss, dass eine KNB und BIA nicht im Rahmen der gesundheitsökonomischen Evaluation angebracht sind, sondern dass diese nicht anhand der beiden Begriffe ‚Angemessenheit‘ und ‚Zumutbarkeit‘ rechtlich abgeleitet bzw. legitimiert werden müssen, ist die KNB genuin im SGB V enthalten. Des Weiteren wird in diesem Abschnitt auf die Nichtanwendung eines universellen Schwellenwertes Bezug genommen und erklärt, dass dies nicht Gegenstand der wissenschaftlichen Bewertung durch das IQWiG sei. Ferner wird darauf hingewiesen, dass in Deutschland kein allgemein akzeptierter Schwellenwert existiert. Diesen Verlautbarungen folgend, kann postuliert werden, dass sich das

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	IQWiG hier höchstwahrscheinlich auf Schwellenwerte im Rahmen von Kosten-Nutzwert-Analysen bezieht, die auf der Effektseite QALYs abbilden. Soll sich daraus eine Unterstützung für einen Entscheidungsträger ableiten lassen, ist aus der gesundheitsökonomischen Literatur hinlänglich bekannt, dass sowohl ein impliziter oder expliziter Schwellenwert notwendig ist als auch eine Aussage zum verfügbaren Budget, um dem Opportunitätskostenansatz folgend eine effiziente Mittelallokation zu ermöglichen (diese Thematik wird bereits in gesundheitsökonomischen Lehrbüchern extensiv abgehandelt). Soll der Entscheidungsträger nur mit inkrementellen Kosten-Effektivitäts-Verhältnissen (ICERs) bzw. Kosten-Nutzwert-Verhältnissen (ICURs) zu den zu bewerteten Gesundheitstechnologien (hier primär Arzneimittel im Rahmen einer der frühen Nutzenbewertung folgenden KNB) ergänzend bzw. supportiv informiert werden bzw. die Ergebnisse der KNB darüber hinaus in die entsprechenden Erstattungsbetragsverhandlungen einfließen, sei darauf hingewiesen, dass aus dem mittlerweile als Lehrbuchwissen geltendem Oregonexperiment (Oregon Health Plan) bekannt ist, dass Entscheidungsträger bei einer einfachen Darstellung von ICURs geneigt sind, die Mittelallokation so zu gestalten, dass sie die Erstattung der gesundheitsökonomisch anhand dieser ICERs bzw. ICURs effizientesten Technologie priorisieren (League Tables), womit das ursprüngliche Experiment auch gescheitert ist, da es zu Priorisierungen gekommen ist, die keiner gesellschaftlichen Akzeptanz sowie keiner medizinischen Rationale folgten bzw. im weiteren Verlauf zwangsläufig nicht mehr alleinig auf das Kosten-Effektivitäts-Kriterium Bezug genommen wurde (Tengs et al. 1996), bspw. aufgrund des Prominence Effekten (Persson et al. 2023). Des Weiteren sei bereits hier darauf hingewiesen, dass sowohl direkte Erhebungsinstrumente von Nutzwerten als auch indirekte Erhebungen bspw. über den EQ-5D erhebliche methodische Herausforderungen mit sich bringen und bei Letzteren auch die Änderungssensitivität für unterschiedliche Krankheitsbilder zu hinterfragen ist (siehe aktuell exemplarisch zur Depression (Stolz et al. 2022)). Diese methodischen Fragestellungen können dazu führen, dass es zu Verzerrungen hinsichtlich der Erstattung von Gesundheits-technologien kommen kann, da der gemeinsame Nenner QALY eben nicht ohne Weiteres geeignet ist, durch Gegenüberstellung oder Rangreihenfolgen eine valide Preisinformation oder einen Mittelallokationsmechanismus zu gewährleisten.
Abschnitt 3.3.4	3.3.4 Anwendungsbegleitende Datenerhebung

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Die Anwendungsbegleitende Datenerhebung (AbD) zur Nutzenbewertung gemäß § 35a SGB V ist als neues Konzept zu begrüßen, da sie grundsätzlich eine Möglichkeit der Evidenzgenerierung nach der Zulassung zum Zwecke der Nutzenbewertung bietet und somit auf besondere Therapiesituationen eingeht, die z.B. eine randomisierte kontrollierte Studie unmöglich machen. Zum allgemeinen Vorgehen: Der G-BA kann eine AbD Konzeptionierung und Planung bereits vor der Zulassung des betreffenden Arzneimittels beauftragen, so dass mit dem Tag der Markteinführung theoretisch die AbD beginnen kann. Dies ist vor allem im Falle von ATMP von großer Bedeutung, um eine vollständige und vollzählige Erfassung aller behandelten Patienten zu ermöglichen. Jedoch muss für die Konzepterstellung und der anschließenden Protokoll-Erstellung durch den pU mehr Zeit eingeplant werden. Bayer nimmt daher wie folgt Stellung: die Konzeptionsphase sollte in einem Zeitrahmen geschehen, der es dem pU ermöglicht, sich in angemessener Weise (z.B. durch Konzeptvorschläge oder Stellungnahmen), zu beteiligen. Dies soll sicherstellen, dass neben den unerlässlichen Anforderungen der EBM auch die praktische Durchführbarkeit der AbD gewährleistet ist. Das Augenmerk liegt dabei vor allem in einer sinnvollen und überschaubaren Anzahl von zu kontrollierenden Confoundern, unter Beachtung der tatsächlich erhebbaren und nicht der theoretisch möglichen Parameter. Auch muss die Fallzahlplanung, die Studiendauer und der zu erwartende Effekt den tatsächlichen Rekrutierungsmöglichkeiten in der vorliegenden (Register)-Plattform entsprechen, also einen machbaren Rahmen aufweisen. Die Fragestellung der grundsätzlichen Realisierbarkeit (Fallzahl, Zeitdauer, Verwendung vorhandener Register) sollte daher noch vor der eigentlichen Konzepterstellung nach PICO für die AbD beantwortet werden. Die frühzeitige Einbindung von Experten und dem betroffenen pU ist ebenfalls eine notwendige Voraussetzung, um die eventuell benötigten Absprachen mit dem Registerbetreiber zu treffen, beispielsweise um fehlende Parameter in die Erhebung aufzunehmen, notwendige Qualitätssicherungs-Maßnahmen zu ergänzen oder auch den Zugang zu den Daten zu ermöglichen. Letzteres ist vor dem Hintergrund der Datenaufbereitung für die spätere Nutzendossier-Erstellung von zentraler Bedeutung, um eine fristgerechte Dossier-Aufbereitung nicht zu gefährden. Zum erwarteten Effekt: Dass bei einer nicht randomisierten Studie, wie im Rahmen der AbD vorgesehen, ein Effektmaß in einem Bereich mit hoher Effektstärke notwendig ist, um einen nicht allein durch Verzerrung erreichten Effekt auszuschließen, ist eine nachvollziehbare und

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	ebenso durch die EMA gelebte Praxis (Djulbegovic et al. 2018). Jedoch sollten nicht allein das Effektmaß und die Lage des Konfidenzintervalls in einem solchen Bereich eine Entscheidung über einen Zusatznutzen begründen. Weitere zu berücksichtigende Faktoren sind etwa die Konsistenz des Effektes bei allen Patienten, das Vorhandensein eines Dose-Response-Gradienten, die Objektivität von patienten-relevanten Endpunkten oder eine plausible Einschätzung über die Auswirkung unbekannter und daher nicht adjustierbarer Confounder (Guyatt et al. 2011).
Abschnitt 4.1	4.1 Einführung Laut Methodenentwurf 7.0 umfasst die KNB gemäß § 35b SGB V neben der Berechnung des inkrementellen Kosten-Effektivitäts-Verhältnisses (ICER) dessen Einordnung sowie eine Ausgaben-Einfluss-Analyse. Es stellt sich hier die Frage, ob das Auftragsinstitut diese Einordnung überhaupt umsetzen sollte, da es keine wissenschaftliche Rationale dafür gibt und letztlich auch ein Zusammenhang mit den jeweiligen finanziellen Ressourcen für die Betrachtungsperiode existiert, der überhaupt erst eine Einordnung ermöglichen würde. Dies ist aber eine explizite Aufgabe der Selbstverwaltung im deutschen Gesundheitswesen, die auch die Erstattungsentscheidungen im Rahmen des SGB V trifft und nicht des IQWiG. Eine Möglichkeit, einen Optionsraum für die Einordnung durch den Entscheidungsträger bzw. den Kostenträger bei Erstattungsbetragsverhandlungen zu visualisieren, besteht über Kosten-Effektivitäts-Akzeptanz-Kurven (CEAC), welche auch vom IQWiG bzw. den einreichenden pUs aufbereitet werden können.
Abschnitt 4.2	4.2 Kosten-Nutzen-Bewertung von Arzneimitteln gemäß § 35b SGB V Bezugnehmend auf Tabelle 6: Übersicht über die Elemente eines Referenzfalls werden im Folgenden Punkte aufgegriffen, die einer kritischen Würdigung bedürfen. Zum Punkt <i>Effektivität sowie weitere klinische Inputparameter</i>, Unterpunkt <i>Das Ergebnis der KNB stellt das Ergebnis der Nutzenbewertung nicht infrage</i> sollte eine Klärung erfolgen, dass dies in beide Richtungen gelten sollte (Ergebnis der Nutzenbewertung mit Zusatznutzen, KNB ohne Effektunterschiede bzw. Nutzenbewertung ohne Zusatznutzen aber KNB mit Effektunterschieden, wobei letztere Konstellation eher relevant für eine KNB nach § 35a (5a) und nicht nach §35b SGB V ist). Das bereits angegebene Beispiel der geringen Änderungssensitivität des EQ-5D bei unterschiedlichen Krankheitsbildern der Depression könnte eine Konstellation bedingen, laut welcher ein Zusatznutzen auf Basis etablierter Depressionsskalen wie der HADS, MADRS oder PhQ-9 in der Nutzenbewertung ableitbar wäre, hingegen kein

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Unterschied bezogen auf indirekt erhobene QALYs gezeigt werden kann. Zu dem Punkt <i>Nutzwerte</i>, Unterpunkte <i>Für die Berechnung von QALYs sollen die in das Modell einfließenden Nutzwerte auf der Bewertung durch Patientinnen und Patienten basieren und Nutzwerte basierend auf Bewertungen durch die Allgemeinbevölkerung sind als dann sinnvoll, wenn sich die Bewertungen nicht von denen durch Patientinnen und Patienten unterscheiden</i>, ist zu klären, ob Patienten hier im Sinne keiner Erhebung durch Leistungserbringer bzw. Proxies gemeint ist oder direkte Erhebungen von Nutzwerten bei entsprechenden Patienten bzw. indirekte Erhebungen mit Tarifen, die auf Patientenpräferenzen basieren, intendiert. Der zweite Unterpunkt wird kaum aus der Empirie heraus erfüllbar sein, denn ein fehlender Unterschied zwischen indirekten Erhebungen von Nutzwerten auf Basis von Präferenzen von Stichproben aus der Allgemeinbevölkerung und solchen Erhebungen auf Basis von Patientenpräferenzen lässt sich ohne Erhebung bei beiden Gruppen und Ergebnisvergleich nicht postulieren. Tarife für Nutzwerte für indirekte Erhebungen auf Basis von Patientenpräferenzen existieren zudem zurzeit nur in Schweden, wo sie prioritär im Vergleich zu Tarifen auf Basis von Präferenzen der Allgemeinbevölkerung zur Anwendung kommen sollen (Burström et al. 2020). Zum Punkt <i>Inflationsbereinigung und Diskontierung</i>, Unterpunkt <i>Kosten und Nutzen, die länger als 1 Jahr dauern, sind sie nach dem 1. Jahr im Basisfall mit einer identischen konstanten Rate von 3 % pro Jahr zu diskontieren</i> ist aufgrund der Tatsache, dass das theoretische Fundament zur Ableitung einer Diskontierungsrate auf der Zeitpräferenzrate für Konsum fußt, welche über den langfristigen Kapitalmarktzins in Deutschland, abgebildet durch Bundesanleihen, seit der Finanzkrise viel niedriger liegt (es wurden teilweise Negativzinsen erzielt) eventuell eine niedrigere Rate anzulegen. Andere Ansätze präferieren als Referenzgröße das inflationsbereinigte (reale) Bruttoinlandsproduktwachstum (van Hout 1998). Auch diese Referenzgröße für den Diskontierungssatz liegt weit unter 3% zurzeit. Zum Punkt <i>Ergebnisdarstellung</i>, Unterpunkt <i>Die Berechnung des ICER zwischen den beiden betrachteten Therapiealternativen bildet das wesentliche Ergebnis, welches tabellarisch und grafisch dargestellt wird. Zur tabellarischen Darstellung gehört auch die Angabe der absoluten Ergebnisse je Therapie sowie der entsprechenden inkrementellen Unterschiede zwischen den Therapien und Des Weiteren werden die Ergebnisse ergänzender Analysen unter Berücksichtigung der Therapiekosten anderer zVTs dargestellt</i> ist anzumerken, dass für patientenindividuelle Therapien bzw. Therapien nach Maßgabe des Arztes ein mischkalkulatorischer Ansatz für die Komparatoren mit entsprechenden Gewichtungen bspw. über Verordnungsanteile angebracht wäre. Diese Option wird in der Tabelle 6 nicht genannt.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Hinsichtlich der Ergebnisdarstellung ergänzender Analysen unter Berücksichtigung der Therapiekosten anderer vom G-BA ebenfalls bestimmter Komparatoren, ohne auf deren Effektivität einzugehen, ist dieses vorgeschlagene Vorgehen weder zielführend, noch leage artis, da es von keinen Effektunterschieden ausgeht, was in der Realität wohl kaum so anzutreffen sein wird. Selbst unter den Vertretern von gleichen Wirkstoffklassen zeigen sich Effektivitätsunterschiede sei es auf der positiven oder negativen Effektseite (bspw. sind Glitazone hinsichtlich Nebenwirkungen unterschiedlich, mit dem Resultat, dass die Zulassung von Rosiglitazon 2010 zurückgerufen wurde). Ferner existieren Indikationsgebiete, in welchen bereits mehrere Arzneimittel einer frühen Nutzenbewertung unterzogen wurden, die zu unterschiedlichen Ausprägungen des Zusatznutzens geführt haben. Beispielhaft sei hier das Indikationsgebiet Prostatakarzinom (nicht metastasiert, hohes Metastaserisiko) genannt, in welchem bereits bewertete Arzneimittel (Darolutamid, Apalutamid und Enzalutamid) sowohl einen geringen als auch einen beträchtlichen Zusatznutzen erzielt haben. Aus diesem Grund ist eine derartige ergänzende Analyse hinfällig.
Abschnitt 4.3.1	4.3.1 Population Da sich eine gesundheitsökonomische Evaluation auch auf Gesundheitstechnologien beziehen kann, die im Rahmen des ThemenChecks bewertet werden, ist eine alleinige Einschränkung auf Arzneimittel nicht nachvollziehbar.
Abschnitt 4.3.2	4.3.2 Intervention und Komparator(en) <i>Um die Bandbreite der Therapiekosten zu berücksichtigen, werden im Rahmen von ergänzenden Analysen inkrementelle Kosten-Effektivitäts-Verhältnisse für weitere von der zVT umfasste Therapieoptionen (z. B die günstigste und die teuerste Therapie) im Sinne einer Kostenvariation berechnet werden (Szenarioanalyse zVT-Kosten). Dies kann nur erfolgen, sofern die verschiedenen Therapien als gleichermaßen zweckmäßig anzusehen sind und nicht davon auszugehen ist, dass die Krankheitskosten zwischen den Therapien erheblich variieren (siehe Abschnitt 4.9). Wie bereits in der Kommentierung zu 4.2 ausgeführt wurde, ist das vorgeschlagene Vorgehen nicht zielführend. Die Kostenvariation kann auf der Kostenseite über Sensitivitätsanalysen untersucht werden, eine Annahme zu gleichbleibenden Effekten unterschiedlicher Komparatoren ist nicht realistisch und die Abbildung der potenziellen Kostenvariation über ICERs</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	unangebracht, da hierdurch Szenarien simuliert werden, die nicht entscheidungsrelevant für eine Erstattungsentscheidung bzw. Erstattungsbetragsverhandlungen sind. Wenn die Kosteneffektivität zu einem dezidierten Indikationsgebiet visualisiert werden soll, bietet sich die Effizienzgrenze an, ohne dass aus dieser ein Schwellenwert abgeleitet werden muss. Diese ermöglicht die Einbeziehung beider konstituierender Größen der Kosteneffektivität (Kosten und Effekte). Therapien, die im Rahmen des SGB V erstattet werden, sind zwangsläufig gleichermaßen zweckmäßig. Die Zweckmäßigkeit zumindest nach dem Wirtschaftlichkeitsgebot (§12 SGB V) stellt eine dichotome Größe dar. Eine ungleichmäßige Zweckmäßigkeit lässt sich sozialrechtlich nicht ableiten.
Abschnitt 4.3.3	4.3.3 Gesundheitsökonomische Zielgröße <i>Insbesondere in den Fällen, in denen der Beschluss des Zusatznutzens gemäß § 35a SGB V maßgeblich auf 1 dominierenden Endpunkt aus der Nutzenbewertung beruht, könnten die Kosten in Relation zu diesem Endpunkt die gesundheitsökonomische Zielgröße der KNB bilden. Wenn dem so sein sollte, stellt sich die Frage, wieso überhaupt Kosten-Nutzwert-Analysen im Methodenentwurf 7.0 als Option der KNB abgehandelt werden, da QALYs zumindest bei der frühen Nutzenbewertung von Arzneimitteln nach §35a SGB V weder vom Auftragsinstitut noch vom G-BA als Outcomes akzeptiert werden und die Nutzenbewertungen auf klinischen Endpunkten fußen. Ferner ist der Gesamtnutzen als Basis für die (Zusatz-)Nutzenausprägung über die Nutzen-Schaden-Abwägung abzuleiten, womit es rein konzeptionell keinen dominierenden Endpunkt als solchen geben kann, da dieser insofern theoretisch existent mittels der Gegenüberstellung mit dem komparativen Schadenspotenzial der Intervention wiederum relativiert wird.</i> <i>Es werden in der Literatur verschiedene Limitationen des QALY-Konzepts diskutiert, insbesondere ethische, methodische sowie konzeptionelle Einschränkungen [Referenzen]. Dazu gehört u. a. auch die Diskussion um die Frage wessen Präferenzen bei der Bestimmung von QALYs zugrunde gelegt werden sollen (siehe Abschnitt 4.6). Das QALY-Konzept stellt dennoch die derzeit am umfassendsten angewandte Methodik dar, die explizit und mit einem theoretisch begründeten Konstrukt die Bewertung der Effektivität von Therapien in gesundheits-ökonomischen Evaluationen unter Berücksichtigung von Präferenzen ermöglicht [Referenzen]. Der Aussage ist zwar zuzustimmen, dennoch ist der Verbreitungsgrad einer angewandten Methodik kein Garant für ihre Richtigkeit,</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	insbesondere dadurch, dass die auf die Utility Theory von Neumann und die multi-attribute Utility Theory axiomatisch fußenden Annahmen zum QALY (Risikoneutralität, "A QALY is a QALY", Mutual utility independence, Additive utility independence) empirisch mehrmals widerlegt wurden (exemplarisch (Dolan 2011; Duru et al. 2002; McGregor & Caro 2006)). Eine Kosten-Nutzwert-Analyse sollte aus diesen Gründen mit einer auf klinischen Endpunkten fußenden Kosten-Effektivität immer begleitet werden, um Verzerrungen oder Trugschlüssen, die aus dem QALY-Konstrukt herrühren, nicht aufzusitzen. Darüber hinaus schließt die Aussage im Methodenentwurf 7.0 QALYs auf Basis von VAS-abgeleiteten Nutzwerten aus, da diese nicht präferenzbasiert sind.
Abschnitt 4.3.4	4.3.4 Perspektive Insofern als Auftraggeber von KNB die Selbstverwaltung im Gesundheitswesen fungiert, ist die ergänzende Einnahme der gesellschaftlichen Perspektive nicht zielführend, da der G-BA inklusive seiner Bänke (Leistungserbringer und Kostenträger) keinen über das gesundheitspolitische Feld hinausgehenden Auftrag als untergesetzlicher Normengeber erfüllt. Anders gestaltet sich die Sachlage bei einer rein theoretisch möglichen Beauftragung einer KNB durch das BMG. Ähnliches gilt für die Perspektive einzelner Sozialversicherungs-träger, deren Wirken nicht im SGB V sondern in anderen Sozialgesetzbüchern geregelt ist.
Abschnitt 4.6	4.6 Nutzwerte (im Fall einer Kosten-Nutzwert-Analyse) Unabhängig von der Sinnhaftigkeit der Bewertung durch die Allgemeinbevölkerung oder betroffener Patienten stellt sich die Frage nach der Erhebung von Nutzwerten per se. Direkte Erhebungsmethoden (TTO und das in der ökonomischen Theorie fundierte SG) mögen zur Entwicklung von Tarifen durchaus angebracht sein, werden aber im Rahmen von klinischen Studien aufgrund des enormen Aufwandes und der intellektuellen (Über-)Beanspruchung potenzieller Zielpopulationen nicht angewendet. Bis jetzt liegen wiederum anders als in Schweden keine auf die Bewertung von betroffenen Patienten fußende Tarife für indirekte Nutzwertenerhebungen vor. Diese zu entwickeln und zu publizieren ist auch nicht die Aufgabe der pharmazeutischen Industrie, sondern entsprechender akademischer Gruppen wie bspw. der EuroQoL-Gruppe und anderen. Das hehre Ziel der Berücksichtigung von Präferenzen betroffener Patienten bei der Erhebung von Nutzwerten scheitert somit an dem Nichtvorhandensein

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	entsprechender Tarife für Deutschland und an der fehlenden Implementierung direkter Nutzwertenerhebungen im Rahmen von klinischen Studien.
Abschnitt 4.7.1	4.7.1 Kostenkategorien Bei den gemachten Literaturverweisen fehlt u.a. eine zwar ältere, aber dennoch spezifisch auf die Perspektive der GKV ausgerichtete Publikation, die viel detaillierter auf kalkulatorische Schritte zu vielen Kostenkategorien als die referenzierte Literatur eingeht, und deren kalkulatorischen Grundprinzipien teilweise auch heute noch Geltung besitzen (Krauth et al. 2005). E) Zukünftige Kosten Die Berücksichtigung von nicht interventionsassoziierten zukünftigen Kosten im Rahmen separater Sensitivitätsanalysen stellt ein weiteres hehres Ziel dar, das sich in praxi ohne starke Annahmen und damit nicht belastbarer Datensätze kaum umsetzen lässt. F) Investitions- und Implementierungskosten Der Abschnitt ist zu abstrakt gehalten. Eine angemessene Berücksichtigung im Rahmen von Sensitivitätsanalysen ohne weiterführende Informationen, wie diese auszulegen ist, erschwert die Berücksichtigung solcher Kosten, insofern relevant für die Kostenbetrachtung. Ferner stellen Investitionskosten je nach Konstellation in der Kostentheorie auch versunkene Kosten dar oder werden über die Zeit abgeschrieben oder über Fixkostendegression marginalisiert, was wiederum für Langzeitbetrachtungen von Relevanz ist.
Abschnitt 4.7.2	4.7.2 Schritte zur Kostenbestimmung Erneut wird auf die gesellschaftliche Perspektive bei den Kostenkomponenten abgestellt. Diese ist aber nicht Gegenstand des SGB V und somit auch nicht des Auftragsinstitutes. Die Verweise auf die gesellschaftliche Perspektive und alle damit zusammenhängenden Schritte sollten aus diesem Grund nicht Gegenstand des Methodenpapiers sein. Sollten für die Bewertung einer Technologie Arbeitsunfähigkeitszeiten relevant sein, können diese auch auf der Effektseite als eigenständiger Outcome abgebildet werden. Dies ermöglicht auch die Vermeidung der Bestimmung von Krankengeldzahlungen aus der engen GKV Perspektive.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Bei dem Abschnitt Bewertung der Ressourcen lautet es <i>Wird der Zeitaufwand von Betroffenen oder Angehörigen in der Kostenbestimmung berücksichtigt, wird dieser mit dem Nettolohn bewertet.</i> Dabei wird nicht weiter spezifiziert wessen Nettolohn. Dies könnte sich auf den durchschnittlichen Nettolohn in Deutschland beziehen, aber genauso wiederum geschlechtergetrennt oder branchenspezifisch. Eine Klarstellung des IQWiG sollte hier erfolgen.
Abschnitt 4.8	4.8 Inflationsbereinigung und Diskontierung Da sich die Diskontierungsrate an den Kapitalkosten orientiert, welche wiederum an langfristigen Zinsen für Bundesanleihen approximativ abgebildet werden können, ist die Rate von 3% mittlerweile zu hoch angesetzt. Eine Näherung an den Durchschnitt der Zinsen über die letzten 12 Monate würde die Realität auf den Kapitalmärkten besser abbilden. Auswahl der Modellierungstechnik <i>Die Datenlage allein sollte nicht die Wahl der Modellierungstechnik bestimmen.</i> Dieser Aussage kann gefolgt werden, allerdings sollte auch bedacht werden, dass bei einer ungünstigen Datenlage mit viel mehr Annahmen modelliert werden muss, womit ein elaboriertes Modell schlechtere Ergebnisse liefern kann als ein aufgrund der Datenlage simplifizierendes Modell.
Abschnitt 4.10	4.10 Ergebnisdarstellung <i>Die von der zVT umfassten Therapieoptionen können unterschiedlich hohe Therapiekosten aufweisen. Um diese Bandbreite zu berücksichtigen, muss daher nicht nur das ICER für den Vergleich zwischen dem zu bewertenden Arzneimittel und der zum Nachweis des Zusatznutzens ausgewählten Vergleichstherapie berechnet werden, sondern auch für die Vergleiche zwischen dem zu bewertenden Arzneimittel und den anderen von der zVT umfassten Therapien (Szenarioanalyse zVT-Kosten). Diese ergänzenden Analysen erfolgen auf Grundlage der Therapiekosten der anderen von der zVT umfassten Therapien (z. B die günstigste und die teuerste Therapie). Diese Szenarioanalyse zVT-Kosten findet unter den Annahmen statt, dass die inkrementelle Effektivität des zu bewertenden Arzneimittels im Vergleich zu den verschiedenen von der zVT umfassten Therapien jeweils gleich ist und die Krankheitskosten nicht erheblich voneinander abweichen. Das vorgeschlagene Vorgehen ist abzulehnen. Es stellt keine</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	eigentliche Szenarioanalyse dar und arbeitet mit zwei sehr starken Annahmen (i. inkrementelle Effektivität der zu bewertenden Technologie im Vergleich zu allen vom G-BA bestimmten zVT Optionen sind gleich sowie ii. eine geringe Abweichung der Krankheitskosten). Insbesondere die dargestellte Grafik (Abbildung 15) suggeriert durch unterschiedliche, nicht in der Realität existente ICER, einen Optionsraum für eine Erstattungsentscheidungsfindung, die auf das geringste ICER abstellt. Da das ICER Kosteninkremente in Relation zu Effektinkrementen setzt, müsste man im Rahmen einer Szenarioanalyse auch die Effektseite variieren. Das unternehmensseitige Auswahlkriterium für die Komparatorauswahl ist bereits durch die klinische Studienevidenz zu Zulassungszwecken determiniert und orientiert sich somit nicht an dem ICER-Vergleich unter den potenziellen Alternativen. Zulassungsstudien bedienen eine Reihe von Ländern gleichzeitig und richten sich nicht nach den Erstattungsgegebenheiten in Deutschland, zumindest für international aktive pharmazeutische Unternehmen. Ferner fehlen auch jegliche Verweise aus der gesundheits-ökonomischen Literatur für eine derartige Ergebnisdarstellung. Wäre man an der Verteilung aller potenziellen Komparatoren auf der Kosten-Effektivitäts-Ebene interessiert, so böte sich dafür wiederum der Effizienzgrenzansatz ohne Ableitung einer ad hoc Schwellengröße aus der Effizienzgrenze.
Abschnitt 4.12	4.12 Einordnung des ICERs Hinsichtlich der institutseigenen Deutung der unbestimmten Rechtsbegriffe aus dem SGB V Angemessenheit und Zumutbarkeit ist bereits weiter oben Stellung genommen worden. Weiter lautet es in diesem Abschnitt <i>Die Darstellung gesundheitsökonomischer Evaluationen aus anderen Ländern ist auftragsspezifisch zu prüfen</i>. Obwohl wenige HTA Agenturen international in ihren methodischen Empfehlungen auf die Übertragbarkeit von gesundheits-ökonomischen Evaluationen aus anderen Kontexten eingehen (Weise et al. 2022) gibt es auch hierfür methodische Ansätze wie bspw. Bayesianische shrinkage Schätzungen, auf die sich eine Task Force der ISPOR (Transferability of Economic Evaluations Across Jurisdictions) bereits 2009 bezogen hat (Drummond et al. 2009). Diese Ansätze sollten entweder im Methodenpapier kurz erläutert werden oder zumindest über entsprechende Literatur-verweise genannt werden.
Abschnitt 9.1.1	9.1.1 Kriterien für den Einschluss von Studien

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	1. Formulierungsvorschlag: <i>„Häufig ergibt sich das Problem, dass für eine Nutzenbewertung relevante Studien die Einschlusskriterien für die in der systematischen Übersicht und der Fragestellung der Nutzenbewertung definierte Patientenpopulation bzw. die definierten Prüf- und Vergleichsinterventionen nicht vollständig erfüllen. Das Institut verfährt hier in der Regel nach den folgenden Kriterien.</i> <i>Liegen für eine Studie Analysen der relevanten Teilpopulation bzw. der relevanten Prüf- und Vergleichsinterventionen vor, wird auf diese Analysen zurückgegriffen.“</i> Allgemeiner Kommentar: Der ursprüngliche Sinn des 80% Kriteriums geht im neuen Text (gegenüber der alten Version 6.1) verloren. Entweder er besteht nicht mehr und eine relevante Teilpopulation kann auch kleiner als 80% sein, oder der Absatz sollte exakter formuliert werden.
Abschnitt 9.3.2	9.3.2 Statistische Signifikanztests und Konfidenzintervalle Wir begrüßen die Erweiterung der methodischen Bandbreite um bayessche Methoden, die für spezifische Situationen, in denen die üblichen frequentistischen Verfahren an ihre Grenzen stoßen, eine Auswertung ermöglichen.
Abschnitt 9.3.6	9.3.6 Adjustierung und Multifaktorielle Verfahren Die Adjustierung für Confounder ist bei nicht-randomisierten Studien notwendig, um den Einfluss ungleich verteilter Variablen in den Studienarmen weitmöglichst auszuschließen. Bei der AbD wie auch bei indirekten Vergleichen einzelner Studienarme gehört daher die sorgfältige Identifizierung und Adjustierung von Confoundern zum methodischen Armamentarium der EBM. Wenn eine AbD oder ein indirekter Vergleich einzelner Studienarme auf bestehenden Registerdaten basiert, so ist die vollumfängliche Erhebung aller systematisch identifizierten Confounder in der Regel unmöglich, da die betreffenden Variablen nicht im Register erhoben werden oder die Dokumentation unvollständig ist (Pufulete et al. 2022). Das mit der angegebenen Referenz beschriebene Vorgehen einer systematischen Literaturrecherche und nachfolgender Expertendiskussion ist für solche Situationen darüber hinaus nicht zweckmäßig, da der Aufwand (SRL plus Expertenbefragung) das Ergebnis laut der Autoren zur Auswertung einer nicht-randomisierten Studie nicht

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	rechtfertigt (Pufulete et al. 2022). Die Confounder-Adjustierung sollte sich demnach nur an den wichtigsten und auch praktisch erheblichen Parametern orientieren, und die Confoundersuche sollte sich je nach Informationslage auf eine der beiden Methoden beschränken.
Abschnitt 9.3.12	9.3.12 Umgang mit variablen Beobachtungszeiten Obwohl der Abschnitt letztlich eine Klarstellung der vom Institut bereits praktizierten Methodik bietet, ist anzumerken, dass die aus Sicht der EBM verständliche Forderung nach vergleichbaren Nachbeobachtungszeiten in der Umsetzung und vor allem bei vulnerablen Populationen an praktische Grenzen bzw. auf geringe Teilnahme aufseiten der Patienten stoßen kann.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

Burström, K., Teni, F. S., Gerdtham, U.-G., Leidl, R., Helgesson, G., Rolfson, O., et al. (2020). Experience-Based Swedish TTO and VAS Value Sets for EQ-5D-5L Health States. *Pharmacoeconomics* 38(8):839-856.

Djulgovic, B., Glasziou, P., Klocksieben, F. A., Reljic, T., VanDenBergh, M., Mhaskar, R., et al. (2018). Larger effect sizes in nonrandomized studies are associated with higher rates of EMA licensing approval. *J Clin Epidemiol* 98:24-32.

Dolan, P. (2011). Thinking about it: thoughts about health and valuing QALYs. *Health Econ* 20(12):1407-1416.

Drummond, M., Barbieri, M., Cook, J., Glick, H. A., Lis, J., Malik, F., et al. (2009). Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value Health* 12(4):409-418.

Duru, G., Auray, J. P., Béresniak, A., Lamure, M., Paine, A., Nicoloyannis, N. (2002). Limitations of the methods used for calculating quality-adjusted life-year values. *Pharmacoeconomics* 20(7):463-473.

Guyatt, G. H., Oxman, A. D., Sultan, S., Glasziou, P., Akl, E. A., Alonso-Coello, P., et al. (2011). GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 64(12):1311-1316.

Krauth, C., Dintsios, C.-M., Brandes, I., Wasem, J. (2005). Die Perspektive der Gesetzlichen Krankenversicherung in der gesundheitsökonomischen Evaluation. Zeitschrift für die gesamte Versicherungswissenschaft 94(2):215-256.

McGregor, M. & Caro, J. J. (2006). QALYs: are they helpful to decision makers? Pharmacoeconomics 24(10):947-952.

Persson, E., Erlandsson, A., Slovic, P., Västfjäll, D., Tinghög, G. (2023). The prominence effect in health-care priority setting. Judgment and Decision Making 17(6):1379-1391.

Pufulete, M., Mahadevan, K., Johnson, T. W., Pithara, C., Redwood, S., Benedetto, U., et al. (2022). Confounders and co-interventions identified in non-randomized studies of interventions. J Clin Epidemiol 148:115-123.

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. Bmj 312(7023):71-72.

Stolz, M., Albus, C., Beutel, M. E., Deter, H. C., Fritzsche, K., Herrmann-Lingen, C., et al. (2022). Assessment of health-related quality of life in individuals with depressive symptoms: validity and responsiveness of the EQ-5D-3L and the SF-6D. Eur J Health Econ.

Tengs, T. O., Meyer, G., Siegel, J. E., Pliskin, J. S., Graham, J. D., Weinstein, M. C. (1996). Oregon's Medicaid ranking and cost-effectiveness: is there any relationship? Med Decis Making 16(2):99-107.

van Hout, B. A. (1998). Discounting costs and effects: a reconsideration. Health Econ 7(7):581-594.

Weise, A., Büchter, R. B., Pieper, D., Mathes, T. (2022). Assessing transferability in systematic reviews of health economic evaluations - a review of methodological guidance. BMC Med Res Methodol 22(1):52.

A.1.8 Boehringer Ingelheim Pharma GmbH & Co. KG

Autorinnen und Autoren

- Elsässer, Amelie
- Emser, Angela
- Engelking, Franziska
- Henschel, Andreas
- Kleine, Eva
- Schepers, Cornelia

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Dr. Amelie Elsässer
Dr. Angela Emser
Dr. Franziska Engelking
Dr. Andreas Henschel
Eva Kleine
Cornelia Schepers
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen) ☒ im Namen folgender Institution / Organisation: Boehringer Ingelheim Pharma GmbH & Co. KG ☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Am 06.12.2022 hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) einen Entwurf mit inhaltlichen Aktualisierungen und Ergänzungen seiner allgemeinen Methoden veröffentlicht und ein Stellungnahmeverfahren veröffentlicht. Boehringer Ingelheim (BI) begrüßt die kontinuierliche Weiterentwicklung dieser allgemeinen Methoden und nimmt im folgenden dazu Stellung.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Kapitel 9, Abschnitt 9.3.2, (S. 170)	<i>„Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an [59], z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.“</i> <u>Anmerkung:</u> Das Beispiel erscheint hier ungeeignet. Schon in Quelle 59 steht, dass die Tests unabhängig sein sollen, was aber für Tests zum gleichen Endpunkt mit verschiedenen Operationalisierungen nicht gegeben ist. Verschiedene Operationalisierungen eines Endpunktes können aus wissenschaftlicher Sicht sinnvoll sein, um die Robustheit der Ergebnisse (Sensitivitätsanalysen) zu untersuchen oder unterschiedliche Fragenstellung zu adressieren (z. B. Gruppenunterschiede basierend auf Mittelwerten vs. Responderanalysen). In diesem Falle ist eine Hauptanalyse zu definieren, wodurch die Notwendigkeit der Adjustierung für multiples Testen entfällt. Im Allgemeinen ist eine Korrektur für multiples Testen nur dann sinnvoll und anwendbar, wenn die zu testenden Hypothesen in ihrer Gesamtheit vorab präspezifiziert sind. Darüber hinaus ist die alpha-Inflation gerade bei positiv korrelierten Endpunkten weniger stark ausgeprägt (Pocock, Geller, & Tsiatis, 1987; Wassmer,

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Reitmeir, Kieser, & Walter, 1999). <u>Vorgeschlagene Änderung:</u> <i>„Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an [59], z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.“</i>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

[1]. Pocock, S. J., Geller, N. L., & Tsiatis, A. A. (1987). The Analysis of Multiple Endpoints in Clinical Trials. *Biometrics*, 43(3), 487. <https://doi.org/10.2307/2531989>

[2]. Wassmer, G., Reitmeir, P., Kieser, M., & Walter, L. (1999). Procedures for testing multiple endpoints in clinical trials: An overview, 82(1), 69–81. [https://doi.org/10.1016/s0378-3758\(99\)00032-4](https://doi.org/10.1016/s0378-3758(99)00032-4)

A.1.9 Bristol-Myers Squibb GmbH & Co. KGaA

Autorinnen und Autoren

- Bluhmki, Tobias
- Böhm, Doris
- Ellis, Aneurin
- Kähm, Katharina

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Bluhmki, Tobias; Dr.
Böhm, Doris
Ellis, Aneurin
Kähm, Katharina; Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Bristol-Myers Squibb GmbH & Co. KGaA
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Mit dem Entwurf der Allgemeinen Methoden 7.0 führt das IQWiG verschiedene Änderungen ein, zu denen die Bristol-Myers Squibb GmbH & Co. KGaA (im Folgenden: BMS) Stellung nehmen möchte.

Neben Anpassungen im Bereich der Arzneimittel-Nutzenbewertung, die in erster Linie den Einschluss von Studien, die Beurteilung statistischer Signifikanz und den Umgang mit variablen Beobachtungszeiten betreffen, erweitert das IQWiG sein Methodenpapier mit dem vorliegenden Entwurf erstmal auch um Konzepte zur anwendungsbegleitenden Datenerhebung (AbD), von denen sich BMS als pharmazeutischer Unternehmer unmittelbar betroffen sieht.

Darüber hinaus präsentiert das IQWiG grundlegende Überarbeitungen im Kapitel zur Kosten-Nutzen-Bewertung. In Anbetracht der seit Einführung der AMNOG-Regelungen fehlenden praktischen Relevanz von Kosten-Nutzen-Bewertungen, sind die Änderungen aus Sicht von BMS eher von theoretischer als von praktischer Bedeutung. Ungeachtet der zweifelhaften Relevanz ist aus Sicht von BMS zu würdigen, dass sich das IQWiG bei seinen Anpassungen nun stärker an internationalen gesundheitsökonomischen Leitlinien zu orientieren versucht. Die Abkehr vom Konzept der Effizienzgrenzen wird seitens BMS begrüßt.

Im Folgenden wird auf die genannten Aspekte detaillierter eingegangen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 3.3.4 (S. 66 ff.)	<u>Anmerkung:</u> Die im Entwurf des IQWiG-Methodenpapier skizzierten Inhalte des AbD-Konzepts entsprechen im Wesentlichen dem Rapid Report A19-43. Erstaunlich ist an dieser Stelle, dass die Beschreibung des Verfahrens und der Methodik im Zuge der bisher entwickelten Konzepte und Beschlüsse zu AbDs keine Anpassungen erfahren haben. BMS möchte auf folgende Aspekte aufmerksam machen, die im Rahmen der Erläuterungen im Entwurf des Methodenpapiers nur

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	unzureichend thematisiert werden und für die eine weitere Ausarbeitung wünschenswert wäre: <i>Fallzahlplanung</i> Um zu prüfen, ob die AbD prinzipiell möglich ist, gibt das IQWiG an, für die Fallzahlabeschätzung epidemiologische Daten wie die Prävalenz und Inzidenz der Erkrankung sowie die Vollständigkeit der Datenquelle (Abdeckung des Registers / Vollständigkeit der Erkrankungsmeldungen) zu berücksichtigen. Unberücksichtigt bleibt hingegen die aktuelle Versorgungssituation im Hinblick auf die Verteilung der Patienten auf die zu vergleichenden Behandlungsoptionen. Das vom IQWiG in bisherigen Konzepten weitgehend angenommene Verhältnis von 1:1 ist im nicht-interventionellen Kontext eher unwahrscheinlich und im Hinblick auf die Fallzahlplanung oft zu optimistisch, insbesondere dann, wenn es sich bei der Vergleichstherapie um eine im Versorgungsalltag etablierte Therapie handelt. <i>Test einer verschobenen Nullhypothese</i> Um Unsicherheiten bei nicht-randomisierten Vergleichen zu berücksichtigen, fordert das IQWiG einen großen Behandlungseffekt und schlägt den Ansatz des Testens einer verschobenen Nullhypothese vor. Der postulierte Schwellenwert für einen relativen Effekt, der aus Sicht des IQWiG groß genug ist, um Unsicherheiten auszuräumen – das IQWiG schlägt für das relative Risiko eine Schwelle im Bereich von 2 bis 5 in Bezug auf die untere Grenze des Konfidenzintervalls vor – ist wissenschaftlich nicht anerkannt und berücksichtigt weder die Art der Erkrankung noch die Art des betrachteten Endpunkts. Nach Ansicht von BMS sollten, nach erfolgter Adjustierung, verbliebene Unsicherheiten vielmehr durch die Überprüfung der getroffenen Annahmen zur Strukturgleichheit sowie anhand geeigneter Sensitivitätsanalysen bewertet werden. Insgesamt bleibt unklar, inwieweit der Zusatznutzen auf Basis der methodischen Vorgaben des IQWiG zu quantifizieren ist. <u>Vorgeschlagene Änderung:</u> Zur Überprüfung der prinzipiellen Realisierbarkeit der AbD sollten die Überlegungen zur Fallzahlplanung alle relevanten Aspekte einbeziehen. Neben epidemiologischen Kennzahlen und der Abdeckung der identifizierten Register sollte auch weiteren Aspekte, wie der Anzahl der zu berücksichtigenden Confounder und der aktuellen Versorgungssituation Rechnung getragen werden. Die Methodik zur Untersuchung von Unsicherheiten im Zusammenhang mit nicht-randomisierten Therapievergleichen sollte ergänzt werden. Die alleinige Forderung einer ausreichenden Effektstärke

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	im Zusammenhang mit der Methodik des Testens einer verschobenen Nullhypothese erscheint BMS nicht zielführend.
Abschnitt 4 (S. 83 ff.)	<u>Anmerkung:</u> Für den Prozess der Kosten-Nutzen-Bewertung ist aus Sicht von BMS ein hohes Maß an Verfahrenstransparenz sicherzustellen. Das IQWiG-Methodenpapier macht keinerlei Aussagen zum Prozess einer Kosten-Nutzen-Bewertung. <u>Vorgeschlagene Änderung:</u> Es sollte ein eigener Abschnitt zum möglichen Verfahrensprozess ergänzt werden. Vor dem Hintergrund der gesetzlichen Transparenzanforderungen ist beispielsweise ein Vorgehen ähnlich dem englischen NICE denkbar.
Abschnitt 4.3.4 (S. 89)	<u>Anmerkung:</u> BMS begrüßt, dass für die Kosten-Nutzen-Bewertung gemäß dem Entwurf der Allgemeinen Methoden 7.0 nicht die reine GKV-, sondern eine patientenzentriertere Perspektive (hier konkret: die GKV-Versichertengemeinschaft) eingenommen werden soll. Um möglichst alle positiven und negativen Folgen einer Therapie abbilden zu können, stellt die Gesellschaftsperspektive jedoch den größten gemeinsamen Nenner dar und sollte daher im Referenzfall Anwendung finden [1]. Diese ist auch international am weitesten akzeptiert. Der Vorschlag des IQWiG – eine Beschränkung auf die GKV bzw. GKV-Versicherte – ist auch ökonomisch nicht nachvollziehbar, da GKV-Versicherte z.B. immer auch Pflegeversicherte bzw. Steuerzahler sind. <u>Vorgeschlagene Änderung:</u> In begründeten Fällen sollte eine gesellschaftliche Perspektive zur Anwendung kommen.
Abschnitt 4.5 (S. 90 f.)	<u>Anmerkung:</u> Das Verbot der Suggestierung von Zusatznutzen „in Situationen, in denen im Rahmen des Verfahrens der Nutzenbewertung kein Zusatznutzen nachgewiesen wurde“, wie es das IQWiG formuliert, ist nicht nachzuvollziehen. Das Ziel eines gesundheitsökonomischen Modells sollte es nicht sein, die Nutzenbewertung zu reproduzieren und um weitere Kostenkomponenten zu ergänzen, sondern eine fundierte Entscheidungsunterstützung unter Unsicherheit und unter Berücksichtigung der zu diesem Zeitpunkt besten verfügbaren Evidenz zu Kosten und Nutzen einer Therapie bzw. ggf. mehreren Vergleichstherapien zu liefern.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Vorgeschlagene Änderung:</u> Die Bandbreite der Unsicherheit sollte in beide Richtungen – negativ wie positiv – berücksichtigt werden, dies gilt auch für den Nutzen einer Therapie.
Abschnitt 4.6 (S. 91)	<u>Anmerkung:</u> Trotz inkonklusiver wissenschaftlicher Literatur, wird für den Referenzfall gefordert, dass die für die Berechnung von QALYs in das Modell einfließenden Nutzwerte auf der Bewertung durch Patientinnen und Patienten basieren. Das IQWiG schreibt selbst, „dass aufgrund der international gängigen Praxis, vorzugsweise auf Bewertungen der Allgemeinbevölkerung zurückzugreifen, die Datenlage im Hinblick auf Bewertungen durch Patientinnen und Patienten eingeschränkt sein kann“. Gleichzeitig wird eingeräumt, dass Nutzwerte basierend auf Bewertungen durch die Allgemeinbevölkerung „dann sinnvoll sein können, wenn sich die Bewertungen nachgewiesen nicht von denen durch Patientinnen und Patienten unterscheiden“. Es bleibt dabei unklar, wie ein solcher Nachweis tatsächlich gelingen soll. Es ist bedauerlich, dass auf diese Weise ein Konzept vorgestellt wird, welches in der gesundheitsökonomisch-wissenschaftlichen Praxis wenig praktikabel ist. <u>Vorgeschlagene Änderung:</u> keine
Abschnitt 4.8 (S. 98)	<u>Anmerkung:</u> Die gleiche Diskontierung von Effekten und Kosten untergräbt aus gesellschaftlicher Perspektive den Präventionsgedanken. <u>Vorgeschlagene Änderung:</u> Eine angemessen deutlich niedrigere Diskontierung von Effekten (beispielsweise in Höhe der Hälfte der Diskontierung der Kosten) ist anzuwenden.
Abschnitt 4.12 (S. 107)	<u>Anmerkung:</u> Es muss umfassende Transparenz über die herangezogenen Bewertungskriterien bestehen. Diese Kriterien sollten sowohl belastbar als auch gesellschaftlich akzeptiert sein und den Zugang zu medizinischen Innovationen nicht behindern. Hier bleibt das IQWiG in seinen Formulierungen so ungenau, dass unklar bleibt, wie eine Kosten-Nutzen-Bewertung mit der darin bestehenden und quantifizierten Unsicherheit unter diesen Rahmenbedingungen überhaupt zu einer transparenteren Entscheidungsfindung und Preisbildung beitragen kann.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	<u>Vorgeschlagene Änderung:</u> Aktuell keine praktische Relevanz, daher keine Änderungsvorschläge
Abschnitt 9.1.1 (S. 159)	<u>Anmerkung:</u> In seiner Überarbeitung der Kriterien zum Einschluss von Studien nimmt das IQWiG zunächst vorwiegend redaktionelle Änderungen an der Beschreibung der sogenannten „80 % Regel“ vor, wonach auch solche Studien in ihrer Gesamtheit für die Nutzenbewertung relevant sind, deren Studienpopulation zu mindestens 80 % mit der bewertungsrelevanten Population übereinstimmen bzw. für die sowohl die Prüfintervention als auch die Vergleichsintervention zu mindestens 80 % der zu untersuchenden Fragestellung entsprechen. Dieses Vorgehen ist gängige Praxis und stellt somit keine Neuerung dar. Anders als in der Vorversion des Methodenpapiers fehlt jedoch die Beschreibung des Vorgehens für den Fall, dass weniger als 80 % der Studienpopulation die Einschlusskriterien hinsichtlich Population bzw. Interventions- und Vergleichstherapie erfüllen. Die Möglichkeit, in diesem Fall bei der Bewertung zum Zusatznutzen auf die Ergebnisse der Gesamtpopulation zurückzugreifen – auch wenn Ergebnisse der bewertungsrelevanten Teilpopulation vorliegen – findet im Entwurf des Methodenpapiers 7.0 keine Erwähnung mehr. BMS hätte sich an dieser Stelle eine Konkretisierung und Berücksichtigung der Ergebnisse des IQWiG Berichts zur Übertragbarkeit von Studienergebnissen auf Teilpopulationen (GA18-01, Version 1.0, 20.06.2018) [2] gewünscht, in dem eine Berücksichtigung der Ergebnisse der Gesamtpopulation diskutiert wird für Situationen, in denen • sowohl Informationen zu Ergebnissen der bewertungsrelevanten und der nicht bewertungsrelevanten Teilpopulation vorliegen, • diese Ergebnisse hinreichend homogen sind (gleichgerichtete Effekte; nicht-signifikanter Interaktions- bzw. Heterogenitätstest), • der Patientenanteil in der bewertungsrelevanten Teilpopulation gering ist, sodass von einer relevanten Powerreduktion auszugehen ist, wenn an Stelle der Gesamtpopulation die bewertungsrelevante Teilpopulation betrachtet wird sowie

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	keine klinisch-inhaltlichen Aspekte vorliegen, die gegen die Übertragbarkeit der Ergebnisse der nicht-bewertungsrelevanten Teilpopulation auf die bewertungsrelevante Teilpopulation sprechen. Stattdessen listet das IQWiG Kriterien auf, die – so zumindest die Interpretation des letzten Abschnitts durch BMS – zu überprüfen sind, um die Anwendung der 80 % Regel zu rechtfertigen. Um Unsicherheiten zu vermeiden, wäre es auch wünschenswert, die generisch formulierten Kriterien „Stärke der Effektmodifikation“, „Anteil der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen“ und „Grad der Abweichungen der Interventionen“ näher auszuführen. <u>Vorgeschlagene Änderung:</u> Der Abschnitt sollte um eine Beschreibung ergänzt werden, unter welchen Umständen die Ergebnisse der Gesamtpopulation einer Studie in der Bewertung zum Zusatznutzen berücksichtigt werden können, wenngleich die bewertungsrelevante Teilpopulation weniger als 80 % der Studienpopulation ausmacht und die Ergebnisse dieser Teilpopulation vorliegen. Die Notwendigkeit zusätzlicher Kriterien für den Fall, dass eine Studie das 80%-Kriterium erfüllt, ist für uns nicht ersichtlich. Zudem bleiben die vom IQWiG benannten zusätzlichen Kriterien unklar. Eine Streichung des entsprechenden Absatzes wird angeregt.
Abschnitt 9.3.2 (S. 170)	<u>Anmerkung:</u> Anders als beim formalen, konfirmatorischen Nachweis der Wirksamkeit im Rahmen von Zulassungsstudien wird in der Nutzenbewertung in der Regel eine separate Bewertung der einzelnen Endpunkte ohne Multiplizitätskorrektur vorgenommen. Wie bereits in der Vorversion merkt das IQWiG im Entwurf der Allgemeinen Methoden 7.0 an, dass es, „wenn sinnvoll und möglich“, Methoden zur Adjustierung für multiples Testen in Erwägung zieht. Das IQWiG nennt hier beispielhaft die Anwendung „unterschiedlicher Operationalisierungen von Endpunkten“ als eine Situation, in der eine Adjustierung für multiples Testen sinnvoll sein könnte. Der Begriff der „unterschiedlichen Operationalisierung“ beinhaltet eine große Bandbreite an Variationen in der Betrachtung von Endpunkten. So können unterschiedliche Operationalisierungen eines Endpunktes beispielsweise einen zusätzlichen klinisch-inhaltlichen Nutzenaspekt adressieren oder auch lediglich unterschiedlichen Auswertungsmethoden zu einem Endpunkt

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	betreffen. Durch den Zusatz „wenn möglich und sinnvoll“ bleibt unklar, in welchen konkreten Situationen eine Adjustierung für multiples Testen vom IQWiG angedacht ist. Durch eine entsprechende Adjustierung bei Auswertungen eines Endpunkts mit verschiedenen Operationalisierungen würde ggf. das bisher im Rahmen der Nutzenbewertung etablierte Vorgehen konterkariert, für die Ableitung des Zusatznutzens die relevanteste Operationalisierung, etwa in Bezug auf Mess- und Ergebnissicherheit oder auch in Bezug auf den relevantesten Estimand, auszuwählen und weitere Operationalisierungen nur ergänzend einzubeziehen. <u>Vorgeschlagene Änderung:</u> Aufgrund der beschriebenen Unklarheit sollte eine Löschung des Satz erwogen werden. Andernfalls erscheint zumindest eine hinreichende Konkretisierung, in welchen Fällen eine Adjustierung für multiples Testen durchgeführt werden sollte, notwendig.
Abschnitt 9.3.6 (S. 185)	<u>Anmerkung:</u> In den Ausführungen werden die wichtigsten Informationen zur Darstellung multifaktorieller Verfahren stichpunktartig genannt. Als ein wesentlicher Aspekt nennt das IQWiG die „Angaben zur Selektion von Variablen und Wechselwirkungen“. Aus Sicht von BMS fehlen an dieser Stelle jedoch konkrete Vorschläge zur formal adäquaten Umsetzung. <u>Vorgeschlagene Änderung:</u> In dem benannten Abschnitt wird eine Konkretisierung des Begriffs „Angaben“ sowie eine genauere Einbettung des im gesamten Kapitel beschriebenen Sachverhalts in die Thematik „Anwendungsbegleitende Datenerhebung“ angeregt, z.B. anhand einer Einbettung des durch das IQWiG veröffentlichten Rapid Reports A29-43 (speziell Kapitel 5.5.2, kausale Methoden) [3].
Abschnitt 9.3.12 (S.190 f.)	<u>Anmerkung:</u> In dem Abschnitt zum Umgang mit variablen Beobachtungszeiten wurden Ausführungen zu systematisch verkürzten Beobachtungsdauern aufgrund unvollständiger Datenerfassung bei patientenberichteten Endpunkten ergänzt. Als Beispiel für entsprechende Situationen führt das IQWiG unter anderem die fehlende Erhebung von patientenberichteten Endpunkten nach Abbruch oder Wechsel der Behandlung an. In Bezug auf den Begriff „systematisch verkürzt“ wird nicht erklärt, im Vergleich zu welcher angestrebten Zeitdauer eine Verkürzung zu

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	verstehen ist bzw. worin für das IQWiG eine „unverkürzte“ Beobachtungsdauer besteht. Es bleibt somit offen, ob die „Verkürzung“ seitens des IQWiG etwa im Vergleich zur Beobachtungsdauer bestimmter anderer Endpunkte, im Vergleich zu einem festgelegten Zeitintervall ab Studienbeginn oder im Vergleich zur verbliebenen Lebensdauer der Patientin oder des Patienten bemessen wird. Darüber hinaus erfolgt die Diskussion des Sachverhalts der systematisch verkürzten Beobachtungsdauer ausschließlich im Kontext unterschiedlicher Beobachtungsdauern (so auch in Bezug auf die Kritik an „dauerhaften“ Veränderungen). Je nachdem, was man unter einer „systematisch verkürzten“ Beobachtungsdauer versteht, gibt es jedoch Situationen, in denen sich Unterschiede in der Beobachtungsdauer zwischen den Studienarmen auch bei „unverkürzter“ Beobachtung ergeben oder in denen trotz „verkürzter“ Beobachtung vergleichbare Beobachtungsdauern zwischen den Studienarmen vorliegen. Klärungsbedarf zu diesen Aspekten besteht aus Sicht von BMS insbesondere vor dem Hintergrund der in dem Abschnitt neu eingeführten Aussage, dass im Falle einer systematisch verkürzten Beobachtungsdauer die Ergebnisse der betroffenen Endpunkte in der Regel keinen dominierenden Einfluss auf die Gesamtaussage zum Zusatznutzen haben dürften. Hierbei stellt sich jedoch die Frage, inwieweit Auswertungen über einen möglichst langen Zeithorizont nach Therapieende im Einzelfall (d.h. im Kontext der jeweiligen Indikation und in Bezug auf den betrachteten Endpunkt) der interessierenden Fragestellung entsprechen bzw. ob für den betrachteten Endpunkt langfristig überhaupt ein einheitlicher Therapieeffekt über die Zeit anzunehmen ist, so dass die Ergebnisse sinnvoll interpretierbar sind und zu einer präzisen Beantwortung der Fragestellung beitragen können. Auch methodische Herausforderungen, wie etwa eine im Zusammenhang mit Therapieabbrüchen und Folgetherapien zunehmende Disproportionalität der Hazards, sind zu bedenken. Eine einheitliche Erhebungsdauer für alle Endpunkte in einer Studie ist aus Sicht von BMS in vielen Konstellationen – abhängig von der Indikation und Therapiesituation – weder für die Nutzenbewertung erforderlich noch zielführend. Im Hinblick auf die praktische Durchführbarkeit von Studien wären entsprechende Forderungen nicht angemessen. <u>Vorgeschlagene Änderung:</u> In dem benannten Abschnitt wird eine Konkretisierung des Begriffs „systematisch verkürzt“ sowie ggf. eine genauere Differenzierung

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	des Sachverhalts unterschiedlicher Beobachtungsdauern und verkürzter Beobachtungsdauern angeregt.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

[1] Jönsson B. Ten arguments for a societal perspective in the economic evaluation of medical innovations. Eur J Health Econ. 2009 Oct;10(4):357-9. doi: 10.1007/s10198-009-0173-2. PMID: 19618224.

[2] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Untersuchung der statistischen Eigenschaften von Verfahren zur Übertragbarkeit von Studienergebnissen auf Teilpopulationen: Arbeitspapier; Auftrag GA18-01 [online]. 2018 [Zugriff: 12.01.2023]. URL: https://www.iqwig.de/download/ga18-01_uebertragbarkeit-von-studienergebnissen-auf-teilpopulationen_arbeitspapier_v1-0.pdf.

[3] Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V; Rapid Report [online]. 2020 [Zugriff: 12.01.2023]. URL: https://www.iqwig.de/download/a19-43_versorgungsnaher-daten-zum-zwecke-der-nutzenbewertung_rapid-report_v1-1.pdf.

A.1.10 Bundesverband Medizintechnologie (BVMed) e. V.

Autorinnen und Autoren

- Winkler, Olaf
- Spranger, Anne

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Olaf Winkler
Anne Spranger
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: BVMed e. V.
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

In verschiedenen Abschnitten wird immer wieder die Wichtigkeit von randomisierten prospektiven Studien als „Goldstandard“ berichtet. Es wird auch darüber berichtet wie wichtig dieser Standard ist aber auch wie schwierig es ist diese Studien zu erstellen. Auch die Methodik der RCTs entwickelt sich weiter, wie zum Beispiel RCTs die aus einem Register erstellt werden. Diese neuen Formen von RCTs können bei bestimmten Krankheitsbildern hilfreich sein, wenn die Patientenpopulation schwierig zu randomisieren ist.

Vorschlag Änderung:

Das IQWiG sollte über die verschiedenen Optionen, die es bei RCTs gibt berichten, und welche Voraussetzungen für Register basierte RCTs erforderlich sind damit diese in einer IQWiG Evaluation Berücksichtigung finden können.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung
Abschnitt 3.2.1 (S.55)	<u>Anmerkung:</u> „Die von Dritten auf entsprechende Anfrage übermittelten Informationen werden bez. Ihrer Vollständigkeit bewertet, da eine selektive Datenübermittlung das Ausmaß der Verzerrung erhöhen kann. Hierfür wird das Ausmaß fehlender Informationen und deren Relevanz (z. B. der Aussage zu einzelnen Fragestellungen oder zu bestimmten Endpunkten) abgeschätzt.“ <u>Vorgeschlagene Änderung:</u> Es ist eine genauere Spezifikation wünschenswert, auf Basis welcher Kriterien das IQWiG die Einschätzung der Relevanz und die nachfolgende Beurteilung vornimmt 3.2.1 Umgang mit möglichem Reporting Bias

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Wir begrüßen die Verdeutlichung und Auserarbeitung des Abschnitts zum Reporting Bias. Die vom IQWiG dargestellten möglichen Verzerrung durch Reporting Bias sind verständlich ausgelegt, allerdings wäre es hilfreich, hier die tatsächlich geforderte Informationsdichte noch einmal deutlich zu machen. Zum Beispiel anhand von Checklisten z.B. die von Ihnen erwähnte STROSA unter 5.4.6 (Swart 2015). Was ist der Threshold unter dem das IQWiG eingereichte Studien nicht als Beleg anerkennt. Eine andere Möglichkeit wäre eine Verdeutlichung der geforderten Information auf der IQWiG Homepage. 3.2.4 Patientenberichtete Endpunkte In Ihrer Ausführung zu PROMs führen Sie aus, randomisierte kontrollierte Studientypen am besten geeignet seien. Hier bitten wir zu beachten, dass die Erfassung von PROMs insbesondere im Bezug zu Medizinprodukten international bereits konsentiert wurde. Insbesondere wurde starke Evidenz für generische und konditionsspezifische PROMs Instrumente erprobt, die auf einen vollen Einschluss der entsprechenden PatientInnen abzielt. Wir bitten zu bedenken, dass das IQWiG umfangreiche Vorarbeiten auf internationaler Ebene (1) (PaRIS, aber auch die standardisierte Erfassung der Ergebnisqualität z.B. in England durch PROMs seit 2009) sowie deutscher Versorgungsforschung durch den Innovationsfonds (EPOS, PROMQual, Pro-P, PROMchronic) mit in das Kapitel einbezieht. Ebenso ist die umfängliche Erfassung von PROMs zu verschiedenen Zeitpunkten auch Grundlage für die wenigen Messungen, die bisher im deutschen Gesundheitssystem vorliegen: Nutzung des EPIC-26-Fragebogens Voraussetzung für die Zertifizierung als Pankreaskarzinomzentrum durch die Deutsche Krebsgesellschaft seit 2019, .DeQS-RL des G-BA für Perkutane Koronarinterventionen (PCI) und Koronarangiographie auch über regelmäßige Patientenbefragungen.
Abschnitt 9.3.12	„Auch für PROs werden bei unterschiedlichen Beobachtungszeiten aufgrund unvollständiger Datenerfassung, z. B. durch Abbruch der Beobachtung bei Progression der Erkrankung, zwar teilweise Überlebenszeitanalysen durchgeführt (z. B. Responderanalysen für die Zeit bis zur Verschlechterung). Es fehlen jedoch in diesen Auswertungen die notwendigen Informationen zu PROs nach Behandlungsabbruch oder -wechsel. Für PROs ist in einer solchen Situation durch die systematisch verkürzte Datenerhebung die Interpretierbarkeit stark eingeschränkt. Nicht sinnvoll interpretierbar sind bei sich deutlich zwischen den Behandlungsarmen unterscheidenden Beobachtungsdauern sogenannte „dauerhafte“

Kommentiert [REDACTED] : DOI: 10.1055/s-0034-1396815

Kommentiert [REDACTED] : <https://www.oecd.org/health/health-systems/OECD-PaRIS-hip-knee-data-collection-guidelines-en-web.pdf>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Veränderungen (Verbesserungen bzw. Verschlechterungen). Denn eine „dauerhafte“ Veränderung ist in dem länger beobachteten Behandlungsarm aufgrund der dafür notwendigen ununterbrochenen Über bzw. Unterschreitung des Schwellenwerts (Responsekriterium) potenziell schwerer zu erreichen. Darüber hinaus ist es in einem solchen Fall auch nicht sachgerecht, von einer „dauerhaften“ Veränderung zu sprechen, da die Beobachtungszeit nur einen verkürzten und ggf. sehr kleinen Anteil der gesamten Studiendauer abdeckt. Solche Auswertungen werden im Falle einer verkürzten oder sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauer daher in der Regel nicht für die Nutzenbewertung herangezogen.“ <u>Anmerkung:</u> In zahlreichen Indikationen (z.B. in der Onkologie) ist es bei der Durchführung von RCTs Standard, dass Patienten z.B. bei einer Progression der Erkrankung die Studie beenden und ggf. die Studienintervention erhalten, wenn sie in den Vergleichsarm der Studie randomisiert wurden. Da insbesondere bei verschiedenen nicht-medikamentösen Interventionen eine Verblindung des Patienten nicht möglich ist, wäre eine Studie ohne dieses Angebot in vielen Fällen nicht erfolgreich durchzuführen. Ein adäquater Einschluss von Patienten in den Vergleichsarm ist oft nicht möglich, da die Patienten eine weitere Teilnahme an der Studie ablehnen, wenn sie erfahren, dass sie in den Vergleichsarm randomisiert wurden. In solchen Studien kann eine größere Zahl von Patienten, die im Kontrollarm nach einer Verschlechterung der Erkrankung mit der Studienintervention weiter therapiert und beobachtet werden z.B. zu verbesserten PRO führen, die das Ergebnis im Kontrollarm fälschlich verbessern und damit den Benefit der Studienintervention fälschlich verschlechtern. Eine Zuordnung von Therapieeffekten der nachgeschalteten Studienintervention zur gescheiterten Vergleichsintervention erscheint nicht sachgerecht. Dem ist bei der Nutzenbewertung Rechnung zu tragen, da eine frühere Progression der Erkrankung im Kontrollarm und damit eine unterschiedliche Beobachtungsdauer in den Behandlungsarmen an sich schon für den Vorteil der Studienintervention spricht.
Abschnitt 2.2.2 Auswahl externer Sachverständiger	Im Abschnitt „Gewährleistung der Unabhängigkeit externer Sachverständiger“ und den vorhergehenden Abschnitten wird auf die Notwendigkeit der Unabhängigkeit eingegangen, insbesondere

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	in Bezug auf die pharmazeutische Industrie und der Medizinprodukteindustrie. Dieses ist im Kontext nachvollziehbar. Dennoch ist gerade bei der Einführung neuer Produkte notwendig, dass der Sachverständige sich intensiv mit der Materie auseinandergesetzt hat, eine Produktschulung erhalten hat und erste Fälle behandelt hat. Speziell neue Medizinprodukte haben im Vergleich zur existierenden Standardtherapie einen innovativen Charakter und weichen in der Anwendung teilweise stark von der bisherigen Standardmethode ab. Um die neue Methode beurteilen zu können reichen medizinische Kenntnisse nicht immer aus, um eine neue Methode umfangreich beurteilen zu können und die erforderlichen Parameter für eine Evaluation zu bestimmen. Vorgeschlagene Änderung: Die unabhängigen externen Sachverständigen müssen vor dem Projektbeginn einen Nachweis erbringen das Sie sich mit der neuen Methode auseinandergesetzt haben. Z.B. Durch den Nachweis einer Teilnahmebestätigung an einem Kongress auf der die neue Methode diskutiert wurde, oder durch eine Präsentation einer Ärztin oder Arztes der die Methode schon mehrfach in Deutschland angewendet hat. Mit diesen Kenntnissen können die unabhängigen Sachverständigen immer noch unabhängig beraten, haben aber Kenntnisse über die neue Methode erlangt, die Sie evtl. dazu befähigen die Potentiale der neuen Methode richtig einzuschätzen und die richtigen validierten Tools für deren Erhebung einzufordern. Wichtig ist das dieser Prozess vor der eigentlichen Erstellung des Projektplans berücksichtigt wird, um die neue Methode mit den Wissen von Experten umfangreich evaluieren zu können.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

A.1.11 Clinical Research platform Into molecular testing, treatment and outcome of non-Small cell lung carcinoma Patients (CRISP), Nationales Netzwerk Genomische Medizin, Lungenkrebs (nNGM), Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie (DGHO)

Autorinnen und Autoren

- Griesinger, Frank
- Wolf, Jürgen
- Wörmann, Bernhard

Stellungnahme zur Entwurfsfassung der *Allgemeinen Methoden* (V. 7) des IQWiG

In der am 6. Dezember 2022 veröffentlichten Version 7 seiner *Allgemeinen Methoden* legt das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) unter anderem eine Überarbeitung seiner Methoden zur Nutzenbewertung und zur anwendungsbegleitenden Datenerhebung vor. Für Stakeholder, die auf unterschiedlichen Ebenen an der Generierung, Auswertung und Nutzung von Real World Evidenz (RWE) aus Gesundheitsversorgungsdaten in Deutschland beteiligt sind, bleibt dieser Entwurf in entscheidenden Punkten hinter den Erwartungen an ein umfassendes, patientenzentriertes, für die Versorgungsrealität in Deutschland relevantes Real World Daten (RWD) Framework zurück:

Zwar ist es unbestritten, dass randomisiert kontrollierte Studien (RCT) in den meisten Settings den Goldstandard der Evidenz zum Nachweis der Effektivität von Therapien darstellen, doch das Potential von RWD wird im jüngsten Entwurf der *Allgemeinen Methoden* des IQWiG nicht ausreichend anerkannt, insbesondere vor dem Hintergrund der Limitationen von RCT in spezifischen Kontexten. Letztere betreffen unter anderem die externe Validität, die Selektion von Studienpopulationen, kurze Follow-up-Zeiten und von der tatsächlichen Versorgungssituation abweichende Studienprotokolle. In sich schnell entwickelnden, innovativen Feldern der Medizin besteht zudem bei der Nutzung von Evidenz aus RCTs für regulatorische Entscheidungen die Schwierigkeit, dass die jeweils zum Vergleich herangezogene Therapie zum Zeitpunkt der Nutzenbewertung oft schon nicht mehr dem Standard of Care entspricht.

Qualitativ-hochwertige RWD aus gut geführten Registern bieten die Möglichkeit, Evidenzlücken zu schließen und Patientinnen und Patienten zeitnah Zugang zu innovativen, wirksamen Therapien zu ermöglichen, wenn RCT nicht ethisch oder zeitgerecht durchführbar oder aufgrund sehr kleiner Patientenpopulationen nicht realisierbar sind (NICE 2022). Dabei nimmt der Bedarf für robuste RWE stetig zu (Patel et al. 2021): Denn kleine Patientenkollektive erschweren nicht nur bei seltenen Erkrankungen (orphan diseases) die Durchführung valider RCTs. So ist mit der fortschreitenden Diversifizierung von Behandlungsoptionen sowie der kontinuierlichen Entwicklung weiterer personalisierter Therapie(-sequenzen) auch in der Krebstherapie eine Unterteilung in immer kleinere Patientenstrata zu beobachten, die durch seltene onkogene Treiber definiert sind (Eichler et al. 2020). Vor diesem Hintergrund muss das Potential von RWE im Rahmen von robusten, nicht-randomisierten Studien auf Grundlage von qualitativ-hochwertigen und für die jeweilige Fragestellung und den spezifischen Anwendungskontext relevanten Daten für die regulatorische Bewertung von Therapien genutzt werden. Dabei sollte die Entscheidung über die Eignung von RWD-Quellen jeweils

kontextbezogen nicht nur aufgrund statistisch-methodischer Erwägungen erfolgen, sondern auch klinisch-fachliche Expertise im jeweiligen medizinischen Spezialgebiet sowie eingehende Kenntnisse des spezifischen Versorgungskontexts berücksichtigen (vgl. Niemeyer et al. 2021).

Die Maßstäbe zur Bewertung von Qualität und Eignung von RWD müssen gewisse Trade-Offs in Bezug auf die Detailschärfe klinischer Variablen (z. B. vereinfachte Erfassung von unerwünschten Ereignissen und Dosierungsänderungen), Sample Size und Nähe zum Versorgungskontext gegenüber im Rahmen von RCTs generierten Daten ermöglichen. Dabei sollte ein formaler Prozess der Konsultation zwischen regulatorischen Entscheidungsträgern und den für die Sammlung, Auswertung und Interpretation von RWD Verantwortlichen auf allen Stufen von der Planung bis zur Einreichung einer Studie insbesondere angesichts des sich derzeit ausdifferenzierenden Methodendiskurses zur Nutzung von RWE für regulatorische Zwecke etabliert werden (NICE 2022; FDA 2022). Methoden der Quantitative Bias Analysis zur Identifizierung des bei nicht-randomisierten Studien erhöhten Risikos der Verzerrung aufgrund (unbekannter) Confounder können diese Konsultationen leiten, indem sie zu adressierende Biases benennen und die transparente Darstellung hinnehmbarer Unsicherheiten ermöglichen (Leahy et al. 2022; Popat et al. 2022; Lash et al. 2014).

Methodisch sind die Definition von Standards und die Bewertung von Anwendungskontexten von RWE gegenwärtig einem grundlegenden Wandel unterzogen – von der herkömmlichen Evidenzhierarchie hin zur Anerkennung der Komplementarität unterschiedlicher Formen der Evidenzgenerierung unter Berücksichtigung jeweils spezifischer Anwendungskontexte (Arlett et al. 2021). Aus diesem Grund besteht ein dringender Bedarf für Pilotstudien in Kooperation von IQWiG, Registerverantwortlichen und Datennutzern mit dem Ziel der Erarbeitung eines umfassenden RWE Frameworks von Best Practices für die Generierung valider Evidenz aus RWD-Studien. So kann das Potential von RWD zur Generierung patientenzentrierter, robuster Forschungsergebnisse erschlossen werden, die den zeitnahen Zugang zu innovativen Therapien auch für kleine Subpopulationen ermöglichen und die Anschlussfähigkeit an methodische und medizinische Entwicklungen auf internationalem Niveau sicherstellen (EMA 2020; NICE 2022; FDA 2022)

Literatur

- Eichler, H.-G., Pignatti, F., Schwarzer-Daum, B., Hidalgo-Simon, A., Eichler, I., Arlett, P., Humphreys, A., Vamvakas, S., Brun, N. and Rasi, G. (2021), Randomized Controlled Trials Versus Real World Evidence: Neither Magic Nor Myth. *Clin. Pharmacol. Ther.*, 109: 1212-1218. DOI: 10.1002/cpt.2083.
- EMA – European medicines agencies network strategy to 2025, (2020) <https://www.ema.europa.eu/en/documents/report/european-union-medicines-agencies-network-strategy-2025-protecting-public-health-time-rapid-change_en.pdf>. Accessed 20 February 2023.
- FDA – U.S. Food and Drug Administration, Advancing Real-World Evidence, 19 October 2022. www.fda.gov/drugs/development-resources/advancing-real-world-evidence-program. Accessed 20 February 2023.
- Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol.* 2014 Dec;43(6):1969-85. DOI: 10.1093/ije/dyu149.
- Leahy TP, Kent S, Sammon C, Groenwold RH, Grieve R, Ramagopalan S, Gomes M. Unmeasured confounding in nonrandomized studies: quantitative bias analysis in health technology assessment. *J Comp Eff Res.* 2022 Aug;11(12):851-859. DOI: 10.2217/ce-2022-0029.
- NICE – National Institute for Health and Care Excellence, Real-World Evidence Framework, 23 June 2022, www.nice.org.uk/corporate/ecd9. Accessed 18 January 2023.
- Niemeyer A, Semler S, Veit C, Hoffmann W, van den Berg N, Röhrig R, Gurisch C, Schlünder I, Beckedorf I. Gutachten zur Weiterentwicklung medizinischer Register zur Verbesserung der Dateneinspeisung und -anschlussfähigkeit. Erstellt für das Bundesministerium für Gesundheit (2021). <https://www.bundesgesundheitsministerium.de/service/publikationen/details/gutachten-zur-weiterentwicklung-medizinischer-register-zur-verbesserung-der-dateneinspeisung-und-anchlussfaehigkeit-1.html>. Accessed 22 February 2023.
- Patel D, Grimson F, Mihaylova E, Wagner P, Warren J, van Engen A, Kim J. Use of External Comparators for Health Technology Assessment Submissions Based on Single-Arm Trials. *Value Health.* 2021 Aug;24(8):1118-1125. DOI: 10.1016/j.jval.2021.01.015.
- Popat, S., Liu, S.V., Scheuer, N. et al. Addressing challenges with real-world synthetic control arms to demonstrate the comparative effectiveness of Pralsetinib in non-small cell lung cancer. *Nat Commun* 13, 3500 (2022). DOI: 10.1038/s41467-022-30908-1.

A.1.12 co.value

Autorinnen und Autoren

- Loke, Luran
- Sawade, Linda
- Thaa, Bastian
- Esser, Marc

Autoren: Laura Loke, Dr. Linda Sawade, Dr. Bastian Thaa, Dr. med. Marc Esser

Stellungnahme co.value zum Entwurf der Allgemeinen Methoden Version 7.0

Berlin, 27.02.2023

Einleitung

co.value ist eine Unternehmensberatung mit Schwerpunkt Market Access und langjähriger Erfahrung in der deutschen Versorgungslandschaft. Die Kunden von co.value – pharmazeutische Unternehmen und Hersteller von Medizinprodukten – erhalten Unterstützung in den unterschiedlichsten Belangen des Market Access: Vom Beratungsgespräch beim G-BA bis hin zur Preisverhandlung.

co.value begrüßt die Überarbeitung des Methodenpapiers und ist der Meinung, dass ein fundierter Diskurs mit den stellungnahmeberechtigten Parteien produktive Anregungen hervorbringen kann. Daher möchten wir uns für die Möglichkeit einer Stellungnahme bedanken und im Folgenden zu ausgewählten Aspekten des Methodenpapiers Stellung nehmen.

co.value ist ein Angebot der co.faktor GmbH

Geschäftsführer: Dr. med. Marc Esser

T: [REDACTED] | [REDACTED]@covalue.de | www.covalue.de

Zu Abschnitt 3.3.4, Nutzenbewertung therapeutischer Interventionen: Konzept für anwendungsbegleitende Datenerhebungen gemäß §35a Abs. 3b SGB V (AbD-Konzepte)

Das IQWiG beschreibt auf S. 67 des Methodenpapiers, dass bei der Erstellung eines Konzeptes für eine AbD die wissenschaftlichen Fragestellungen, die sich aus der Nutzenbewertung ergeben, anhand des PICO-Schema geprüft werden sollen. Das IQWiG führt weiter aus, dass es sich bei der Datenerhebung aufgrund der prinzipiellen Fragestellung der Nutzenbewertung in der Regel um eine vergleichende Studie handelt, welche jedoch ohne Randomisierung ausgeführt wird.

Grundsätzlich wird die Generierung versorgungsnaher Daten begrüßt. Es ist in einigen Fällen allerdings fraglich, inwieweit innerhalb einer AbD eine höhere Evidenz generiert werden kann als die im Rahmen der Nutzenbewertung vorgelegten Daten. Doch genau das ist das oberste Ziel einer AbD, wie auch in der Verfahrensordnung des G-BA beschrieben:

„Eine anwendungsbegleitende Datenerhebung dient dazu, die Evidenzgrundlage für die Bewertung des Zusatznutzens des Arzneimittels nach dem 1. Abschnitt dieses Kapitels durch die Erhebung von Daten aus der Versorgung von Patientinnen und Patienten mit dem Arzneimittel, auch im Vergleich zur Versorgung mit anderen Arzneimitteln oder Behandlungsmöglichkeiten, zu verbessern.“¹

Insbesondere gilt dies, wenn im Rahmen der Nutzenbewertung randomisierte kontrollierte Studien durch den pU vorgelegt wurden. In diesem Fall ist es nicht möglich, im Rahmen einer AbD höherwertige Evidenz zu gewinnen: Eine Randomisierung im Rahmen der AbD ist per Gesetz nicht möglich (vgl. § 35a Abs. 3b Satz 6 SGB V), sodass hier nicht von einer Verbesserung der Datenlage im Vergleich zu den in der Nutzenbewertung vorgelegten Daten ausgegangen werden kann.

¹ Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses. Veröffentlicht im Bundesanzeiger Nr. 84a (Beilage) vom 10. Juni 2009; https://www.g-ba.de/downloads/62-492-3055/VerfO_2022-10-20_iK_2023-02-04.pdf

Weiterhin stellt die Notwendigkeit einer vergleichenden Studie insbesondere für Orphan Drugs eine Herausforderung dar, wohingegen bei Orphan Drugs in der Nutzenbewertung kein Vergleich gegenüber einer zweckmäßigen Vergleichstherapie (zVT) gefordert wird. Ungeachtet dessen, dass für Orphan Drugs bis zu einer Umsatzschwelle von 30 Millionen Euro kein Zusatznutzen gegenüber einer zVT gemäß §35 SGB V nachgewiesen werden muss, wurden auch für Orphan Drugs in der Vergangenheit vom G-BA im Rahmen einer AbD zumeist vergleichende Datenerhebungen gefordert (vgl. Fedratinib², Valoctocogen Roxaparvovec³, Autologe Anti-CD19-transduzierte CD3+ Zellen⁴, Risdiplam⁵).

Bei der Konzepterstellung prüft das IQWiG, ob die vorliegende Fragestellung sich aus bereits laufenden Datenerhebungen beantworten lässt oder ob Datenplattformen existieren, welche als primäre Datenquelle für die Beantwortung der Fragen dienen können. Neben studienindividuellen Datenerhebungen können insbesondere (Indikations-)Register eine geeignete Datenquelle für Studien zur Beantwortung der AbD-Fragestellung darstellen.

Dabei weist das IQWiG darauf hin, dass in einer nicht randomisierten Studie aufgrund potenziell unbekannter Confounder aus den in der Studie beobachteten Effekten erst ab einer bestimmten Effektstärke eine verlässliche Aussage zum Nutzen oder Schaden einer Intervention abgeleitet werden kann. Basierend auf den vorliegenden Erkenntnissen, ob innerhalb eines akzeptablen Zeitraums ausreichend Patient*innen rekrutiert und beobachtet werden können, soll entschieden werden, ob eine AbD grundsätzlich ausreichend aussagekräftige Ergebnisse für die Nutzenbewertung generieren kann.

² Gemeinsamer Bundesausschuss. Beschluss über eine Änderung der Arzneimittel-Richtlinie: Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V) Fedratinib (Myelofibrose); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen. 03.11.2022; https://www.g-ba.de/downloads/39-261-5707/2022-11-03_AM-RL-XII_Fedratinib_AbD-010_Forderung_BAnz.pdf

³ Gemeinsamer Bundesausschuss. Beschluss über eine Änderung der Arzneimittel-Richtlinie: (AM-RL) Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V) Valoctocogen Roxaparvovec (schwere Hämophilie A); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen. 02.02.2023; https://www.g-ba.de/downloads/39-261-5866/2023-02-02_AM-RL-XII_Valoctocogen_Roxaparvovec_AbD-002_Forderung.pdf

⁴ Gemeinsamer Bundesausschuss. Beschluss über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Autologe Anti-CD19-transduzierte CD3-positive Zellen (rezidiviertes oder refraktäres Mantelzell-Lymphom); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen. 21.07.2022; https://www.g-ba.de/downloads/39-261-5540/2022-07-21_AM-RL-XII_Autologe-Anti-CD19-transduzierte-CD3-positive-Zellen_AbD-008_Forderung_BAnz.pdf

⁵ Gemeinsamer Bundesausschuss. Beschluss über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Risdiplam (spinale Muskelatrophie); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen. 21.02.2022; https://www.g-ba.de/downloads/39-261-5536/2022-07-21_AM-RL-XII_Risdiplam_AbD-004_Forderung_BAnz.pdf

Die Berücksichtigung dieser Faktoren wird ausdrücklich begrüßt. Fraglich ist jedoch auch hier, inwiefern eine hinreichende Generierung von Daten durch die Erfassung in nichtrandomisierten Registerstudien für Orphan Drugs möglich ist. Insbesondere bei sehr seltenen Erkrankungen stellt die Rekrutierung einer ausreichenden Teilnehmerzahl für die Darstellung einer relevanten Effektstärke eine erhebliche Herausforderung dar. Hinzu kommt, dass es insbesondere bei seltenen Erkrankungen umso schwieriger sein dürfte, alle relevanten Confounder zu identifizieren und in der Datenerhebung zu berücksichtigen. Dies sollte bei der Veranlassung von AbDs für Orphan Drugs grundsätzlich berücksichtigt werden.

Generell möchten wir darauf hinweisen, dass die Realisierbarkeit einer stichhaltigen Datenerhebung innerhalb eines bestehenden Indikationsregisters auch hinsichtlich erhobener Endpunkte und der bestehenden Infrastruktur der Register Herausforderungen für die Umsetzung einer AbD birgt. Häufig sind Register, insbesondere solche für seltene Erkrankungen, nicht auf eine umfassende Datenerhebung ausgelegt, wie sie für die AbD erforderlich ist. Beispielsweise wurde dies auch anhand der Rapid Reports des IQWiG zur Konzepterstellung für eine AbD für Fedratinib⁶ und Zolgensma⁷ ersichtlich. Neben zu geringen Patientenzahlen können in der Praxis häufig nicht alle Daten im Rahmen des Registers so erhoben werden, wie es für die Nutzenbewertung nötig wäre. Aus Dienstleistersicht bemängeln wir in diesem Zusammenhang, dass eine adäquate Aufbereitung der Daten für die Nutzenbewertung gerade in solchen Fällen eine Herausforderung darstellen dürfte.

Bei alledem gilt zu beachten, dass stets die Realisierbarkeit und Angemessenheit bei der Konzepterstellung einer AbD der ausschlaggebende Faktor für die tatsächliche Veranlassung sein sollte. Die Berücksichtigung dieser Faktoren wird in der Verfahrensordnung des G-BA ausdrücklich vorausgesetzt, um eine AbD als erforderlich zu beurteilen und somit veranlassen zu können¹. In bisherigen Konzepten zur Durchführung einer AbD erscheinen die Hürden zumeist unverhältnismäßig im Vergleich zu dem anhand der neuen Daten gewonnenen Nutzen⁸. Wünschenswert wäre daher, dass das IQWiG in seiner Ausarbeitung zum Konzept einer AbD stärker auf erhebliche Hindernisse und Aufwände hinweist, welche es zur Durchführung einer AbD zu bewältigen gilt.

⁶ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Rapid Report A21-142 Anwendungsbegleitende Datenerhebung Fedratinib. Version 1.0 vom 29.06.2022; https://www.iqwig.de/download/a21-142_anwendungsbegleitende-datenerhebung-fedratinib_kurzfassung_rapid-report_v1-0.pdf

⁷ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. IQWiG-Bericht Nr. 977. Konzept für eine anwendungsbegleitende Datenerhebung - Onasemnogen-Abeparvovec. Version 1.0 vom 01.10.2020; https://www.iqwig.de/download/a20-61_anwendungsbegleitende-datenerhebung-onasemnogen-abeparvovec_rapid-report_v1-0.pdf

¹ Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses. Veröffentlicht im Bundesanzeiger Nr. 84a (Beilage) vom 10. Juni 2009; https://www.g-ba.de/downloads/62-492-3055/VerfO_2022-10-20_iK_2023-02-04.pdf

⁸ Verband forschender Arzneimittelhersteller e.V. Anwendungsbegleitende Datenerhebung: eine erste Zwischenbilanz. 2022; <https://www.vfa.de/de/wirtschaft-politik/anwendungsbegleitende-datenerhebung>

Zu Abschnitt 3.5, Nutzenbewertung therapeutischer Interventionen: Diagnostische Interventionen

Wir begrüßen die generelle Überarbeitung und deutliche Erweiterung des Abschnitts 3.5 (Diagnostische Verfahren, S. 159 ff), die dem pharmazeutischen Unternehmer nun spezifischere Vorgaben für die Erstellung des Dossiers und bereits bei der Ausrichtung von klinischen Studien für eine erfolgreiche Interventionsbewertung ermöglicht.

Zu Abschnitt 7, Evidenzbasierte Gesundheitsinformation für Bürgerinnen und Bürger

Im Abschnitt 7.2 (Themenauswahl und Identifizierung der Informationsbedürfnisse) wird auf S. 133 der Wechsel des Instruments zur Qualitätsbewertung von systematischen Übersichtsarbeiten auf AMSTAR 2⁹ beschrieben.

Wir begrüßen diese Entscheidung, da die Entwicklung von AMSTAR 2 auf den Erfahrungswerten aus rund 10 Jahren AMSTAR-Anwendung basiert und somit verspricht, eine lohnenswerte Ablösung des inzwischen in die Jahre gekommenen Oxman-Guyatt-Index mit seiner ursprünglichen Entwicklung im Jahre 1991 zu sein¹⁰.

Neben dem Vorteil der Erweiterung auf 16 Items und der Unterteilung in kritische und nicht-kritische Domänen sind beispielsweise der sinnvolle Umgang mit „not applicable“- und „cannot answer“-Angaben und insbesondere die Möglichkeit der Bewertung von systematischen Übersichtsarbeiten mit nicht randomisierten Studien Gründe, warum AMSTAR 2 derzeit als besonders verlässliches Bewertungsmittel zur Interpretation von systematischen Übersichtsarbeiten zu medizinischen Interventionen in vielen Publikationen beschrieben wird^{11,12}.

Unklar ist, warum diese Aktualisierung des Instruments zur Qualitätsbewertung von systematischen Übersichtsarbeiten nicht auch im Abschnitt für die Informationsbeschaffung umgesetzt wird (siehe Anmerkungen zu **Abschnitts 8.2**).

⁹ Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017 Sep 21;358; <http://dx.doi.org/10.1136/bmj.j4008>.

¹⁰ Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991 Jan 1;44(11):1271-8; 1271-1278. [https://dx.doi.org/10.1016/0895-4356\(91\)90160-b](https://dx.doi.org/10.1016/0895-4356(91)90160-b).

¹¹ Li L, Asemota I, Liu B, Gomez-Valencia J, Lin L, Arif AW et al. AMSTAR 2 appraisal of systematic reviews and meta-analyses in the field of heart failure from high-impact journals. *Syst Rev* 2022 Dec;11(1):1-8; <https://doi.org/10.1186/s13643-022-02029-9>

¹² De Santis KK, Lorenz RC, Lakeberg M, Matthias K. The application of AMSTAR2 in 32 overviews of systematic reviews of interventions for mental and behavioural disorders: A cross-sectional study. *Res Synth Methods* 2022 Jul;13(4):424-33; <https://doi.org/10.1002/jrsm.1532>.

Zu Abschnitt 8, Informationsbeschaffung

Im **Abschnitt 8.1.2** (Suche in Studienregistern) wird auf S. 150 die Ergänzung der neuen EMA-Datenbank „Clinical Trials Information System“ (CTIS) beschrieben, mit dem Hinweis, dass die bisher verwendete EMA-Datenbank „EU-CTR“ ab 2025 nur noch als Archiv dient. Es wird außerdem darauf hingewiesen, dass es eine Übergangsfrist für die Veröffentlichung von Daten zur Registrierung und Veröffentlichung von Studienberichten klinischer Prüfungen im CTIS bis 2025 geben wird. Demnach sollen zur Informationsbeschaffung in Studienregistern zukünftig beide Datenbanken der EMA herangezogen werden.

Wir weisen darauf hin, dass hierfür auch eine rechtzeitige Anpassung der Dossievorlage für Modul 4 nötig sein wird, um dies in der Infobeschaffung entsprechend klarzustellen und zu berücksichtigen. Die Verpflichtung zur Suche in EU-CTR sollte spätestens 2025 entfallen.

Aus der Straffung des **Abschnitts 8.2** (Fokussierte Informationsbeschaffung, S. 157 ff.) unter Wegfall der ursprünglichen Kapitel 8.2.2 bis 8.2.5 (Recherche nach qualitativer Forschung, gesundheitsökonomischen Fragestellungen, und zu Addenda von §-137e- oder §-137-h-Bewertungen) ergibt sich ein offenerer und weniger formalistischer Ansatz für die Darstellung und Bewertung der entsprechenden Fragestellungen, den wir im Allgemeinen begrüßen. Auch wurde die bisherige Redundanz von Punkt 8.2.5 und 8.5 im Methodenpapier 6.1 nun beseitigt. Insgesamt ist der Abschnitt 8.2 somit schlanker und zielführender formuliert.

Unklar ist jedoch, warum bei der Bewertung von systematischen Übersichtsarbeiten weiterhin das inzwischen veraltete Tool AMSTAR verwendet wird, zumal in Abschnitt 7.2 des Methodenpapiers 7.0 auf das neuere Tool AMSTAR 2 verwiesen wird. Hier wäre ein Abgleich zu erwägen.

Im **Abschnitt 8.4** (Recherche nach Leitlinien) sind die Ergänzungen auf S. 156 zu den Leitlinien-Datenbanken, welche bei der Bewertung durch das IQWiG insbesondere berücksichtigt werden, positiv zu nennen, da so dem pharmazeutischen Unternehmer und damit auch dem Dienstleister eine klarere Orientierung bei der Recherche nach Leitlinien verschafft wird.

Allerdings sind beispielsweise nationale Versorgungsleitlinien (NVL) nur über *Leitlinien.de*¹³(Portal des ÄZQs) aktuell und gültig abrufbar (im Vergleich zur AWMF). Daher plädieren wir dafür, dass auch diese Webseite an dieser Stelle explizit mit aufgeführt wird. Hinsichtlich der genannten Datenbank *ECRI Guidelines Trust* gibt es jedoch weiterhin nur wenige Organisationen, die der Veröffentlichung bei ECRI überhaupt zugestimmt haben.

Aus dem Abschnitt geht jedoch weiterhin nicht hervor, welche expliziten Fragestellungen für eine ausschließliche Recherche in MEDLINE qualifiziert wären, da in dieser Datenbank tatsächlich nur ein Bruchteil der verfügbaren Leitlinien überhaupt anzufinden sind.

¹³ <https://www.leitlinien.de/>

Abschnitt 9, Informationsbewertung

Im **Abschnitt 9.3.3** (Beurteilung klinischer Relevanz) wird auf S. 173 weiterhin der mit dem Methodenpapier 6.0 eingeführte Schwellenwert von 15 % des Skalenumfangs als Responsekriterium aufgeführt, damit die entsprechende Responderanalyse für die Bewertung durch das IQWiG berücksichtigt wird. So ist es inzwischen auch in der Dossievorlage des G-BA zur Nutzenbewertung (Modul 4) definiert.

Jedoch benennt der G-BA auf seiner Webseite unter „Antworten auf häufig gestellte Fragen zum Verfahren der Nutzenbewertung“ explizit den Fragebogen EORTC-QLQ-C30 (und die entsprechenden krankheitsspezifischen Ergänzungsmodule) als Ausnahme von dieser Regel:

„Für die EORTC-Fragebögen sind nur Auswertungen zum Responsekriterium 10 Punkte im Dossier darzustellen.“¹⁴

10 Punkte im EORTC QLQ-C30 entsprechen 10 % des Skalenumfangs. Es ist gängige Spruchpraxis des G-BA, bei EORTC-Fragebögen den etablierten Schwellenwert von 10 % als Responsekriterium zu verwenden (z.B. kürzlich im Verfahren zu Idecabtagen vicleucel¹⁵), und auch das IQWiG selbst zieht für die Bewertung dieser Responderanalysen inzwischen eine Responderschwelle von 10 statt 15 Skalenpunkten heran (Trastuzumab deruxtecan¹⁶).

Die EORTC-QLQ-C30-Fragebögen¹⁷ sind ein weltweit verbreitetes Standardinstrument zur Erfassung der Lebensqualität von onkologischen Patient*innen und werden somit häufig in Dossiers für die Nutzenbewertung im Anwendungsgebiet von onkologischen Indikationen dargestellt. Unseres Erachtens nach sollte für den EORTC QLQ-C30 der Schwellenwert von 10 % für die klinische Relevanz von Responderanalysen auch Einzug in die neue Version des Methodenpapiers finden.

¹⁴ <https://www.g-ba.de/themen/arzneimittel/arzneimittel-richtlinie-anlagen/nutzenbewertung-35a/faqs/#wie-soll-vor-dem-hintergrund-des-seit-veroffentlichung-des-methodenpapiers-60-am-5112020-neu-unterbreiteten-vorgehens-des-iqwig-zur-beurteilung-klinischer-relevanzschwellen-bei-komplexen-skalen-in-der-dossiererstellung-mit-der-bestimmung-von-klinischen-relevanzschwellen-bei-komplexen-skalen-umgegangen-werden>

¹⁵ Gemeinsamer Bundesausschuss. Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO. Wirkstoff: Idecabtagen vicleucel. 01.04.2022; <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/781/#nutzenbewertung>

¹⁶ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Trastuzumab-Deruxtecan (Mammakarzinom, nach ≥ 2 Vortherapien) – Addendum zum Projekt A22-81 (Dossierbewertung). Version 1.0; 13.01.2023; https://www.g-ba.de/downloads/92-975-6225/2023-02-02_Addendum-IQWiG_Trastuzumab-Deruxtecan_D-837.pdf

¹⁷ <https://qol.eortc.org/core/>

A.1.13 Daiichi Sankyo Deutschland GmbH

Autorinnen und Autoren

- Bauer, Robert
- Fasan, Annette
- Menzler, Jacob
- Claes, Christa
- Bilgeri, Andrea
- Ratsch, Boris Alexander

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Bauer, Robert; Dr.
Fasan, Annette; Dr.
Menzler, Jacob, Dr.
Claes, Christa, Dr.
Bilgeri, Andrea
Ratsch, Boris Alexander, Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Daiichi Sankyo Deutschland GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>
Daiichi Sankyo Deutschland GmbH (im Folgenden Daiichi Sankyo genannt) ist als forschendes Pharma-Unternehmen im vfa organisiert und teilt die in der Stellungnahme des vfa beschriebenen Positionen zum Entwurf der Allgemeinen Methoden. Die vorliegende Stellungnahme adressiert ergänzend weitere spezifische Aspekte. Konkretisierung in Abschnitt 9.3.3 Im Unterabschnitt „Relevanz bei Skalen“ des Abschnitts 9.3.3 ist aus Sicht von Daiichi Sankyo eine explizite Beschreibung wünschenswert, welche MIDs für die Skalen des SF-36, des EORTC QLQ-C30 sowie den entsprechenden validierten krankheitsspezifischen Ergänzungsmodulen maßgeblich für die Nutzenbewertung sind.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 9.1.1: „Kriterien für den Einschluss von Studien“ (Seite 159)	<u>Anmerkung:</u> In den Entwurf der Allgemeinen Methoden wurde der folgende Satz aufgenommen: „<i>Falls geeignete Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium (Population bzw. Interventionen) vorliegen, muss über die Berücksichtigung der Studie anhand der Stärke der Effektmodifikation und dem Anteil der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen bzw. dem Grad der Abweichung der Interventionen entschieden werden.</i>“ Es bleibt unklar, welche Informationen als „geeignete Informationen über eine Effektmodifikation durch das Einschlusskriterium“ gelten könnten. Des Weiteren bleibt unklar wie die Stärke der Effektmodifikation konkret beurteilt werden soll und auf welche Weise das Institut der Frage nachgehen wird, ob potenzielle Effektmodifikationen klinisch begründet sind.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Zudem besteht aus Sicht von Daiichi Sankyo das Risiko, dass Entscheidungen über die Berücksichtigung von bewertungsrelevanten Studien mittels Effektmodifikation auf Basis von Zufallseffekten in kleinen Subpopulationen getroffen werden könnten. <u>Vorgeschlagene Änderung:</u> Es sollte eine Konkretisierung in Bezug auf die Beurteilung von Effektmodifikationen erfolgen. Zudem muss dringend sichergestellt werden, dass Entscheidungen über die Berücksichtigung von bewertungsrelevanten Studien nicht auf Basis von Zufallseffekten in kleinen Subpopulationen getroffen werden. Auch muss sichergestellt werden, dass die klinische Plausibilität von potenziellen Effektmodifikationen im Nutzenbewertungsverfahren hinreichend geprüft wird.
Abschnitt 9.3.12: „Umgang mit variablen Beobachtungszeiten“ (Seite 191)	<u>Anmerkung:</u> In den Entwurf der Allgemeinen Methoden wurde folgender Satz aufgenommen: <i>„Darüber hinaus können im Falle systematisch verkürzter Beobachtungsdauer die Ergebnisse davon betroffener Endpunkte (z. B. unerwünschte Ereignisse oder PROs) die Gesamtaussage zum Nutzen oder Zusatznutzen einer medizinischen Intervention in der Regel nicht dominieren.“</i> <u>Beobachtungsdauer und Estimands</u> Von systematisch verkürzten Beobachtungszeiten kann im Zusammenhang mit der Anwendung des Treatment Policy Estimands gesprochen werden, wenn zum Beispiel Fragebögen bis zum individuellen Behandlungsende (intercurrent event), nicht aber bis zum generellen Studienende erhoben werden. Es sollte berücksichtigt werden, dass insbesondere bei PROs und unerwünschten Ereignissen auch andere Estimands für die Nutzenbewertung relevante Fragestellungen adressieren können [1]. In diesem Zusammenhang sollten auch die Empfehlungen des SISAQOL-Konsortiums für onkologische Studien berücksichtigt werden, in denen solche Daten als fehlend definiert werden, die für die Analyse einer gegebenen Forschungsfrage oder eines Estimands von Bedeutung sind [2]. <u>Vorgeschlagene Änderung:</u> Im Methodenpapier sollte nicht generell von systematisch verkürzten Beobachtungsdauern ausgegangen werden. Vielmehr

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	sollte die Relevanz unterschiedlicher Estimands für die jeweiligen Behandlungssituationen gewürdigt werden. Die Anwendung nur eines Estimands (i.e. Treatment Policy Estimand) über alle Nutzenbewertungen hinweg ist aus Sicht von Daiichi Sankyo nicht angemessen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. Lawrance et al. (2020): What is an estimand & how does it relate to quantifying the effect of treatment on patient-reported quality of life outcomes in clinical trials?; doi: <https://doi.org/10.1186/s41687-020-00218-5>
2. Coens et al. (2020): International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. Lancet Oncology; doi: 10.1016/S1470-2045(19)30790-9

A.1.14 Deutsche Gesellschaft für Gesundheitsökonomie e. V. (dggö)

Autorinnen und Autoren

- Kifmann, Mathias
- Brosig-Koch, Jeannette
- Herr, Annika
- Wuppermann, Amelie
- Kairies-Schwarz, Nadja
- Liersch, Sebastian
- Leidl, Reiner
- Brettschneider, Christian
- Chernyak, Nadja

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Kifmann, Mathias; Prof. Dr.
Brosig-Koch, Jeannette; Prof. Dr.
Herr, Annika; Prof. Dr.
Wuppermann, Amelie; Prof. Dr.
Kairies-Schwarz, Nadja; Prof. Dr.
Liersch, Sebastian; Dr.
Leidl, Reiner; Prof. Dr.
Brettschneider, Christian; Dr.
Chernyak, Nadja; Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Deutsche Gesellschaft für Gesundheitsökonomie e.V. (dggö)
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>
Die Deutsche Gesellschaft für Gesundheitsökonomie (dggö) beurteilt den neuen Methodenentwurf positiv. Insbesondere begrüßt sie die Überarbeitung der Methoden zur Kosten-Nutzen-Bewertung (KNB) in Kapitel 4 und den damit verbundenen Verzicht auf die Effizienzgrendarstellung.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 1.4 (S.9): Ausführungen zu Gesundheitsökonomie	<u>Anmerkung:</u> In diesem Abschnitt sollten noch Literaturquellen ergänzt werden. <u>Vorgeschlagene Änderung:</u> Ergänzung Quelle [1] als anerkannter internationaler Standard und Quelle [2] zur systematischen Entscheidungsanalyse.
Abschnitt 4.3.2 (S.88): Wahl der Komparatoren	<u>Anmerkung:</u> Der Wahl der Komparatoren kommt in der ökonomischen Evaluation eine entscheidende Bedeutung zu. Im vierten Absatz finden sich folgende Sätze: „Die zweckmäßige Vergleichstherapie stellt eine nach dem allgemeinen Stand der Erkenntnisse zweckmäßige Therapie im Anwendungsgebiet dar. Es soll unter anderem vorzugsweise eine Therapie sein, die sich in der praktischen Anwendung bewährt hat und gegen die nicht das Wirtschaftlichkeitsgebot spricht (5. Kapitel 1. Abschnitt § 6 der Verfahrensordnung des G-BA) [284].“ Das Konzept der zweckmäßigen Vergleichstherapie ist unter dieser Operationalisierung recht unbestimmt und kann relativ leicht Anlass zu vertieften

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Diskussionen im Entscheidungsprozess nach Berichtslegung führen. Die hier aufgeführten Kriterien einer (a) nach dem aktuellen Stand der Erkenntnisse zweckmäßigen Therapie, (b) die sich in der in der praktischen Anwendung bewährt hat und (c) gegen die nicht das Wirtschaftlichkeitsgebot spricht, kann auf Seiten der Verhandlungspartner:innen (regulatorische Instanz und pharmazeutisches Unternehmen) zu entgegengesetzten Interpretationen führen. Aus dem vorangegangenen Abschnitt geht zwar hervor, dass „der Vergleich zwischen dem zu bewertenden Arzneimittel und der für den Nachweis des Zusatznutzens herangezogenen zVT“ erfolgt, jedoch öffnet die Preis/Kosten-Dimension, die in der nachgelagerten KNB relevant wird, Raum für Diskussionen. <u>Vorgeschlagene Änderung:</u> Es sollte eine transparentere Darstellung der Auswahl der Komparatoren erfolgen.
Kapitel 4.3.5 (S. 90): Zeithorizont	<u>Anmerkung:</u> Die Wahl eines festen 5-Jahres Zeithorizonts für längerfristige Modellierungen wirkt pragmatisch, aber kann wie im Text auch angedeutet – etwa in schweren Krankheitsbildern der Onkologie – unnötig lang, bei chronischen Erkrankungen wie Diabetes zu kurz sein. Daher sollten begründete Abweichungen möglich sein. Hier steckt sicherlich die Intention dahinter, die Analysen möglichst zu standardisieren und vergleichbar zu machen. <u>Vorgeschlagene Änderung:</u> Ähnlich wie bei der Diskontierungsrate sollten Variationen im Rahmen von Szenarioanalysen empfohlen werden, wie z.B. 1 Jahr, 5 Jahre, lebenslang.
Kapitel 4.5 (S. 90): Effektivität und weitere klinische Inputvariablen	<u>Anmerkung:</u> Die Aussage „Das Ergebnis der KNB stellt das Ergebnis der Nutzenbewertung nicht infrage“ muss nicht zwangsläufig zutreffen, da die Nutzenbewertung mehr bzw. andere Endpunkte als die ‚KNB‘ betrachten kann, die ‚KNB‘ Aussagen über längere Zeiträume als die Nutzenbewertung

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	machen und andere Methoden zur Erfassung der Unsicherheit verwenden kann. <u>Vorgeschlagene Änderung:</u> Explizite Klärung des Umgangs mit dieser Situation.
Kapitel 4.6 (S. 91 ff.): Nutzwerte	<u>Anmerkung:</u> Bei der Verwendung von QALYs im Rahmen einer Analyse aus der GKV-Perspektive ist zu beachten, dass die Value-Sets zur indirekten Bewertung – etwa der EQ-5D-5L von [3] – typischerweise die Präferenzen einer Repräsentativstichprobe der gesamten (deutschen) Bevölkerung erfassen und nicht die der GKV-Versicherten. Generell müssen verwendete Präferenzen für den Entscheidungskontext relevant sein. Zudem beruhen QALYs in vielen Fällen auf einem Erhebungsbogen zur Lebensqualität, etwa dem EQ-5D oder Derivaten des SF36. Hier sollte klargestellt werden, ob für die Verwendung darauf basierender QALYs eine Validierung der Anwendung des Lebensqualitätsmessinstruments im Indikationsgebiet gefordert wird, wie dies für die Lebensqualitäts- oder Symptommessung im Rahmen der Nutzenbewertung der Fall ist. Schließlich sind bei den indirekten Bewertungsverfahren mit Value-Sets die Antworten zu den Lebensqualitätsdimensionen die Grundlage für die jeweilige Anknüpfung der Populationspräferenzen. Um konsistente Entscheidungen zu unterstützen, ist zu prüfen, ob für QALYs bestimmte Verfahren zur Erhebung des Gesundheitszustands und der Präferenzen verlangt werden, und ob auch für die Präferenzmessung Nachweise psychometrischer Eigenschaften erforderlich sind. Die in der Modellierung übliche (wohl manchmal alternativlose) Verwendung von Nutzwerten aus Studien zu anderen Bevölkerungen und/oder anderen Bewertungsverfahren sollte als Einschränkung gekennzeichnet werden. <u>Vorgeschlagene Änderung:</u> In der Anmerkung genannt: • Sicherstellung, dass verwendete Präferenzen für den Entscheidungskontext relevant sind.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	• Klarstellung, ob eine Validierung des Lebensqualitätsmessinstruments für QALYs im Indikationsgebiet gefordert wird. • Kennzeichnung der Verwendung von Nutzwerten aus Studien zu anderen Bevölkerungen und/oder anderen Bewertungsverfahren als Einschränkung.
Kapitel 4.6 (S. 91): Nutzwerte	<u>Anmerkung:</u> Während EQ-5D, HUI-2 und HUI-3 Instrumente im Rahmen indirekter Messverfahren sind, ist der SF-6D der Name des Scores, der auf Basis der Messung mit dem SF-36 oder SF-12 berechnet werden kann. <u>Vorgeschlagene Änderung:</u> Kennzeichnung des SF-6D als Score.
Kapitel 4.7.2 (S. 95): Kostenbestimmung	<u>Anmerkung:</u> Es stellt sich an dieser Stelle die Frage, wie im Rahmen der Perspektive der GKV-Versichertengemeinschaft mit den Kosten informeller Pflege und Unterstützungsleistungen umzugehen ist. Diese Leistungen werden zwar unentgeltlich erbracht, bergen jedoch teilweise gemäß §37 SGB V das Potential einer Kassenleistung, falls die Unterstützung wegfällt. Eine Intervention, die diesen Leistungsbedarf und somit das Kostenpotential reduziert, hätte einen Vorteil aus genannter Perspektive. <u>Vorgeschlagene Änderung:</u> Kosten informeller Pflege und Unterstützungsleistungen sollten zumindest in Sensitivitätsanalyse berücksichtigt werden.
Kapitel 4.7.2 (S. 95): Kostenbestimmung	<u>Anmerkung:</u> Die Kombination verschiedener Ansätze muss konsistent sein, also etwa Doppelzählungen vermeiden. Zudem sollten keine unterschiedlichen Perspektiven miteinander verrechnet werden (etwa Arzneimittelpreise nach bundesweitem Apothekenabgabepreis und Krankenhauskosten nach individueller Kostenrechnung eines Hauses).

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Vorgeschlagene Änderung:</u> Expliziter Hinweis auf Vermeidung von Doppelzählungen und einheitlicher Perspektive bei der Kostenbestimmung.
Kapitel 4.7.2 (S. 96): Kostenbestimmung	<u>Anmerkung:</u> Die GKV-Perspektive und die volkswirtschaftliche Perspektive sind zwei konzeptionell unterschiedliche Herangehensweisen – eine GKV-Perspektive erfordert somit auch, die von der GKV nach SGB V zu zahlenden Transfers zu berücksichtigen. <u>Vorgeschlagene Änderung:</u> Es sollten bei der Kostenbestimmung auch die von der GKV nach SGB V zu zahlenden Transfers berücksichtigt werden.
Kapitel 4.9 (S. 98) Gesundheitsökonomisches Studiendesign:	<u>Anmerkung:</u> Der erste Absatz dieses Kapitels könnte so verstanden werden, dass allein Modelle zum Standard der ökonomischen Evaluation gehören und klinisch motivierte Piggy-Back Studien als unzuverlässig und zweitklassig gelten. Aus Sicht der dggö sollte sich das zu wählende Studiendesign an der Fragestellung orientieren. Bei Innovationen mit kurz- oder mittelfristigen Konsequenzen kann eine Piggy-Back Studie durchaus zu einer besseren Schätzung der Kosten-Effektivitäts-Relation führen; bei langfristigen Konsequenzen ist eher die modellbasierte Evaluation geeignet. <u>Vorgeschlagene Änderung:</u> Differenzierte Diskussion von Modellen und Piggy-Back-Studien im Kontext der Fragestellung. Sofern es das Ziel ist, dass Modelle aufgrund der Berücksichtigung mehrerer Quellen und eines längeren Zeithorizonts präferiert werden, sollte dies so dargestellt werden.
Kapitel 4.11 (S. 106, Absatz 9) Umgang mit Unsicherheit / Sensitivitätsanalysen	<u>Anmerkung / Vorgeschlagene Änderung:</u> Die Cost-Effectiveness Acceptability Curve (CEAC) sollte nicht nur „ggf.“ berücksichtigt werden, sondern stärker

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	eingefordert werden, da sie wichtige zusätzliche Information liefert.
Kapitel 4.12 (S. 107): Einordnung des ICERs	<u>Anmerkung / Vorgeschlagene Änderung:</u> An dieser Stelle wäre es wünschenswert, wenn ebenfalls explizit auf die Ergebnisse der Cost-Effectiveness Acceptability Curve (CEAC) hingewiesen werden würde. Die Bestimmung des ICERs an sich ist zwar der erste Schritt, aber die CEAC liefert zusätzlich entscheidungsrelevante Information über die Verlässlichkeit des ICERs.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

- [1] Husereau, D., Drummond, M., Augustovski, F., de Bekker-Grob, E., Briggs, A. H., Carswell, C., ... & Staniszewska, S. (2022). Consolidated health economic evaluation reporting standards (CHEERS) 2022 explanation and elaboration: a report of the ISPOR CHEERS II Good Practices Task Force. *Value in Health*, 25(1), 10-31.
- [2] Siebert, U. (2022). Transparente Entscheidungen in Public Health mittels systematischer Entscheidungsanalyse. In: Schwartz FW et al. (Hrsg.) *Public Health. Gesundheit und Gesundheitswesen*, 4. Aufl., Elsevier/Urban & Fischer, München, Kap. 57.2.
- [3] Ludwig, K., Graf von der Schulenburg, J. M., & Greiner, W. (2018). German value set for the EQ-5D-5L. *Pharmacoeconomics*, 36, 663-674

A.1.15 Deutsche Gesellschaft für Innere Medizin e. V. (DGIM)

Autorinnen und Autoren

- Ertl, Georg
- Sauerbruch, Tilman

**Kommentare der DGIM zum Entwurf Allgemeine Methoden IQWiG Version 7.0
vom 06.12.2022**

Seite 7 (Lebenszeit, Beschwerden, Nebenwirkungen, Lebensqualität):

Neben den schon immer vom Institut hervorgehobenen Endpunkten bei der Beurteilung von medizinischen Maßnahmen (Lebensverlängerung, Nebenwirkungen, Besserung der Symptome) ist die Lebensqualität, insbesondere bei chronisch kranken und älteren Patienten ein wichtiger und heute häufig miterfasster Endpunkt. Hier ist allerdings darauf zu achten, dass Lebensqualität nach strikten vorgegebenen Methoden sorgfältig erfasst und statistisch ausgewertet wird. Positive Veränderungen von krankheitsspezifischen Symptomen und die Verminderung von Nebenwirkungen sind dabei unbedingt zu berücksichtigen. Hier bleibt zum Teil offen, wie das IQWiG mit den vielen verschiedenen krankheitsbezogenen Messverfahren zur Erfassung der Lebensqualität und den unterschiedlichen Stadien der Erkrankungen (z.B. akut, chronisch, palliativ) umgeht (Zuordnung der Systeme, statistische Auswertung, Definition von signifikanten Unterschieden etc.).

Seite 10:

In diesem Kapitel wird auf die Kosten-/Nutzen-Bewertung eingegangen. Hier wäre durch einen Dialog IQWiG – G-BA eine größere Transparenz hinsichtlich der Preisverhandlungen zu wünschen, bisher liegen zumindest der DGIM keine entsprechenden Informationen vor.

Seite 11:

Hinsichtlich der verschiedenen „Produkte“ des Instituts erlauben wir uns die Frage, wer welche Rückkopplung zu den verschiedenen Fragen bekommt und – wenn ja - ob diese sich ggf. auch in einer Änderung der Methoden spiegeln. Gibt es hier auch externe Evaluationen?

Seite 12:

Bei der Potenzialbewertung gem. § 132e ist eine Anhörung nicht vorgesehen. Die DGIM betont noch einmal, dass sie es für gerechtfertigt und notwendig hält, hier eine Anhörung durch wissenschaftliche Fachgesellschaften mitaufzunehmen. Dies betrifft auch § 137h SGB V.

Seite 14 (Tabelle 1; Arbeitspapier):

Laut IQWiG handelt es sich um *Information zu versorgungsrelevanten Entwicklungen in der Medizin oder zu methodischen Aspekten*: Soweit wir hier korrekt recherchiert haben, hat es in den vergangenen 4 Jahren nur 3 fertiggestellte Papiere gegeben. Kann das Institut ermessen, zu welchen Konsequenzen sein Beitrag GA16-02 (Kommunikation von Unsicherheit im Gesundheitswesen) geführt hat?

Seite 14 (Tabelle 1, Potenzialbewertung und Bewertung gemäß § 137h SGB V):

Soweit das uns ersichtlich ist, sind seit 2017 nur 17 Berichte erstellt worden. Das erscheint uns wenig. Liegt hier ein methodisches Problem vor?

Seite 18, Rapid Reports:

Rapid Reports werden vorrangig mit dem Ziel einer zeitnahen Information zu relevanten Entwicklungen im Gesundheitswesen (z. B. neue Technologien, Publikation wichtiger Meilensteinstudien) erstellt, heißt es in Ihrem Bericht. Von den ca. 20 (?) Rapid Reports der vergangenen 4 Jahre befassten sich nahezu die Hälfte mit dem Zusammenhang zwischen Leistungsmenge nicht medikamentöser Interventionen und Qualität. Hier heißt es in ihrem Papier, dass der Entwurf vor der Fertigstellung optional externen Gutachtern vorgelegt werden könne. Da es sich um eine sehr wichtige Fragestellung handelt, möchten wir sie fragen, wie häufig war das hier der Fall und warum nur optional?

Seite 21: AbD-Konzept:

Hier werden externe Sachverständige für medizinische Expertise sowie betroffene Patientenorganisationen eingebunden. Nach wie vor ist es uns nicht klar, wie repräsentativ diese Gruppen sind und wie belastbar die Recherche ist.

S. 22: Kosten-Nutzen-Bewertung gemäß § 35b SGB V:

Hier haben wir keinen Bericht innerhalb der letzten 4 Jahre gefunden. Kann das sein? Siehe auch unseren Kommentar zur Seite 10!

Seite 26 (§ 137h):

Es wird dargelegt, dass das Institut optional zur Unterstützung der Bewertung eine eigenen Recherche durchführen wird. Was heißt optional? Welches sind die Kriterien?

Seite 30 Arbeitspapiere:

Siehe Seite 14

Seite 37:

Beauftragung externer Sachverständige für die Dossier-Bewertung (§ 35b SGB V):

Konnten inzwischen durchgehend Sachverständige rekrutiert werden? Gibt es eine transparente Liste dieser Sachverständigen, die einsehbar ist?

Seite 39 (Review der Produkte des Institutes):

Die Auswahl erfolgt durch das Institut, d. h. derjenige der den Text erstellt, bestimmt auch die Reviewer. Dies entspricht eigentlich keinem unabhängigen Review-Prozess. Sind die Reviews einsehbar?

Seite 42 unten (Patientenrelevanter Nutzen):

Hier geht es darum, wie ein Patient sich fühlt, seine Funktionen und Aktivitäten wahrnehmen kann etc.: Vielleicht nochmal ganz deutlich machen, was bei der Erfassung berücksichtigt werden muss.

Seite 43 (Patientenrelevante Zielgröße 2):

Inwieweit sind Beschwerden und Komplikationen a priori definiert?

Seite 43 unten (Patientenrelevante Endpunkte, Einbeziehung Betroffener):

Auch hier handelt es sich ja um Kriterien, die einerseits generell, andererseits krankheitsspezifisch definiert werden müssen. Kann das Institut hierzu valide Aussagen machen, oder handelt es sich um eine Pseudo-Genauigkeit?

Seite 44 (Surrogat-Endpunkte):

Hier sollten unbedingt die wissenschaftlichen Fachgesellschaften an einer definierten Vorgehensweise beteiligt werden. Dies gilt auch für die Definition des patientenrelevanten Nutzens (Seite 45 oben).

Seite 47 (Unerwartet auftretende, unerwünschte Ereignisse):

Hier sollte man unbedingt auf die verschiedenen vorhandenen Register zurückgreifen und auch im Dialog mit den Fachgesellschaften die Qualität dieser Register besprechen (teilweise wird das später auch in ihrem Methodenreport adressiert).

Seite 55 (Dramatischer Effekt):

Möglicherweise ist es sinnvoll, wenn das Institut in seinem Methodenreport einmal versucht zusammenzustellen, in welchen Bereichen der Medizin dramatische Therapieerfolge zu erwarten sind (Wie werden chirurgische und individuelle Eingriffe hier eingeordnet?).

Seite 56 (Mindeststudiendauer):

Auch hier wäre der Dialog mit den jeweils wissenschaftlichen Fachgesellschaften zur Definition oder den Anforderungen an die Mindeststudiendauer sehr entscheidend.

Seite 59 (Anwendung von Arzneimitteln außerhalb der in der Fachinformation beschriebenen Zulassung):

Drug Repurposing spielt in der Medizin eine zunehmende Rolle (siehe Behandlungsansätze SARS-CoV-2!). Auch hier müsste sich das IQWiG in Zukunft einmal Gedanken machen, wie man die Wirksamkeit solcher Medikamente erfasst bzw. ob die Frage für das Institut relevant ist.

Seite 59 (Nicht-Anwendbarkeit von Studienergebnissen, wenn die Altersgruppen außerhalb der Zulassung):

Dies ist in praxi möglicherweise ein Problem, da in einer Zulassungsstudie natürlich eine begrenzte Altersspannweite umfasst wird, in praxi aber bei zunehmender Veränderung der Altersstruktur solche Patienten, auch wenn sie älter sind (ggf. auch Kinder, siehe Hepatitis C) mit den entsprechenden Medikamenten behandelt werden. Diese berührt die Problematik von Pharmakokinetik und Dynamik in verschiedenen Populationen (Geschlecht, Alter, Schwere der Krankheit etc.) Darauf wird im Methoden-Report wenig (gar nicht?) eingegangen.

Seite 66 (anwendungsbegleitende Datenerhebung):

Auch hier sollte man natürlich auf Register zurückgreifen. Aber auch das wird in den nächsten Seiten gestreift. Wichtig wäre es auch Register zu gleichen Krankheiten zusammenzuführen, wie oben schon beschrieben und im Dialog mit den wissenschaftlichen Fachgesellschaften (auch BMG und GBA) die Vorgaben für die Register zu besprechen.

Seite 69 (Nicht-medikamentöse therapeutische Interventionen):

Hinsichtlich der Rapid Reports in diesem Themenkomplex verweisen wir auf unseren Kommentar auf S. 18

Seite 72 (Wirkungen und Nebenwirkungen diagnostischer Interventionen):

Es gibt eine Reihe von diagnostischen Interventionen, die auch mit einem Schaden einher gehen können (Strahlenbelastung, Biopsie, Endoskopie, Sedierung bei diagnostischen Eingriffen). Hier sind wahrscheinlich völlig unterschiedliche Methoden zu der Erfassung notwendig.

Seite 74 (Anreicherungsdesign):

Sind diese verschiedenen Studiendesigns alle ethisch vertretbar?

Seite 79 (Überdiagnose):

Wir verweisen noch einmal darauf, wenn auch dieser Begriff häufig benutzt wird, dass er in sich nicht logisch ist. Eine Diagnose ist eine Diagnose, man kann nicht von Überdiagnose sprechen, allenfalls von einer in dieser Situation irrelevanten Diagnose, zu häufigen und häufig falschen Diagnose oder Überdiagnostik im Sinne von zu viel Diagnostik.

Seite 80 (Präventionsprogramme, die von Bevölkerungsgruppen mit erhöhtem Krankheitsrisiko weniger in Anspruch genommen werden, diese Aspekte stehen besonders im Fokus von Bewertungen des Institutes):

Auf welche Bewertungen bezieht sich das Institut hier konkret?

Seite 99 (Gesundheitsökonomische Modelle für gesundheitsökonomische Evaluation):

Ist eine *Reduktion der Komplexität* heutzutage noch der Stand in Anbetracht der enormen Datenmengen die vorliegen, zugänglich sind und heutzutage auch über die Datenverarbeitung analysiert werden können?

Seite 99 (Glaubwürde und weniger glaubwürdige Modelle):

Was heißt das konkret?

Seite 110 (Ergänzende Befragung von Expertinnen und Experten zur Kostenbestimmung):

Welche Experten sind hier gemeint? Wie werden sie konkret ausgewählt, wie häufig findet dies statt?

Seite 117 (Mindestmengen):

Hier sollten unbedingt die Fachgesellschaften in die Begutachtung mit involviert werden (siehe auch oben).

Seite 118 (Versorgungsdaten):

Analyse von Versorgungsdaten ist u.E. sehr wichtig. Methodik wird hier sehr knapp abgehandelt. Inwieweit können diese sektorübergreifend erfasst und analysiert werden? Wurde das Institut während der Corona-Pandemie hier eingebunden?

Seite 120 (Datenquelle, Tabelle 9):

Kommt das Institut (z.B. indirekt über das BMG) nicht an die Routinedaten der gesetzlichen Krankenkassen oder auch der Kassenärztlichen Vereinigung? Dies sollte im Methodenreport kurz ausgeführt werden. Wenn nicht, ergäbe sich heraus möglicherweise das Risiko eines Selektions-Bias.

Seite 123 (Themenvorschläge für ...):

Wie ist hier das Verhältnis von Vorschlägen zu Auswahl? Das wäre einmal interessant zu wissen, ggf. auch darzulegen.

Seite 137 (Gesundheitsinformation, Beteiligung von Betroffenen):

Hier auch wieder die Frage, wie repräsentativ ist die Auswahl der Personen etc. (möglicherweise für die gleichen Krankheiten immer dieselben?)

Seite 172 (Responder-Analysen):

Hier verweisen wir auf unseren Kommentar vom vergangenen Jahr (15% Schwelle). Hat es hier inzwischen Rücksprache mit den Fachgesellschaften gegeben? Gibt es andere Lösungsvorschläge?

Seite 201 (Tabelle 12, gesundheitsbezogene Lebensqualität):

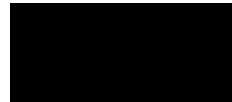
Hier auch wieder die Frage, ob wirklich die Methodik bestehen und diese zu erfassen (unterschiedliche Krankheiten etc.).

22.02.2023

Deutsche Gesellschaft für Innere Medizin e.V.



Prof. Dr. Tilman Sauerbruch



Prof. Dr. Georg Ertl

A.1.16 Deutsche Gesellschaft für Kardiologie – Herz- und Kreislaufforschung e. V. (DGK)

Von: Leistungsbewertung [REDACTED]
Gesendet: Dienstag, 28. Februar 2023 15:27
An: Methoden
Cc: AWMF | Geschäftsstelle; [REDACTED]; [REDACTED]
Betreff: Unterstützung der GMDS+IBS-DR Stellungnahme | IQWiG - Allgemeine Methoden 7.0: Entwurf zur Stellungnahme - Anlage | DGK V2023_015
Anlagen: 20230122_Stellungnahme_IQWiG_Version7.0_IBS-DR_GMDS.pdf; Allgemeine Methoden - Entwurf für Version 7.0.pdf
Priorität: Hoch

Sehr geehrte Damen und Herren,

wir unterstützen die anhängige Stellungnahme der GMDS und der IBS-DR seitens der DGK vollumfänglich! Die Methoden sind hinsichtlich der Medizinischen Fachrichtungen global formuliert, dennoch stehen wir gerne zur Verfügung, sollten Detailfragen zum kardiologischen Fach aufkommen. Welche Beispiele der Reports ggfs. herangezogen werden, sind ja im Vorfeld nicht abzusehen. Würde eine Anhörung geplant, würden wir unseren Experten, Herrn Professor Tim Friede anmelden wollen.

Bitte lassen Sie uns die entsprechenden Anforderungen zukommen, sobald ein Termin bekannt werden sollte.

Für Rückfragen jedweder Art stehe ich Ihnen gerne zur Verfügung
und sende beste Grüße

i. A. [REDACTED]
Ausschüsse Qualität- / Leistungsbewertung | Projektgruppen
Gesundheitspolitik

Deutsche Gesellschaft für Kardiologie
- Herz- und Kreislaufforschung e.V.
German Cardiac Society
Grafenberger Allee 100, 40237 Düsseldorf
Tel.: +49 211 [REDACTED]
Fax: +49 211 6 [REDACTED]
E-Mail: leistungsbewertung@dgk.org

dgk.org
facebook.com/DGK.org
twitter.com/DGK_org

Vorstand: Prof. Dr. Stephan Baldus (Präsident) - Prof. Dr. Andreas M. Zeiher - Prof. Dr. Holger Thiele
Geschäftsführer und besonderer Vertreter nach § 30 BGB: Dr. rer. med. Konstantinos Papoutsis
Sitz: Bad Nauheim - Eingetragen beim AG Friedberg unter VR [REDACTED]

A.1.17 Deutsche Gesellschaft für Orthopädie und Unfallchirurgie (DGOU), Deutsche Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC), Deutsche Gesellschaft für Unfallchirurgie (DGU)

Autorinnen und Autoren

- Kladny, Bernd
- Pennig, Dietmar

GESCHÄFTSSTELLE

DGOU- / DGOOC- / DGU-Geschäftsstelle: [REDACTED]

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

Per eMail: methoden@iqwig.de

DGOU e. V. / DGOOC e. V. / DGU e. V.

Tel.: [REDACTED]
Fax: [REDACTED]
www.dgou.de

Berlin, 28.02.2023

**Gemeinsame Stellungnahme
der Deutschen Gesellschaft für Orthopädie und Unfallchirurgie (DGOU),
der Deutschen Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC) und
der Deutschen Gesellschaft für Unfallchirurgie (DGU)
zum Entwurf der Allgemeinen Methoden, Version 7.0**

Sehr geehrte Frau [REDACTED],

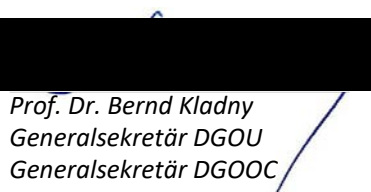
die Deutsche Gesellschaft für Orthopädie und Unfallchirurgie (DGOU), die Deutsche Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC) und die Deutsche Gesellschaft für Unfallchirurgie nehmen Bezug auf die Anfrage des IQWiG zum Entwurf der Allgemeinen Methoden, Version 7.0 vom 26.01.2023.

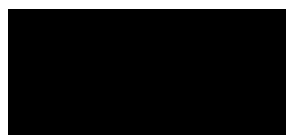
Hierzu verweisen wir auf unsere Stellungnahme zur Version 5.0 aus dem Jahre 2017, welche wir diesem Schreiben nochmals beifügen. Die hier enthaltenen Äußerungen möchten wir abermals bekräftigen.

Gestatten Sie uns an dieser Stelle die Bemerkung, dass wir seit fast zwei Dekaden eine eigene Arbeitsgemeinschaft Evidenzbasierte Medizin implementiert haben und aufgrund unserer Expertise bei Themen mit Bezug zu Orthopädie und Unfallchirurgie ein grundsätzliches Sprachrecht erwarten können.

Für Ihre Fragen und Vorschläge stehen wir jederzeit gern zur Verfügung.

Mit freundlichen Grüßen


Prof. Dr. Bernd Kladny
Generalsekretär DGOU
Generalsekretär DGOOC


Prof. Dr. Dietmar Pennig
Stellv. Generalsekretär DGOU
Generalsekretär DGU

ANLAGE

Vorstand DGOU (gemäß §26 BGB Abs. 1)

Präsident: Prof. Dr. Maximilian Rudert, Stellvertretender Präsident: Prof. Dr. Steffen Ruchholtz
Generalsekretär: Prof. Dr. Bernd Kladny, Stellvertretender Generalsekretär: Prof. Dr. Dietmar Pennig

DGOU-Bankverbindung: [REDACTED], IBAN: [REDACTED], SWIFT-BIC: [REDACTED]

DGOU-Steuer-Nr. [REDACTED], Amtsgericht [REDACTED], VR [REDACTED]



PRÄSIDENT 2017

Prof. Dr. med. [REDACTED]

DGOU Geschäftsstelle [REDACTED]

Herrn [REDACTED]

AWMF-Geschäftsstelle
[REDACTED]

Per E-Mail: [REDACTED]

DGOU Geschäftsstelle

[REDACTED] – [REDACTED]

[REDACTED]

Tel. [REDACTED] [REDACTED]

Fax [REDACTED] [REDACTED]

[REDACTED]

www.dgou.de

Berlin, 23.01.2017

Stellungnahme zum Entwurf „Allgemeine Methoden“, Version 5.0

Sehr geehrter Herr Makoschey,

zum Entwurf „Allgemeine Methoden“, Version 5.0 des IQWiG bittet die Deutsche Gesellschaft für Orthopädie und Orthopädische Chirurgie (DGOOC), sowie die Deutsche Gesellschaft für Unfallchirurgie (DGU) und die Deutsche Gesellschaft für Orthopädie und Unfallchirurgie (DGOU) nach Beratung mit Prof. Dr. med. Tobias Renkawitz, Leiter des DGOU-Gremiums „Evidenzbasierte Medizin“ (EbM-Gruppe), um Einarbeitung der folgenden Punkte in die Stellungnahme der AWMF:

1. Die im politischen Entscheidungsprozess oftmals so wichtigen „externen Sachverständigen“ will das IQWiG in einer „Sachverständigendatenbank“ führen. Darin können sich „*Interessentinnen und Interessenten unter Angabe ihres Fachgebiets und ihrer fachlichen Expertise*“ registrieren lassen. „*Für die zu vergebenden Projekte wird anhand einer Kriterienliste die oder der jeweils am besten geeignete Bewerberin oder Bewerber des entsprechenden Fachgebiets aus der Sachverständigendatenbank ausgewählt und beauftragt.*“ Auf Grundlage der EbM-Kompetenz der Deutschen Gesellschaft für Orthopädie und Unfallchirurgie (DGOU) wäre die Einräumung eines Vorschlagsrechtes für die DGOU folgerichtig. Innerhalb und außerhalb der „DGOU-EbM-Gruppe“ gibt es für die unterschiedlichsten Bereiche unseres Faches kompetente Ansprechpartner.
2. Ebenso (siehe auch Pkt. 1) wäre es ebenso folgerichtig, auf Grundlage der EbM-Kompetenz der Deutschen Gesellschaft für Orthopädie und Unfallchirurgie (DGOU), auch für die „externen reviewer“ („2.2.4 Review der Produkte des Instituts“- gemeint sind Berichte, Dossierbewertungen etc) ein Vorschlagsrecht für die DGOU einzuräumen. „*Die Identifikation externer Reviewer kann durch eine entsprechende Recherche, durch die Kenntnis der Projektgruppe, durch das Ansprechen von Fachgesellschaften, durch eine Bewerbung im Rahmen der Ausschreibung für die Auftragsbearbeitung usw. erfolgen.*“

Mit freundlichen Grüßen

[REDACTED]
Prof. Dr. med. Ingo Marzi
Präsident der DGOU
Präsident der DGU

[REDACTED]
Prof. Dr. med. Andrea Meurer
Stellv. Präsidentin der DGOU
Präsidentin der DGOOC

A.1.18 Deutsche Gesellschaft für Pharmazeutische Medizin e. V. (DGPharmMed)

Autorinnen und Autoren

- Bamberger, Manuela
- Breitkopf, Simone
- Brandewiede, Bernard

Stellungnahme

Entwurf der Allgemeinen Methoden Version 7.0 des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Die Deutsche Gesellschaft für Pharmazeutische Medizin (DGPharMed) begrüßt die Weiterentwicklung der Allgemeinen Methoden des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), vor allem mit Blick auf den wissenschaftlichen Fortschritt und die sich ändernden Bedingungen im medizinischen Alltag.

Gleichzeitig möchte die DGPharMed darauf hinweisen, dass in der Ergänzung des Abschnitts 3.3.4 zur Beschreibung der Inhalte von Konzepten für anwendungsbegleitende Datenerhebungen gemäß §35a Abs. 3b SGB V (AbD) die maßgebliche Beachtung des klinischen Alltags und wesentlicher Grundprinzipien klinischer Forschung noch unzureichend dargestellt und eingefordert ist.

Als medizinische Fachgesellschaft und Mitglied der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) nimmt die DGPharMed daher zu folgenden Abschnitten des Entwurfs der Allgemeinen Methoden Version 7.0 Stellung:

2.1.3.2. AbD-Konzept (S.21)

Anmerkung:

AbD als Sonderform nichtinterventioneller Studien (NIS) sind komplexe Projekte, deren Planung und Umsetzung fundiertes Wissen und insbesondere Praxiserfahrung erfordern. Aus Sicht der DGPharMed ist die Einbindung von erfahrenen ExpertInnen aus dem Bereich der klinischen Forschung, insbesondere aus den entsprechenden methodischen Fachgesellschaften der AWMF, daher dringend anzuraten.

Vorgeschlagene Änderung:

Bei der Erarbeitung des AbD-Konzepts werden externe Sachverständige für die medizinische Expertise **und Fachleute mit Expertise in klinischer Forschung** sowie Betroffene bzw. Patientenorganisationen für die Patientenperspektive eingebunden.



3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte)

Anmerkung:

Mithilfe des AbD-Konzepts soll die prinzipielle Realisierbarkeit der AbD geprüft werden. Aus Sicht der DGPharMed muss die Prüfung der Realisierbarkeit ebenfalls eine kritische Auseinandersetzung mit der Frage beinhalten, ob die AbD im konkreten Einzelfall aussagekräftigere Daten generieren kann als alternative Studienformen. Beispielsweise kann das ärztliche Ordnungsverhalten zu einem erheblichen Selektionsbias führen, wenn mit dem zu bewertenden Arzneimittel und der zweckmäßigen Vergleichstherapie (ZVT) unterschiedliche Patientenkollektive adressiert werden. Auch bei fehlenden realistischen Therapiealternativen kann ein fairer Vergleich zwischen den Interventionen völlig unmöglich werden. Über alternative Studiendesigns, beispielsweise indirekte Vergleiche auf Basis historischer Kontrollen, kann im Einzelfall ggf. wesentlich aussagekräftigere Evidenz generiert werden.

An dieser Stelle sei ausdrücklich klargestellt, dass die Durchführung der AbD als randomisierte Studie ethisch fragwürdig und aufgrund nationaler und internationaler rechtlicher Vorgaben unzulässig ist. Darüber hinaus wären bei einer Randomisierung die Vorgaben der §§40-42 AMG anwendbar, die im Versorgungsalltag nicht erfüllbar sind.

Der Aspekt der Realisierbarkeit beinhaltet weiterhin Fallzahlberechnungen, für die die Einbindung von Expertise im Bereich des praktischen Studienmanagements dringend anzuraten ist. Andernfalls besteht die Gefahr formal-methodisch plausibel zu scheinende Annahmen zu treffen, die unter Realbedingungen jedoch nicht umsetzbar sind.

Die Erfahrung zeigt außerdem, dass der Zeit- und Ressourcenbedarf zur Anpassung der Registerstruktur häufig unterschätzt wird. In Hinblick auf die Erfassung patientenberichteter Daten zu Symptomatik und Lebensqualität stellt sich weniger die Frage nach der Anpassung identifizierter Indikationsregister, sondern die Frage nach der Rechtsgrundlage. Zur Verwendung von Fragebögen im konkreten Einzelfall sei auf die gemeinsamen Empfehlungen des Bundesinstituts für Arzneimittel und Medizinprodukte (BfArM) und des Paul-Ehrlich-Instituts (PEI) zur Anzeige von Anwendungsbeobachtungen und nichtinterventionellen Studien [1] sowie auf Artikel 2 (2) (3) der Verordnung EU No. 536/2014 (Clinical Trials Regulation) verwiesen.

Im Methodenpapier fehlt aktuell außerdem eine kritische Prüfung der Eignung der IT-Infrastruktur bestehender Indikationsregister. Die DGPharMed weist darauf hin, dass manche der derzeit von Registern zur Datenhaltung genutzten IT-Lösungen, beispielsweise das Tabellenkalkulationsprogramm Microsoft Excel®, für klinische Forschungsprojekte wie AbD vollkommen ungeeignet sind. Zur Erhöhung der Datenqualität stehen neben dem bisher bemühten Konzept der stichprobenartigen Source Data Verification mit datengetriebenen Maßnahmen nach dem Vorbild des Risk Based Quality Management innovativere Möglichkeiten zur Verfügung.



Vorgeschlagene Änderung:

In die Erstellung der AbD-Konzepte sollte zukünftig Sachverstand aus den Bereichen klinische Entwicklung, Studienmanagement, Datenmanagement und IT-Lösungen in der klinischen Forschung einfließen.



Dr. med. univ. Manuela Bamberger
Sprecherin des Fachbereichs Klinische Prüfung



Dr. med. Simone Breitkopf
Sprecherin des Fachbereichs Market Access
Beisitzerin, Ressort Nationale & Internationale Kooperationen / Mitgliedschaften



Bernard Brandewiede
Beisitzer, Ressort: Fachbereiche

[1] Gemeinsame Empfehlungen des Bundesinstituts für Arzneimittel und Medizinprodukte und des Paul-Ehrlich-Instituts zur Anzeige von Anwendungsbeobachtungen nach § 67 Absatz 6 Arzneimittelgesetz und zur Anzeige von nichtinterventionellen Unbedenklichkeitsstudien nach § 63f Arzneimittelgesetz vom 15. Dezember 2022.

URL: <https://www.bfarm.de/SharedDocs/Downloads/DE/Arzneimittel/Zulassung/klin-pr/nichtInterventStud/Gemeinsame-Empfehlungen-zu-AWB-und-PASS.pdf>

A.1.19 Deutsche Krankenhausgesellschaft e. V. (DKG)

Autorinnen und Autoren

- Brenske, Michael

Institut für Qualität und
Wirtschaftlichkeit im Gesundheitswesen
-Stellungnahme zu den „Allgemeinen
Methoden Version 7.0“ -
Im Mediapark 8 (Köln Turm)
50670 Köln

Stellungnahme der Deutschen Krankenhausgesellschaft zum Entwurf „Allgemeine Methoden Version 7.0“ des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) vom 06.12.2022

Sehr geehrter Herr Professor Windeler,
sehr geehrte Damen und Herren,

mit dem Entwurf der „Allgemeinen Methoden Version 7.0“ legen Sie erneut eine umfangreiche Darstellung Ihrer Arbeitsweisen vor. Die damit geschaffene Nachvollziehbarkeit und Transparenz bei der Erstellung Ihrer unterschiedlichen Produkte begrüßen wir sehr. Dies trifft gleichermaßen auf den Umstand zu, dass die „Allgemeinen Methoden“ damit einer regelmäßigen Überarbeitung und Weiterentwicklung unterliegen und die Möglichkeit besteht, über ein Stellungnahmeverfahren in einen inhaltlichen Austausch zu treten.

Gern möchten wir daher die Gelegenheit nutzen, Ihnen Anregungen für eine Überarbeitung des vorgelegten Entwurfes zu übermitteln. Angesichts des mittlerweile erreichten Umfangs des Methodenpapiers war es jedoch nicht möglich, sämtliche Darstellungen einer ausführlichen Prüfung unterziehen zu können. Somit ist nicht ausgeschlossen, dass sich neben den nachfolgend aufgeführten Aspekten noch weiterer Bedarf für Anmerkungen ergeben kann. Insofern ist unsere Stellungnahme als nicht abschließend zu betrachten und wir behalten uns vor, ggf. weitere Aspekte in zukünftige Diskussionsprozesse einzubringen. Im Einzelnen möchten wir an dieser Stelle auf folgende Punkte eingehen:

Umbenennung des Produkts „HTA-Bericht“ in „ThemenCheck-Bericht“

In Version 7.0 werden HTA-Berichte im Verfahren des ThemenChecks Medizin nunmehr ThemenCheck-Berichte genannt. Dies ist insofern schade, als dass der Begriff Health Technology Assessment (HTA) im Zuge nationaler als auch internationaler Standards

und Diskussionen eine global verständliche Verankerung hat.^{1,2,3,4} Mit dem Begriff ThemenCheck-Bericht wird hingegen ein spezifischer Sonderweg beschritten, aus dem methodisch nicht sofort erkenntlich wird, was damit gemeint ist. So stellt sich die Frage, warum diese Änderung im Entwurf für Version 7.0 erfolgt ist.

Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) in Kapitel 3.3.4

Der Abschnitt beschreibt kurz die im AbD-Konzept erarbeiteten Themen. Ausführlichere Darlegungen finden sich im Rapid Report A19-43 des IQWiG.⁵ Unter anderem wird dargelegt, dass aus einer nicht randomisierten Studie „aufgrund potenziell unbekannter Confounder aus den beobachteten Effekten erst ab einer bestimmten Effektstärke eine Aussage zum Nutzen oder Schaden einer Intervention abgeleitet werden“ kann. Dies wird z. B. in einem Bereich von 2 bis 5 für das relative Risiko (bzw. 0,2 bis 0,5 bei mortalitäts-senkenden Interventionen) bezogen auf die untere bzw. obere Konfidenzintervallgrenze dargestellt. Im Rapid Report A19-43 des IQWiG heißt es dazu, dass sich die konkrete Schwelle im Einzelfall durch die Qualität der Daten ergibt.⁵ Es ist daher aus unserer Sicht erforderlich, ggf. im Rahmen einer den Rapid Report A19-43 ergänzenden Publikation festzulegen und zu erläutern, wie die Qualität der beispielsweise im Rahmen eines Propensity Score Matching erreichten Balanciertheit der Confounder in der Studienpopulation beurteilt werden kann. Auf dieser Basis könnte dann der Schwellenwert der erforderlichen Effektstärke nachvollziehbar abgeleitet werden.

Diagnostische Verfahren in Kapitel 3.5

Mit Blick auf die diagnostisch-therapeutische Kette zum Nutznachweis eines diagnostischen Verfahrens, verweisen die „Allgemeinen Methoden“ auf die Verfahrensordnung des Gemeinsamen Bundesausschusses (G-BA), welcher diesem Prinzip folge. Dazu bleibt festzuhalten, dass der G-BA nicht zwingend die höchste Evidenzstufe fordert, sondern auch andere Evidenzstufen mit Studien zur Testgenauigkeit berücksichtigt, die nicht

¹ World Health Organization (2022): Health Technology Assessment, Internet: www.who.int/health-topics/health-technology-assessment#tab=tab_1 (Zugriff 20.12.2022)

² Cochrane Deutschland (2022): Health Technology Assessment (HTA), Internet: www.cochrane.de/hta (Zugriff 20.12.2022)

³ Bundesärztekammer (2022): Health Technology Assessment, Internet: www.bundesaerztekammer.de/themen/aerzte/qualitaetssicherung/health-technology-assessment (Zugriff 20.12.2022)

⁴ National Institute for Health and Care Excellence/ NICE (2022): Technology appraisal guidance, Internet: www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance (Zugriff 20.12.2022)

⁵ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG, 2020): Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V. Rapid Report A19-43. IQWiG Berichte Nr. 863, Internet: https://www.iqwig.de/download/a19-43_versorgungsnaher-daten-zum-zwecke-der-nutzenbewertung_rapid-report_v1-1.pdf (Zugriff 11.01.2023)

unbedingt mit einer Untersuchung des therapeutischen Nutzens verbunden sein müssen. Die Verfahrensordnung spricht differenziert vom diagnostischen oder therapeutischen Nutzen. Wenngleich ein Nutznachweis mittels diagnostisch-therapeutischer Kette den Idealfall darstellt, sind die Aufwände größer und zeitintensiver, so dass in der Verfahrensordnung realistischer Weise auch Evidenzniveaus ohne gleichzeitige Untersuchung zum therapeutischen Nutzen zum Tragen kommen. Systematische Reviews mit Studien zur Testgenauigkeit bekleiden bereits das Evidenzniveau IIa, welches ein hohes, obschon nicht das höchste Niveau abbildet.⁶

Würde man z.B. unter akuten Bedingungen einer neuen Pandemie, die schnelles Handeln erfordert, die diagnostisch-therapeutische Kette zwingend einfordern, so wären Corona-Tests erst viel später verfügbar gewesen, wodurch das Präventionspotenzial vermindert worden wäre. Nach Aussage des Bundesministeriums für Gesundheit, dem Robert Koch Institut und der Bundeszentrale für gesundheitliche Aufklärung sind Corona-Tests wichtige Instrumente bei der Eindämmung der Corona-Pandemie.⁷ Auch Cochrane Reviews wie etwa von Dinnes et al. 2022 weisen auf die Bedeutung der diagnostischen Testgenauigkeit hin.⁸ So beruht der bisherige Einsatz maßgeblich auf Studien zur diagnostischen Genauigkeit, nicht aber auf dem Abwarten von Studien zur diagnostisch-therapeutischen/-präventiven Kette als Ganzes, wenngleich diese in Ergänzung wünschenswert sind. Entsprechend beziehen viele pragmatische Maßnahmen zur Pandemiebekämpfung auch indirekte Linked-Evidence-Aspekte ein. Insofern begrüßen wir im IQWiG-Methodenpapier die Einlassung, dass auch Studien, die getrennt voneinander zum einen die diagnostische Genauigkeit und zum anderen therapeutische/ präventive Interventionen untersuchen, bei vergleichbarer Patientenselektion durch logische Verknüpfung miteinander verbunden werden können, um einen Nutznachweis zu erbringen. Wir unterstützen die Aussage im Methodenpapier, dass im Zuge des Linked-Evidence-Ansatzes Studien zur Testgüte von primärem Interesse sind.

Zudem möchten wir hinsichtlich des diagnostischen Erkenntnisgewinns grundsätzlich auf prognostische Kriterien, Ursachenklärung von Krankheitszuständen und die Planungen für Patienten und Behandelnde hinweisen, die auch ohne unmittelbare therapeutische Konsequenz von Interesse und für die Lebensqualität relevant sein können.

Wir begrüßen im IQWiG-Methodenpapier den Einsatz internationaler Instrumente wie etwa QUADAS-2 (a revised tool for the Quality Assessment of Diagnostic Accuracy Studies, Methodenpapier-Referenz 766, 767) bei der Bewertung der Ergebnissicherheit, oder wie etwa STARD (Standards for Reporting of Diagnostic Accuracy, Methodenpapier-Referenz 78,79) beim Ein- oder Ausschluss nicht im Volltext publizierter Studien. Dabei handelt es sich um Instrumente, die z.B. auch von der Cochrane Screening and

⁶ Gemeinsamer Bundesausschuss (2023): Verfahrensordnung des Gemeinsamen Bundesausschusses, Internet: www.g-ba.de/informationen/richtlinien/42/ (Zugriff 02.01.2022)

⁷ Bundesministerium für Gesundheit, Robert Koch Institut und Bundeszentrale für gesundheitliche Aufklärung (2023): Corona-Test, Internet: www.zusammengegegen corona.de/testen/ (Zugriff 02.01.2022)

⁸ Dinnes et al. (2022): Rapid, point-of-care antigen tests for diagnosis of SARS-CoV-2 infection, Internet www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD013705.pub3/full/de (Zugriff 02.01.2022)

Diagnostic Tests Methods Group (SDTM) zitiert werden.⁹ Hier könnte das PRISMA-DTA-Instrument (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Diagnostic Test Accuracy Studies) ergänzt werden, welches als spezifischer Standard für die Berichterstattung von systematischen Reviews über Testgütestudien dient.¹⁰

Kosten-Nutzen-Bewertung medizinischer Interventionen in Kapitel 4

Mit der grundlegenden Überarbeitung der Kosten-Nutzen-Bewertung medizinischer Interventionen hat das IQWiG weitreichende Anpassungen in der Methodik der gesundheitsökonomischen Evaluation vorgenommen. Aufgrund des Umfangs der Änderungen bezieht sich die Stellungnahme nur auf ausgewählte Aspekte, die aus Sicht der DKG besonders relevant erscheinen.

Auswahl der Komparatoren in Kapitel 4.3.2

Bislang hat die vom IQWiG beschriebene Methodik auf die Herleitung einer Effizienzgrenze als zentrales Ergebnis der gesundheitsökonomischen Analyse abgezielt. Dabei galt es alle im Indikationsgebiet versorgungsrelevanten Therapieoptionen in die Kosten-Nutzen-Bewertung einzubeziehen. Mit der Version 7.0 des Methodenpapiers, wird das Vorgehen nun grundlegend überarbeitet. Die Darstellung der Effizienzgrenze steht nicht länger im Vordergrund. Damit ändert sich auch die Auswahl der Komparatoren. Der Referenzfall für die Kosten-Nutzen-Bewertung nach § 35b SGB V beschreibt nun, dass in der Regel zwei Therapieoptionen betrachtet werden - das zu bewertende Arzneimittel und die in der vorangegangenen Nutzenbewertung nach § 35a SGB V herangezogene zweckmäßige Vergleichstherapie. Für den Fall, dass weitere Komparatoren einbezogen werden sollen, ist dies im Auftrag zu konkretisieren.

Mit der beschriebenen Fokussierung auf die herangezogene zweckmäßige Vergleichstherapie (zVT) als Komparator, wurde die Methodik der Kosten-Nutzen-Bewertung deutlich verschlankt. Anzumerken ist, dass die Auswahl des Komparators die Ergebnisse der Kosten-Nutzen-Bewertung maßgeblich determinieren wird. Insbesondere bei der Betrachtung nur eines Komparators muss dieser mit hoher Sorgfalt ausgewählt werden und sollte die Behandlung widerspiegeln, die der aktuellen Versorgungspraxis am besten entspricht.¹¹ Mit der Bestimmung der zweckmäßigen Vergleichstherapie im Rahmen des Nutzenbewertungsverfahrens wird dieser Anforderung aus unserer Sicht Rechnung ge-

⁹ Cochrane Screening and Diagnostic Tests Methods Group (SDTM, 2023): Cochrane DTA Support & Guidance, Internet: [www.methods.cochrane.org/sdt/](http://methods.cochrane.org/sdt/) (Zugriff 02.01.2023)

¹⁰ PRISMA (Transparent Reporting of Systematic Reviews and Meta-Analyses, 2023): PRISMA for Diagnostic Test Accuracy, Internet: www.prisma-statement.org/Extensions/DTA (Zugriff 02.01.2023)

¹¹ European Network for Health Technology Assessment (EUnetHTA) (2020): Practical considerations when critically assessing economic evaluations, Internet: https://www.eunetha.eu/wp-content/uploads/2020/03/EUnetHTA-JA3WP6B2-5-Guidance-Critical-Assessment-EE_v1-0.pdf (Zugriff 19.01.2023)

tragen. Die vom IQWiG beschriebene zusätzliche Szenarioanalyse von zVT-Kosten ist hierbei jedoch essentiell um die Variation möglicher Kosten anderer Therapieoptionen abzubilden.

Weiterhin beschreibt das IQWiG, dass die Ergebnisse der vorangegangenen Nutzenbewertung nach § 35a SGB V in die Kosten-Nutzen-Bewertung einbezogen werden sollen. Darüber hinaus wird explizit erwähnt, dass die Ergebnisse der vorangegangenen Nutzenbewertung nicht in Frage zu stellen sind. So soll beispielsweise im Falle, dass die Nutzenbewertung keinen Zusatznutzen festgestellt, auch in der Kosten-Nutzen-Bewertung kein Zusatznutzen suggeriert wird. Es ist anzumerken, dass die synergetische Verwendung der Ergebnisse der Nutzenbewertung nach § 35a SGB V aus unserer Sicht insbesondere bei der Kosten-Nutzwert-Analyse nur begrenzt möglich sein wird und in diesem Fall mit einer deutlichen Erweiterung der für die Analyse notwendigen Daten zu rechnen ist.

Nicht nachvollziehbar ist aus unserer Sicht, wie konkret mit Orphan Drugs zu verfahren ist, für die keine reguläre Nutzenbewertung erfolgt. Hierfür beschreibt die Version 7.0 des Methodenpapiers, dass im Rahmen des Auftrags für die Kosten-Nutzen-Bewertung ein Komparator durch den G-BA ggf. erst festzulegen ist. Dies ist nach unserem Verständnis in zweierlei Hinsicht kritisch zu bewerten. Erstens, durchlaufen Orphan Drugs nach Marktzugang nicht das reguläre Nutzenbewertungsverfahren. Es ist keine zweckmäßige Vergleichstherapie durch den G-BA festzulegen und der Zusatznutzen gilt bereits durch die Zulassung als belegt (§ 35a SGB V Abs. 1 Satz 11). Für den Fall, dass im Rahmen der Kosten-Nutzen-Bewertung nun eine zweckmäßige Vergleichstherapie bestimmt werden muss, wird dieses Prinzip aus unserer Sicht verletzt. Zweitens ist für uns nicht ersichtlich, inwieweit die vom IQWiG benannte Prämisse, dass das Ergebnis der vorangegangenen Nutzenbewertung durch die anschließende Kosten-Nutzen-Bewertung nicht in Frage gestellt werden darf, im Fall einer eingeschränkten Nutzenbewertung von Orphan Drugs aufrechterhalten werden kann. Der Zusatznutzen für Orphan Drugs muss im Nutzenbewertungsverfahren nicht durch einen Vergleich mit weiteren Therapiealternativen nachgewiesen werden. Liegen keine geeigneten Daten für einen Vergleich vor, kann der Zusatznutzen in der Regel nicht quantifiziert werden. Wird nun im Rahmen der Kosten-Nutzen-Bewertung ein Vergleich vorgenommen, kann das Ergebnis der vorangegangenen Nutzenbewertung aus unserer Sicht durchaus in Frage gestellt werden. Deutlich wird dies am Beispiel von Orphan Drugs, die die Umsatzschwelle überschritten haben und für die sich eine reguläre Nutzenbewertung angeschlossen hat. In diesen Fällen können sich die Ergebnisse hinsichtlich des Zusatznutzens durchaus unterscheiden.¹²

¹² Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (2021): Evidenz zu Orphan Drugs, Internet: https://www.iqwig.de/download/ga21-01_evidenz-zu-orphan-drugs_arbeitspapier_v1-0.pdf (Zugriff 19.01.2023)

Gesundheitsökonomische Zielgröße in Kapitel 4.3.3

Um die Ergebnisse der vorangegangenen Nutzenbewertung nach § 35a SGB V für die gesundheitsökonomische Evaluation synergetisch nutzbar zu machen, bietet sich aus Sicht des IQWiG die Kosten-Effektivitäts-Analyse an. Ein wichtiger Punkt ist hierbei die Auswahl der Endpunkte, die in der Kosten-Nutzen-Bewertung betrachtet werden sollen. Hierbei soll der Fokus laut IQWiG-Methodenpapier auf die Endpunkte gelegt werden, die im Rahmen der Nutzenbewertung entscheidungsleitend waren. Dies setzt voraus, dass klar erkenntlich ist, welche Endpunkte als entscheidungsleitend gelten können.

Neben der Kosten-Effektivitäts-Analyse, sieht das IQWiG-Methodenpapier auch die Möglichkeit der Kosten-Nutzwert-Analyse vor. Während bei der Kosten-Effektivitäts-Analyse klinische Endpunkte betrachtet werden, wird bei der Kosten-Nutzwert-Analyse das qualitätsadjustierte Lebensjahr (quality adjusted life year, QALY) als Nutzenmaß zugrunde gelegt. Vor dem Hintergrund, dass nach § 35b SGB V internationale Standards der Gesundheitsökonomie zu beachten sind, ist es aus unserer Sicht zu begrüßen, dass dieser Ansatz in der Methodik des IQWiG dargestellt wird. Vorteile des QALY Konzepts liegen beispielsweise in der Möglichkeit indikationsübergreifender Vergleiche.¹³ Zur Erfassung der QALYs schlägt das IQWiG vor, direkte Verfahren, den indirekten Verfahren vorzuziehen. Es beruft sich hierbei auf den Stellenwert der Perspektive der Patientinnen und Patienten für das Institut und auf die Vorgaben des SGB V. Zu bedenken ist, dass die Datenlage hierzu jedoch eingeschränkt sein kann, was auch im IQWiG-Methodenpapier erwähnt wird. Wenngleich auch wir den Stellenwert des Patienten-Nutzens als hoch einordnen, wird im internationalen Kontext oftmals auf Daten aus indirekten Verfahren zurückgegriffen.

Ergebnisdarstellung in Kapitel 4.10

Wie eingangs beschrieben zielt die gesundheitsökonomische Analyse nach dem QWWiG-Methodenpapier nicht länger auf die Darstellung einer Effizienzgrenze ab. Das zentrale Ergebnis bildet nun das inkrementelle Kosten-Effektivitäts-Verhältnis (incremental cost-effectiveness-ratio, ICER) für die zu bewertende Therapie und die herangezogene zweckmäßige Vergleichstherapie. Abweichende Therapiekosten anderer Behandlungsoptionen werden ebenfalls betrachtet (Szenarioanalysen zVT-Kosten). Es wird keine explizite Preisempfehlung abgeleitet. Die Kosten-Nutzen-Bewertung soll dazu dienen, ergänzende Informationen für die Preisverhandlungen bereitzustellen.

Nach unserer Einschätzung eignet sich die gesundheitsökonomische Evaluation grundsätzlich für das benannte Ziel, erweiterte Informationen, die im Rahmen der Nutzenbe-

¹³ European Network for Health Technology Assessment (EUnetHTA) (2015): Endpoints used for Relative Effectiveness Assessment: HEALTH-RELATED QUALITY OF LIFE and UTILITY MEASURES, Internet: https://www.eunetha.eu/wp-content/uploads/2018/01/Endpoints-used-for-Relative-Effectiveness-Assessment-Health-related-quality-of-life-and-utility-measures_Amended-JA1-Guideline_Final-Nov-2015.pdf (Zugriff 19.01.2023)

wertung nach § 35a SGB V nicht erfasst werden (können), zu generieren. Eine wesentliche Stärke liegt beispielsweise darin, dass die Betrachtung eines längeren Zeithorizonts erfolgen kann. Dies ist insbesondere für den Fall bedeutsam, in dem die alleinige Schätzung der Jahrestherapiekosten nicht zielführend ist, wie zum Beispiel bei der Einmalgabe eines Arzneimittels.

Im Hinblick auf die Änderung der Methodik hinsichtlich der Ableitung einer Effizienzgrenze begrüßen wir die grundlegende Überarbeitung. Die vom IQWiG zuvor beschriebene Methodik war unserer Sicht mit Limitationen behaftet und führte nur in idealtypischen Marktkonstellationen zu angemessenen Ergebnissen. Es bleibt aus unserer Sicht jedoch unklar, inwieweit anhand des ICER das intendierte Ziel erreicht werden kann, die Angemessenheit eines Preises zu bewerten. Wie hoch die Zahlungsbereitschaft für einen Zugewinn an Nutzen ist, ist eine Werteentscheidung.

Qualitätsbewertung systematischer Übersichten mittels AMSTAR 2 in Kapitel 7

Wie bereits in Kapitel 9.2.1 als Alternative beschrieben, könnte In Ergänzung zum AMSTAR-2-Instrument (A MeaSurement Tool to Assess systematic Reviews, Methodenpapier-Referenz 662) auch in Kapitel 7 das ROBIS-Instrument (Risk Of Bias In Systematic reviews, Methodenpapier-Referenz 765) zum Einsatz kommen, um systematische Übersichtsarbeiten einer Qualitätsbetrachtung zu unterziehen.

Literatursuche in Kapitel 8

Zur Literatursuche in Kapitel 8 möchten wir anregen, auch die „Specialized Registers“ relevanter themenbezogener Cochrane Review Groups und deren Erfahrungen mit Suchstrategien und Suchstrings zu nutzen (siehe auch Chapter 4, Cochrane Collaboration Handbook 2022, Cochrane Collaboration Review Groups 2022).^{14,15}

Kriterien für den Einschluss von Studien in Kapitel 9.1.1

Hinsichtlich der Einschlusskriterien in systematische Übersichten (SR) von Studien, welche die für die Nutzenbewertung relevante definierte Patientenpopulation bzw. Prüf- und Vergleichsintervention nicht vollständig erfüllen, wäre es hilfreich, Referenzen anzugeben, die das Vorgehen im IQWiG-Methodenpapier stützen. Da die Studienselektion ein ausgesprochen sensibler Gesichtspunkt bei der Durchführung von SR ist und erheblichen

¹⁴ Cochrane Collaboration (2022): Cochrane Handbook for Systematic Reviews of Interventions | Cochrane Training, Internet: www.training.cochrane.org/handbook/current (Zugriff 20.12.2022)

¹⁵ Cochrane Collaboration (2022): Review Groups | Cochrane, Internet: www.cochrane.org/about-us/our-global-community/review-groups (Zugriff 20.12.2022)

Einfluss auf die Ergebnisse von systematischen Übersichten hat, ist eine differenzierte Betrachtung angezeigt.¹⁶

Aus unserer Sicht wäre auch die Zusammenarbeit mit Fachexperten des jeweiligen Forschungsgebiets maßgeblich, um sachgerechte Entscheidungen zur medizinischen Relevanz zu treffen. Die im IQWiG-Methodenpapier favorisierte Verwendung von Analysen relevanter Subgruppen oder die Festlegung von Cut-Offs auf mindestens 80% erfüllter Kriterien, um einen Einschluss zu begründen, sind von quantitativer Natur, die den qualitativen Aspekt der Prüfung von Einschlusskriterien vernachlässigt. So wäre wünschenswert, sowohl quantitative als auch explizit qualitative Aspekte ergänzend und/ oder alternativ einzubeziehen. Ideal wäre bei der Prüfung die Kombination von spezifischer Fachexpertise aus dem jeweiligen medizinischen Forschungsfeld mit unabhängiger methodischer Expertise; letztere kann insofern unabhängiger prüfen, da sie die jeweilige Forschergemeinschaft weniger kennt und deshalb unvoreingenommener einschätzen kann. Auch eine Gewichtung von Einschlusskriterien, die z.B. in Teilen über 80% und in Teilen unter 80% liegen oder quantitativ nicht abschließend bestimmbar sein können, berührt qualitative Aspekte, so dass ein flexiblerer Umgang sachgerecht wäre. Cut-Offs bieten zweifelsohne Orientierung, die qualitative Gesichtspunkte inhaltlicher Merkmale ergänzen. Zudem sollte betont werden, dass es sich um Definitionen handelt, die vorab bestimmt werden, um der Beliebigkeit beim Erstellen von SR vorzubeugen. Die Möglichkeit eines mehrschrittigen Prozesses zur Herleitung klarer Vorab-Definitionen von Population, Intervention und Vergleichsgruppe kann dabei hilfreich sein.¹⁷

Adjustierung und multifaktorielle Verfahren in Kapitel 9.3.6

Das Kapitel wird um Ausführungen zur Adjustierung für Confounder im Rahmen von anwendungsbegleitenden Datenerhebungen (AbDs) ergänzt. Es wird darauf eingegangen, dass im Rahmen einer AbD zur Adjustierung für Confounder insbesondere die Propensity-Score-Methode eine Rolle spielt. Eine Grundvoraussetzung ist, dass für alle Patientinnen und Patienten, die in die Auswertung eingehen, die zu vergleichende Intervention eine Therapieoption darstellt (Positivität). Es muss eine ausreichende Überlappung zwischen den beiden zu vergleichenden Gruppen bezüglich der Propensity Scores gegeben sein und die beiden zu vergleichenden Gruppen müssen bezüglich aller relevanten Confounder ausreichend balanciert sein. Es bleibt unklar, mit welchen Methoden die relevanten Confounder und ihre Wechselwirkungen identifiziert werden sollten und nach welchen Kriterien, ggf. auf der Basis kausaler Grafiken eine Auswahl der ins Modell eingehenden Confounder getroffen werden soll. Diese Erläuterungen sind vor dem Hintergrund, dass bei der überwiegenden Anzahl der Verfahren im Kontext des

¹⁶ Cochrane Collaboration (2022): Cochrane Handbook for Systematic Reviews of Interventions (insbesondere Part 2, Chapter 3 und 4), Internet: www.training.cochrane.org/handbook/current (Zugriff 02.01.2022)

¹⁷ Cochrane Collaboration (2022): Cochrane Handbook for Systematic Reviews of Interventions (insbesondere Part 2, Chapter 3 und 4), Internet: www.training.cochrane.org/handbook/current (Zugriff 02.01.2022)

Arzneimittelmarktneuordnungsgesetzes (AMNOG-Verfahren), bei denen pharmazeutischen Unternehmen Daten aus indirekten Vergleichen mit Adjustierung mittels Propensity Score Matching eingereicht haben, diese Analysen abgelehnt wurden¹⁸¹⁹. Die Ablehnungen erfolgten häufig anhand der Aussage, es seien keine ausreichenden systematischen Literaturrecherchen erfolgt oder die Auswahl der Confounder sei nicht ausreichend begründet oder nicht sachgerecht. Diese Situation zeigt, dass dringender Erläuterungsbedarf besteht. Die Beantwortung dieser Fragen ist auch deshalb dringend, weil in jedem der bislang vom IQWiG erstellten Konzepte und vom G-BA beauftragten anwendungsbegleitenden Datenerhebungen zur adäquaten Adjustierung für Confounder die Propensity-Score-Methode empfohlen wird.

Ergänzungen in Kapitel 9.3.12 zur Auswirkung systematisch verkürzter Beobachtungszeiten bei der Erhebung patientenberichteter Endpunkte (PROs)

Hinsichtlich möglicher Verzerrungsquellen stimmen wir zu, dass Studien mit verkürzten Beobachtungszeiten bei patientenberichteten Endpunkten (PROs) schwieriger zu interpretieren sind. Die Forderung jedoch nach kategorischem Ausschluss können wir nicht folgen; vielmehr wären pragmatische Vorschläge hilfreich, wie fehlende Werte verhindert, wie PROs auch nach Behandlungswechsel bzw. -abbruch erhoben oder wie vorhandene, wenngleich unvollständige Daten doch noch sinnvoll genutzt werden könnten, um den realen Gegebenheiten der Studiendurchführung gerecht werden und deren Hindernissen differenziert begegnen zu können. Damit wird die umfangreiche Literatur zum Themenfeld „Missing Data“ angesprochen, die zum Teil Eingang in die „Allgemeinen Methoden“ des IQWiG gefunden hat (siehe z.B. Methodenpapier-Referenz 133, 240, 269, 487, 561). Nützlich wäre auch der stärkere Einbezug aktuellerer Literatur, die weitere konstruktive Ideen enthält, wie etwa Aiyegbusi et al. 2022,²⁰ Coens et al. 2020²¹ oder Mercieca-Bebber et al. 2016.²²

¹⁸ Gemeinsamer Bundesausschuss (2014): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Ipilimumab (neues Anwendungsgebiet), Internet: https://www.g-ba.de/downloads/40-268-2835/2014-06-05_AM-RL-XII_Ipilimumab_nAwg_2013-12-15-D-090_TrG.pdf (Zugriff 26.01.2023).

¹⁹ IQWiG (2022): [A22-04] Amivantamab (NSCLC) – Nutzenbewertung gemäß § 35a SGB V. IQWiG-Berichte – Nr. 1329, Internet: https://www.iqwig.de/download/a22-05_amivantamab_kurzfassung_nutzenbewertung-35a-sgb-v_v1-0.pdf (Zugriff 26.01.2023).

²⁰ Aiyegbusi et al. (2022): Key considerations to reduce or address respondent burden in patient-reported outcome (PRO) data collection | Nature Communications, Internet: www.nature.com/articles/s41467-022-33826-4 (Zugriff 20.12.2022)

²¹ Coens et al. (2020): Scholarly Article or Book Chapter | International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium | ID: 76537b278 | Carolina Digital Repository (unc.edu), Internet: <https://cdr.lib.unc.edu/concern/articles/76537b278?locale=en> (Zugriff 20.12.2022)

²² Mercieca-Bebber et al. (2016): Design, implementation and reporting strategies to reduce the instance and impact of missing patient-reported outcome (PRO) data: a systematic review | BMJ Open, Internet: <https://bmjopen.bmj.com/content/6/6/e010938> (Zugriff 20.12.2022)

Bekanntermaßen haben wir schon zu Vorgänger-Versionen der „Allgemeinen Methoden“ Stellungnahmen abgegeben. Deren Inhalte haben – soweit sie nicht schon berücksichtigt wurden – prinzipiell auch weiterhin Bestand und wurden, zur Vermeidung von Redundanzen, hier insofern nicht sämtlich wiederholt. Wir hoffen, dass wir Ihnen wertvolle Hinweise für die Überarbeitung des Entwurfes geben konnten.

Diese Stellungnahme ist als institutionelle Meinungsäußerung aufzufassen, an der verschiedene Mitarbeiterinnen und Mitarbeiter der DKG beteiligt waren. Im Falle einer mündlichen Anhörung gehen wir davon aus, eine Vertretung der DKG zu entsenden.

Mit freundlichen Grüßen
In Vertretung



Dr. Michael Brenske
Leiter des Geschäftsbereichs VI - Evidenzbasierte medizinische Versorgung und G-BA

A.1.20 Deutsches Netzwerk Versorgungsforschung e. V. (DNVF)

Autorinnen und Autoren

- Icks, Andrea
- Stausberg, Jürgen
- Pieper, Dawid
- Ernstmann, Nicole

Zusammenfassung Stellungnahme IQWiG Allgemeine Methoden Version 7 verschiedener Arbeitsgruppen des Deutschen Netzwerk Versorgungsforschung

Prof. Dr. Andrea Icks (Düsseldorf) für die AG Gesundheitsökonomie

Kap. 1.4 und Kap. 4

Zu den Kapiteln 1.4 und 4 "Gesundheitsökonomie" und „Kosten-Nutzen-Bewertung medizinischer Interventionen“ schließt sich das DNVF e.V. der Stellungnahme der Deutschen Gesellschaft für Gesundheitsökonomie (DGGÖ) an.

Prof. Dr. Max Geraedts (Marburg), Prof. Dr. Silke Kuske (Düsseldorf), Felix Walther (Dresden) – AG Qualitäts- und Patientensicherheitsforschung

Kap. 3.1

Wir begrüßen die einheitliche und konsistent angewendete Terminologie und klaren Erläuterungen im Kapitel 3.1. Im Hinblick auf die Kerndefinition „höheren Schadens/ Nutzens, vergleichbaren Schadens/ Nutzens und geringeren Schadens/ Nutzens“ möchten wir anregen, die relationalen Klassifizierungen in „hoch, vergleichbar und gering“ anhand eines Beispiels im Generellen als auch unter dem Punkt der Interventions- und damit Patient:innensicherheit darzulegen. Das würde dem Verständnis und der Einordnung sehr helfen. Darüber hinaus sollte angemerkt werden, ab wann ein höherer / geringerer Nutzen / Schaden vorliegt. Eventuelle könnte schon in diesem Abschnitt auf Effektstärken bzw. das später genannte Konzept der Einstufung hingewiesen werden.

Zudem sollten unseres Erachtens, anstatt Patientenzufriedenheit alleine zu benennen, wie heute üblich auch Patientenerfahrungen als Nutzendimension einbezogen werden; sollte eine Intervention zwar eine geringe Verbesserung der Funktionalität aber gleichzeitig weitaus schlechtere Patientenerfahrungen (z. B. inakzeptabel, verstößt gegen übliche Präferenzen/Werte der Betroffenen, zu hoher Aufwand, ...) kann es durchaus sein, dass aufgrund der Patientenerfahrungen eine Intervention bevorzugt werden sollte, die aber einen geringeren Nutzen im Bereich der Funktionalität besitzt. Daher ist unseres Erachtens auch hierdurch ein höherer Nutzen möglich.

Prof. Dr. Jürgen Stausberg (Essen), Dr. Anna Niemeyer (Berlin), Prof. Dr. Michael Hauptmann (Neuruppin), Dr. Sandra Kluge (Hamburg) für die AG Register

Kap. 3.3.4

Kommentare zu Abschnitt 3.3.4 Anwendungsbegleitende Datenerhebungs-Konzepte

1. Eine Datenplattform hat für den Zeitraum und den Umfang der AbD die zur jeweiligen Nutzenbewertung erforderlichen Qualitätskriterien zu erfüllen. Es ist dabei unerheblich, ob diese Qualitätskriterien bereits zum Zeitpunkt der Erstellung des AbD-Konzepts erfüllt werden. Die Prüfung bei der Erstellung des AbD-Konzepts würde sich daher auf die prinzipielle Eignung bzw. das Potenzial einer bestehenden Plattform beziehen. Die Möglichkeit einer qualifizierten Bewertung bestehender Datenplattformen unter Darstellung ihres Potentials wäre daher sinnvoll. Es wird daher vorgeschlagen, im vierten Spiegelpunkt des 2. Absatzes folgenden

Einschub einzufügen: „... gebeten. *Über eine Bekanntgabe der Recherche wird auch weiteren Registerbetreibern die Möglichkeit zur Einreichung dieser Informationen gegeben. Anhand dieser ...*“.

2. Bestehenden Datenplattformen sollte die Möglichkeit gegeben werden, sich potentiell auf eine anwendungsbegleitende Datenerhebung vorzubereiten. Falls seitens des IQWiG aus den genannten Quellen (Literaturstellen 242,248,299,500,691) von einer konkreten AbD unabhängige Qualitätskriterien identifiziert werden, sollten diese beschrieben sein. Für diesen Fall wird vorgeschlagen, einen neuen Spiegelpunkt in Absatz 2 mit Auflistung prinzipieller Qualitätskriterien aufzunehmen.
3. Insbesondere bei Arzneimitteln für seltene Erkrankungen kann nicht regelhaft davon ausgegangen werden, dass im Rahmen der AbD eine Nutzenbewertung im Vergleich gegenüber einer zweckmäßigen Vergleichstherapie erfolgen kann. In diesen Fällen können alternativ Daten zum Erkrankungsverlauf aus einer bestehenden Datenplattform oder aus anderen Quellen herangezogen werden. Es wird daher vorgeschlagen, in Spiegelpunkt 5 in Absatz 2 vor den letzten Satz folgenden Einschub vorzunehmen: *Die Verwendung vorhandener Daten aus dieser oder anderen Datenquellen für den Vergleich mit dem Arzneimittel der AbD ist zu beschreiben und zu begründen.*
4. Bei der Dauer einer AbD nach Fertigstellung des AbD-Konzepts sind neben der Betriebs- und Auswertungsphase auch die Zeiträume für die Erstellung des Studien- bzw. Registerprotokolls sowie für die Umsetzung einer neuen bzw. die Anpassung einer bestehenden Datenplattform zu berücksichtigen. Es wird daher folgende Ergänzung als neuer Satz 1 in Spiegelstrich 6 in Absatz 2 vorgeschlagen: *Die Dauer einer AbD umfasst die Planung mit Erstellung des Studien- und Registerprotokolls, die Umsetzung oder Anpassung der Datenplattform, den Betrieb mit Rekrutierung und Beobachtung der einzelnen Patientinnen und Patienten sowie die Auswertung.* Der bisherige Satz 2 wären dann anzupassen: „~~Dazu~~*-Zum Betrieb* gehört zum einen ...“
5. Als Folge des Vorschlags Nummer 5 wäre der letzte Satz in Spiegelstrich 7 in Absatz 2 anzupassen: „... berücksichtigt. Basierend auf den vorliegenden Erkenntnissen zum Arzneimittel und ggf. zur zweckmäßigen ...“.
6. Beschreibung des Umfangs der AbD (alternativer Textvorschlag): Eine in einer nicht randomisierten Studie beobachtete Assoziation der Intervention mit einem Nutzen oder Schaden für die Patienten gegenüber einer zweckmäßigen Vergleichstherapie kann aufgrund potenzieller Confounder nur nach sorgfältiger Prüfung der Intervention zugeordnet werden. Teil dieser Prüfung sind die Ergebnisse der Adjustierung für gemessene Confounder, die Möglichkeit residualen Confoundings durch ungenaue Messung der gemessenen Confounder, die Existenz ungemessener Confounder sowie die Richtung und Stärke deren Assoziation mit der Intervention und dem Endpunkt. Diese Punkte werden bei der Frage, ob die geplante AbD prinzipiell sinnvoll durchführbar ist, berücksichtigt. Dies schließt bei seltenen Erkrankungen die in der Regel geringe oder sehr geringe Fallzahl ein. Basierend auf den vorliegenden Erkenntnissen zum Arzneimittel und zur zweckmäßigen Vergleichstherapie wird abgeschätzt, ob innerhalb eines akzeptablen Zeitraums prinzipiell ausreichend Patientinnen und Patienten mit ausreichend detaillierten Daten über potenzielle Confounder rekrutiert und beobachtet werden können, um mit einer AbD grundsätzlich aussagekräftige Ergebnisse für eine Nutzenbewertung generieren zu können.
7. Bei sachgerechter Durchführung einer AbD gemäß AbD-Konzepts sollte eine verlässliche Berücksichtigung der Ergebnisse folgen. In Absatz 3 wird daher ein weiterer Spiegelpunkt vorgeschlagen: *„Alle Gesichtspunkte, die für eine Berücksichtigung der Ergebnisse der AbD relevant sind.“*

Kap. 9.3.6 (S. 177):

8. Unter Positivität versteht man, dass die im Rahmen der AbD (anzupassender Text.) „zu vergleichenden Interventionen (statt zu vergleichenden Interventionen à zu vergleichenden Behandlungsformen in der Versorgungspraxis)“ für alle Patientinnen und Patienten, die in die Auswertung eingehen, eine Therapieoption darstellen. Dies ist im Rahmen der Einschlusskriterien für die AbD zu prüfen und zu beschreiben. In der zur Verfügung stehenden Datenbasis muss eine ausreichende Überlappung zwischen den beiden zu vergleichenden Gruppen bezüglich der Verteilung des Propensity Scores gegeben sein. Weiterhin ist es erforderlich, dass die beiden zu vergleichenden Gruppen bezüglich aller relevanten Confounder ausreichend balanciert sind.

Begründung:

Gemäß IQWiG Rapid Report ‚Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V‘³ sind in Registern grundsätzlich nicht interventionelle und interventionelle vergleichende Studiendesigns möglich. Die hier gewählte Formulierung – wenngleich sie sich auch im PICO-Schema wiederfindet – schließt den nicht interventionellen Ansatz, d.h. die Datenerhebung im Rahmen der ärztlichen Routine (mit bereits zugelassenen Medikamenten) in der Versorgung, nicht eindeutig mit ein.

Prof. Dr. Dawid Pieper für die AG Evidenzsynthese

Kap. 5

5.3 (insbesondere 5.3.2.): Bei Mindestmengen handelt es sich um eine Intervention auf Populationsebene. Die betroffenen Bürger*innen bzw. Patient*innen können nicht wie bei Interventionen auf Individualebene sich dafür oder dagegen entscheiden. Insofern erscheint es sinnvoll dies bei der Betrachtung der Outcomes zu berücksichtigen. Bezüglich der Zielgrößen heißt es: „Diese können sich auf die Mortalität, Morbidität, gesundheitsbezogene Lebensqualität einschließlich der Aktivitäten des täglichen Lebens und der Abhängigkeit von der Hilfe anderer Personen beziehen.“ Genannt sind hier nur Zielgrößen auf Individualebene. Bei Einführung einer Mindestmenge sind aber auch Zielgrößen auf Systemebene relevant, wie z.B. die Entfernung oder Fahrzeit zum nächstgelegenen Krankenhaus, das die Leistung erbringt. Auch wenn Interventionsstudien im Kontext Mindestmengenregelung selten sind, so erlauben sie grundsätzlich auch die Analyse von Mengenausweitungen.

5.4 (insbesondere 5.4.4.): Es ist zu hinterfragen, in wie weit Daten der Statistischen Bundesamtes als Befragungsdaten deklariert werden können. Regelhaft beruhen diese nicht auf einer Befragung, sondern auf einer standardisierten Abfrage.

Prof. Dr. Nicole Ernstmann für die AG Gesundheitskompetenz

Kap. 7.4

Die AG begrüßt die Erweiterung von Kapitel 7.4 zur Auswahl und Bewertung der in Patienteninformationen eingehenden externen Evidenz. Es wird deutlich, dass bei der Auswahl der Evidenz systematische Übersichtsarbeiten im Vordergrund stehen. In diese Richtung gehen auch die meisten Ergänzungen. Grundsätzlich hat der Ansatz, Übersichtsarbeiten bei der Evidenzauswahl in den

Vordergrund zu stellen, den Nachteil, dass hochwertige Übersichtsarbeiten mit einer profunden Studienbasis meistens nur bei wenigen Fragestellungen zu erwarten sind. Die systematische Recherche nach Primärstudien sollten daher, je nach Fragestellung, einen gleichwertigen Stand haben und so in dem Kapitel dargestellt werden. Zu den Bewertungsmethoden für Primärstudien könnten Verweise auf die entsprechenden Kapitel zur allgemeinem Nutzenbewertung genügen.

Im Einzelnen wurde ergänzt, dass identifizierte systematische Übersichtsarbeiten mittels AMSTAR-2 bewertet werden sollen. Dieses Vorgehen ist zu begrüßen, da dieses Tool sich in der Praxis bewährt hat. Es wird allerdings nicht deutlich, wie nach der Bewertung weiter mit den Übersichtsarbeiten verfahren wird. So wird lediglich beschrieben, dass eine geeignete Übersichtsarbeit „nur geringe methodische Mängel“ aufweisen darf. Hier wäre es wünschenswert, zumindest die Domänen des AMSTAR-2 Instruments (evtl. mit Priorisierung) aufzuführen, in denen nur geringe Mängel vorliegen sollten. Bei den weiteren genannten Punkten ist deren relative Bedeutung nicht ersichtlich. Die Nutzung eines mittels AMSTAR-2 ermittelten Score zur Definition eines Mindeststandards ist aus unserer Sicht nicht empfehlenswert. Weiter wurde ergänzt, dass die in eine systematische Übersichtsarbeit einbezogenen Studien „einer adäquaten Bewertung des Verzerrungspotentials unterzogen worden sein müssen“. Hier wäre ebenfalls eine nähere Erläuterung wünschenswert, welche Instrumente dazu aus Sicht des IQWiGs geeignet wären. Der bereits bestehende in der Voraufgabe angeführte Anspruch an Autoren von Übersichtsarbeiten, den Umgang mit grauer Literatur zu bedenken, ist zu begrüßen, da dies oft noch unzureichend beachtet wird. Dies dürfte auch im Hinblick auf die Akzeptanz von Patienteninformationen förderlich sein.

Die Folgekapitel zur Entwicklung und Darstellung der Patienteninformationen wurden im Rahmen der Überarbeitung nicht angepasst. Diese orientieren sich bereits weitreichend an den Leitlinien zur Entwicklung von Patienteninformationen und sind aus unserer Sicht sehr gut ausgeführt. Auch die mittlerweile sehr zahlreichen Varianten der Betroffenenbindung in den Erstellungsprozess zeigen eine starke Nutzerorientierung, die wir sehr begrüßen. Erfahrungen zeigen allerdings, dass gesundheitsinformationen.de einen noch zu geringen Bekanntheitsgrad in der Bevölkerung hat. Strategien zur Erhöhung des Bekanntheitsgrads finden sich leider nicht in Kapitel 7. Ein Absatz dazu, sofern als erforderlich angesehen, wäre wünschenswert.

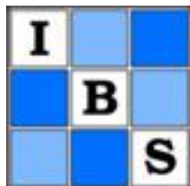
Zusammenfassend besteht noch Potential, den Umgang mit einzubeziehender Evidenz für die Entwicklung von Patienteninformationen weiter auszuführen. Zum anderen ist die bereits bestehende, aus unserer Sicht sehr profunde, Ausarbeitung der Aufbereitung der Informationen und der Einbeziehung Betroffener in den Prozess der Informationsgenerierung, positiv hervorzuheben.

Prof. Dr. Dawid Pieper für die AG Evidenzsynthese

Kap. 8

8.1.1. (insbesondere E): Hier wird festgehalten: „Zusätzlich zur PRESS-Checkliste wird die Suchstrategie gegen ein unabhängiges Set relevanter Referenzen getestet“. Hier wird nicht klar, was mit dem unabhängigen Set gemeint ist und wo dieses herkommt. Für die Entwicklung der Recherchestrategie wird bereits ein Set relevanter Referenzen getestet. Somit können diese nicht gemeint sein, da diese nicht unabhängig wären. Der Schluss liegt somit nahe, dass aus dem Set anfänglich bekannter Referenzen zwei Sets gebildet werden. Dies sollte präziser formuliert werden.

**A.1.21 Deutsche Region der Internationalen Biometrischen Gesellschaft (IBS-DR),
Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie
e. V.**



**Stellungnahme zum Entwurf der Version 7.0 des IQWiG-Papiers
„Allgemeine Methoden“
durch die gemeinsame Präsidiumskommission
„Methodenaspekte in der Arbeit des IQWiG und IQTiG“ der GMDs und IBS-DR**

Im „Allgemeinen Methodenpapier“ werden die gesetzlichen und wissenschaftlichen Grundlagen der Arbeit am Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) erläutert. Die derzeit gültige Fassung ist die Version 6.1 vom 24. Januar 2022. Am 6. Dezember 2022 hat nun das IQWiG einen Entwurf der neuen Version 7.0 vorgelegt. In der Version 7.0 wurden im Vergleich zur Version 6.1 neben redaktionellen Modifikationen auch inhaltliche Änderungen durchgeführt. Grundlegende Überarbeitungen erfolgten in den Kapiteln zu *IQWiG im Gesundheitswesen* (Kapitel 1), *Produkten des Instituts* (Kapitel 2), *Nutzenbewertung medizinischer Interventionen* (Kapitel 3), *Kosten-Nutzen-Bewertung medizinischer Interventionen* (Kapitel 4), *Informationsbeschaffung* (Kapitel 8) und *Informationsbewertung* (Kapitel 9). Zudem soll in Zukunft die Qualitätsbewertung systematischer Übersichten mit AMSTAR 2 statt mit dem Oxman-Guyatt-Index durchgeführt werden; hierzu wurden entsprechende Änderungen in Kapitel 7 „Evidenzbasierte Gesundheitsinformation für Bürgerinnen und Bürger“ eingefügt. Aus biometrischer Sicht sind insbesondere einige Änderungen in Kapiteln 3 und 9 relevant; aus Sicht der Gesundheitsökonomie und der entscheidungsanalytischen Modellierung sind insbesondere einige Textpassagen in den Kapiteln 1.4 und 4 relevant. Auf diese gehen wir unten näher ein.

Die Kommission begrüßt die regelmäßige Überarbeitung der „Allgemeinen Methoden“. Die stetige Fortentwicklung relevanter Methoden macht dies unerlässlich. Wie aus den spezifischen Kommentaren unten deutlich wird, gibt es einige Abschnitte, die einer Aktualisierung bedürfen. Aus Sicht der Kommission ist es auch verständlich, dass es schwierig ist, alle Bereiche immer gleichmäßig aktuell zu halten. Hier wäre insbesondere der Hinweis, sich in den Bereichen Epidemiologie, Versorgungsforschung (z.B. spezielle Studiendesigns wie Stepped Wedge Designs zur Evaluation von Versorgungskonzepten) und Patient Preference Studies zu verstärken oder externe Beratung einzuholen. Die Kommission hat weiterhin wahrgenommen, dass einige Punkte auf Rapid Reports zurückgehen. Diese können als solche nicht kommentiert werden. Wenn Sie ins Methodenpapier eingehen, um z.B. bis dato bestehende Widersprüche zwischen Rapid Report und Methodenpapier aufzulösen, können sie sehr wohl kommentiert werden. Allerdings setzt dies eine ausreichend ausführliche Darstellung im Methodenpapier voraus. Hier gehen wir im weiteren Verlauf beispielhaft auf den Rapid Report A19-43 „Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V“ ein. Die Kommission ist offen für eine fachliche Diskussion mit dem IQWiG in Bezug auf geeignete Methoden und unterstützt das Institut in dieser Hinsicht gerne.

Die Kommission hat zu den folgenden speziellen Punkten des Papiers Kommentierungen durchgeführt:

Abschnitt 1.4 Gesundheitsökonomie und Abschnitt 4 Kosten-Nutzen-Bewertung medizinischer Interventionen

Die Einordnung der Methode 'gesundheitsökonomische Evaluation' in den Gesamtkontext der umfassenderen Disziplin Gesundheitsökonomie in Abschnitt 1.4 wird begrüßt. Ebenso wird die ausführliche Darstellung der Bedeutung der gesundheitsökonomischen Evaluation einschließlich der Begriffsbestimmungen in diesem Abschnitt begrüßt, die der internationalen Terminologie entsprechen.

Abschnitt 3.1.1 Definition des patientenrelevanten medizinischen Nutzens bzw. Schadens

Im Abschnitt 3.1.1 wird als einzige Möglichkeit zur Beteiligung an der Wahl der patientenrelevanten Endpunkte die Einbeziehung einzelner Betroffener erwähnt. Während dies dem allgemeinen Vorgehen zur Einbindung von Betroffenen in die Erstellung von IQWiG-Produkten entspricht, sollten aus Sicht der Kommission neuere Entwicklungen und systematische Möglichkeiten für die Zielsetzungen in diesem Abschnitt diskutiert werden. Die Kommission verweist insbesondere auf die Empfehlungen von IMI PREFER zur Planung, Durchführung, Analyse und Interpretation von Präferenzstudien, die entscheidungsrelevante Informationen z.B. zur Wahl der patientenrelevanten Endpunkte und zur Abwägung zwischen Endpunkten liefern sollen. Im Rahmen der Empfehlungen hat IMI PREFER (siehe The PREFER consortium (2022)) ein umfassendes Rahmenwerk vorgelegt, das eine positive Bewertung vom Committee for Medicinal Products for Human Use (CHMP) basierend auf einer EMA Methodenqualifizierung erhalten hat (siehe EMA (2022)). Diese Qualifizierung basiert auf gemeinsamen EMA/EUnetHTA Qualification Advice. Zudem empfiehlt die Kommission, dass sogenannte „Core Sets of Patient-reported Outcomes“ berücksichtigt werden, sofern diese für die entsprechende Erkrankung/Intervention publiziert wurden (siehe z.B. Stojkov et al (2021)).

Abschnitt 3.2.1 Umgang mit möglichem Reporting Bias

Im Abschnitt 3.2.1 wird der Umgang mit möglichem Reporting Bias beschrieben. Insbesondere wird hier auf zusätzliche Suchen in Registern klinischer Studien eingegangen. In diesem Zusammenhang möchte die Kommission auf neuere Entwicklungen in der statistischen Methodik hinweisen, die auch die Berechnung von Effektschätzern und Konfidenzintervallen unter Berücksichtigung von zusätzlichen Informationen aus solchen Registern berücksichtigen (siehe z.B. Huang et al (2021, 2022)).

Abschnitt 3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte)

Dieser Abschnitt ist neu hinzugekommen, um den Anforderungen der anwendungsbegleitenden Datenerhebungen gemäß § 35a Abs. 3b SGB V Rechnung zu tragen. Die „Beschreibung grundsätzlicher Anforderungen an die Auswertung der Daten“ verweist auf die im Rapid Report A19-43 beschriebenen Konzepte zur Auswertung versorgungsnaher Daten zum Zwecke der Nutzenbewertung von Arzneimitteln. Hier wird hauptsächlich auf „die adäquate Adjustierung für Störgrößen (Confounder)“ abgehoben. Aus Sicht der Kommission sollte an dieser Stelle darauf hingewiesen werden, dass beim Vorliegen von zeitabhängigem Confounding ggf. spezifische Kausalinferenz-Verfahren zur Kontrolle dieses Confounding-Typs erforderlich sind wie u.a. die parametrische g-Formula, marginal structural models mit inverse probability weighting, oder structural nested models mit g-estimation (Robins et al, 2004; Abschnitt „Modelling for explanation“ in Jahn et al, 2022). Aus Sicht der Kommission ist es ferner wichtig zu betonen, dass Confounding jedoch nicht die einzige Quelle von Bias beim Vergleich von Therapiegruppen in nicht-randomisierten Studien darstellen. Neben vielen weiteren möglichen

Aspekten ist eine wesentliche Verzerrungsquelle die Festlegung des Zeitpunkts „Null“, zu dem die Einschlusskriterien erfüllt sind, die Patienten den Therapiegruppen zugewiesen werden und die Nachbeobachtung zur Erfassung der Endpunkte beginnt. Dieser Zeitpunkt Null ist in randomisierten Studien in natürlicher Weise durch den Zeitpunkt der Randomisierung gegeben. In nicht-randomisierten Studien ist dies nicht der Fall und nicht adäquate Festlegungen können zu erheblichen Verzerrungen führen; als Stichworte seien hier *time-dependent bias* und *immortal time bias* genannt (siehe z.B. Hernán et al (2016)). Die Kommission empfiehlt daher, diese Aspekte, die u.a. zu einer geeigneten expliziten Nachbildung einer randomisierten Studie gehören wie die *target trial emulation* (Hernán und Robins, 2016; Hernán et al, 2022), die inzwischen auch von anderen HTA Organisationen in deren Methodenpapieren aufgenommen wurden (National Institute for Health and Care Excellence, 2022), stärker zu betonen.

Wie einleitend bemerkt, wünscht sich die Kommission auch Rapid Reports kommentieren zu können. Da im Abschnitt 3.3.4 des Entwurfs der Version 7.0 der Rapid Report A19-43 zitiert wird, nimmt die Kommission dies zum Anlass diesen zu kommentieren. Aus Sicht der Kommission werden die Vor- und Nachteile verschiedener Auswertungsverfahren in diesem in unausgewogener Weise diskutiert. So wird dort z.B. berichtet, dass der Nachteil von multiplen Regressionsanalysen darin besteht, dass „diese Analysen je nach Fallzahl bzw. Ereigniszahl nur eine begrenzte Anzahl an Confoundern aufnehmen können, da sie sonst unzuverlässige Resultate liefern oder mathematisch gar nicht funktionieren“, während „der große Vorteil der Verfahren, die auf Propensity Scores beruhen, ist, dass im Vergleich zur direkten Adjustierung deutlich mehr mögliche Confounder einbezogen werden können“. Das mag in einzelnen Fällen so richtig sein. Allerdings basiert auch der Propensity Score auf einem Regressionsmodell (häufig einer logistischen Regression) und somit ergeben sich dort ebenso Bedingungen für eine Konvergenz. Insbesondere bei kleinen Datensätzen sind die vermeintlichen Vorteile der Propensity Scores mitunter nicht gegeben (siehe Friedrich und Friede, 2020). Aus Sicht der Kommission ist eine ausgewogenere Diskussion, die sich an den Zielen der jeweiligen Auswertung orientiert, in der Abwägung der Auswertungsmethoden sinnvoller.

Aus Sicht der Kommission gelten die grundsätzlichen Erwägungen zur kausalen Inferenz im Zusammenhang mit anwendungsbegleitenden Datenerhebungen (AbD) auch für Vergleiche mit historischen Kontrollen.

Im Abschnitt 3.2.2 wird ein „dramatischer Effekt“ definiert. Dieser liegt vor, wenn der Effekt statistisch signifikant zum Niveau 1% ist und auf der Skala des relativen Risikos eine Größenordnung von 5 bis 10 erreicht. Bei AbDs liegt jedoch ein sogenannter dramatischer Effekt vor, wenn die untere Grenze des 95% Konfidenzintervalls oberhalb von 2 („z. B. in einem Bereich von 2 bis 5 für das relative Risiko (bzw. 0,2 bis 0,5 bei mortalitätssenkenden Interventionen) bezogen auf die untere bzw. obere Konfidenzintervallgrenze“, Abschnitt 3.3.4) liegt. Die Kommission empfiehlt, die Definitionen zu vereinheitlichen.

Abschnitt 3.5 Diagnostische Verfahren

In Abschnitt 3.5 wird darauf eingegangen, dass diagnostische Tests mit einem kontinuierlichen Ergebnis gemeinsam mit entsprechenden Grenzwerten angegeben werden sollen, die eine Einteilung in ein positives oder negatives Ergebnis erlauben. Aus Sicht der Kommission sind an dieser Stelle nicht nur diagnostische Grenzwerte relevant, sondern auch die vollständige Darstellung der Ergebnisse über eine Receiver-Operating-Characteristic-Kurve (ROC-Kurve) und die Berechnung der zugehörigen Area under the curve (AUC).

Die Kommission begrüßt es, dass im Methodenpapier im Abschnitt 3.5.1.1 nun explizit auf Test-treatment-Studien eingegangen wird (Pepe, 2003; Hot et al, 2021). Jedoch sind diese erst der abschließende Schritt bei der Entwicklung und Bewertung neuer diagnostischer Verfahren. Es wäre wünschenswert herauszustellen, dass diagnostische Verfahren bereits vorab in verschiedenen Studienphasen (siehe z.B. Table 8.1 in Pepe (2003)) hinsichtlich ihrer Genauigkeit in Form von beispielsweise Sensitivität und Spezifität geprüft werden müssen, bevor es zu Test-treatment-Studien kommt.

Im Abschnitt 3.5.1.2 zu „Linked Evidence“ findet sich die Abbildung 14, die für sich genommen nur sehr schwer oder gar nicht nachvollziehbar ist. Insbesondere ist nicht klar, wie sich einige Pfade verzweigen. Die Kommission empfiehlt, die Abbildung zu überarbeiten und evtl. auch in einer ausführlicheren Bildunterschrift oder Legende weitere Informationen zu geben, die ein einfacheres Verständnis erlauben. Zudem wird in Abschnitt 3.5.1.2 nicht explizit auf Screening verfahren eingegangen. Hier sollte aus Sicht der Kommission zumindest kurz angesprochen werden, dass ähnliche Methoden auch zur Nutzen-Schaden-Bewertung von Screeningverfahren eingesetzt werden können, sofern keine aktuelle direkte Evidenz vorliegt. Dies ist insbesondere der Fall, wenn für ein Screeningverfahren der Nutzen bereits als gesichert gilt und ein entsprechendes Screeningprogramm bereits etabliert ist, aber verschiedene Varianten (Startalter, Stopalter, Screeningintervall, Screeningtestkombinationen) des Screenings verglichen werden (siehe z.B. IQWiG (2022)).

Abschnitt 4 Kosten-Nutzen-Bewertung medizinischer Interventionen

Die Kommission begrüßt die Erweiterung im Abschnitt 4 um die tabellarische explizite Beschreibung des Referenzfalls (Tabelle 4), deren Format sich an internationalen Standards bzw. Lehrbücher anlehnt.

Im Abschnitt 4 wird an verschiedenen Stellen auf „Modellierungen“ eingegangen. Es wird vorgeschlagen, in Abschnitt 4 zu Beginn darzulegen, dass es sich beim Begriff „Modellierungen“ um entscheidungsanalytische Modelle handelt, um diese von statistischen Modellen abzugrenzen (z.B.: S. 84 erste Zeile wie folgt ergänzen: „...werden in der Regel Modellierungen *mit entscheidungsanalytischen Modellen* durchgeführt (Siebert, 2003). In Abschnitt 4.9 Gesundheitsökonomisches Studiendesign wäre es hilfreich, im Unterabschnitt ‚Auswahl der Modellierungstechnik‘ neben der Referenz [596] auch auf die internationalen Empfehlungen zu den genannten einzelnen Verfahren hinzuweisen: „Die Konzeption des Modells ist in der Regel ausschlaggebend dafür, welche Modellierungs-technik (Entscheidungsbaum, Markov-Modell, diskrete Ereignissimulation, etc.) zum Einsatz kommt ([596]; Siebert et al, 2012; Karnon et al, 2012; Pitman et al, 2012). Am Ende des einführenden Unterkapitels in Abschnitt 4.9 (S. 99) sollte darauf hingewiesen werden, dass Inputparameter für entscheidungsanalytische Modellierungen, die sich auf Interventionseffekte beziehen, mittels der entsprechenden Verfahren der Kausalinferenz abzuleiten sind (Kühne et al, 2022; Jahn et al, 2022).

Abschnitt 9.1 Qualitätsbewertung von Einzelstudien

Auf den Seiten 158-159 findet sich eine Liste von Stellungnahmen zur Beurteilung der Verwendbarkeit einer Quelle. Die Kommission empfiehlt, diese um die „**RE**porting of studies **C**onducted using **O**bservational **R**outinely-collected **D**ata“ (RECORD) Statements zu erweitern (siehe record-statement.org; Benchimol et al (2015); Langan et al (2018)).

Bei der Überarbeitung des Abschnitts 9.1.1 zu Kriterien für den Einschluss von Studien wurde nun die „80% Regel“ aufgenommen, die besagt, dass eine Studie eingeschlossen wird, „falls auf mindestens

80% der eingeschlossenen Patientinnen und Patienten das Kriterium (Population) zutrifft bzw. falls das Einschlusskriterium bezüglich der Prüfintervention (Interventionsgruppe der Studie) und der Vergleichsintervention (Vergleichsgruppe der Studie) jeweils zu mindestens 80% erfüllt ist“. Die Kommission begrüßt dies, da dies konsistent mit dem bisherigen Vorgehen des Instituts ist und in diesem Punkt nun Klarheit schafft.

Abschnitt 9.3 Spezielle biometrische Aspekte

Im Abschnitt 9.3.3 werden Kriterien und Verfahren zur Beurteilung der klinischen Relevanz beschrieben. Analog zum Methodenpapier 6.0 hat der G-BA einen Schwellenwert von 15% der Spannweite der Skala für PROs in die Verfahrensordnung aufgenommen. Dieser Schwellenwert gilt für alle PROs, für die keine MCID validiert ist oder für die die MCID unterhalb dieser Schwelle liegt. Auf der G-BA FAQ Seite findet sich seit kurzem der Hinweis, dass für Auswertungen des Fragebogens EORTC QLQ-C30 sowie den entsprechenden validierten krankheitsspezifischen Ergänzungsmodulen Analysen zum Responsekriterium 10 Punkte im Dossier darzustellen sind (www.g-ba.de/themen/arzneimittel/arzneimittel-richtlinie-anlagen/nutzenbewertung-35a/faqs, Zugriff: 16.1.2023). Dieses wird so bereits in aktuellen Bewertungen vom IQWiG aufgenommen (siehe z.B. die Bewertung zu Selinexor, Verfahren Nr. 886, IQWiG-Berichte Nr. 1480). Hier wird vom IQWiG kommentiert: "Für den EORTC-QLQ-C30 und seine Zusatzmodule sind im Verfahren der frühen Nutzenbewertung jedoch Auswertungen zu einer Responderschwelle von 10 Punkten erforderlich". Der Hersteller hatte nur Analysen mit einer Responderschwelle von 15 Punkten gemäß IQWiG Methodenpapier 6.1 und der Verfahrensordnung vorgelegt. Nach Ansicht der Kommission sollte diese Ausnahme im Methodenpapier 7.0 Berücksichtigung finden.

Im Abschnitt 9.3.6 werden Verfahren zur Adjustierung für Confounder im Rahmen von anwendungsbegleitenden Datenerhebungen beschrieben. Hierzu hat die Kommission bereits oben beziehend auf den Abschnitt 3.3.4 Anmerkungen gemacht.

Im Abschnitt 9.3.7 werden Verfahren für Metaanalysen dargestellt. Unter B) Heterogenität heißt es: „Zu diesem Zweck gibt es spezielle statistische Methoden wie z. B. das I^2 -Maß [355]. Für dieses Maß existieren Untersuchungen, die eine grobe Einschätzung der Heterogenität zulassen (z. B. über die Kategorien wahrscheinlich unbedeutend [0 bis 40 %], mittelmäßig [30 bis 60 %], substanziell [50 bis 90 %] und erheblich [75 bis 100 %] [174]).“ Die Kommission empfiehlt die Hinweise zum I^2 -Maß zu entfernen, da dessen Verwendung über die Jahre zunehmend in Kritik geraten ist. Die Einschätzung des Ausmaßes der Heterogenität kann alternativ über Schätzung der Standardabweichung tau erfolgen; auch für dieses Maß sind Kategorien des Ausmaßes in der Literatur bekannt (siehe z.B. Tabelle 1 in Friede et al (2017) oder Tabelle 2 in Röver et al (2021)).

Unter A) Allgemeines heißt es: „Kontextabhängig kommen auch alternative Verfahren zur metaanalytischen Zusammenfassung in Betracht, wie z. B. bayessche Verfahren“. Hier würde die Kommission begrüßen, wenn das Institut eine Empfehlung zur Wahl der a priori Verteilungen bei Bayes-Verfahren geben könnte. Insbesondere ist hier die a priori Verteilung für den Heterogenitätsparameter tau von Interesse (Röver et al, 2021).

In der Ergänzung zu Abschnitt 9.3.12 wird die Auswirkung systematisch verkürzter Beobachtungszeiten bei der Erhebung von Patient Reported Outcomes (PROs) diskutiert. Dabei wird auf die Operationalisierung der sogenannten "dauerhaften" Veränderung eingegangen. Bei systematisch verkürzten Beobachtungszeiten in einer der beiden Behandlungsgruppen kann es für die Zeit bis zur Verschlechterung zu einer Verzerrung zu Gunsten der Behandlungsgruppe mit der längeren Beobachtungszeit führen, da die Patienten länger die Chance haben, wieder eine Verbesserung zu

erreichen. Im Falle der Zeit bis zur Verbesserung ist von einer Verzerrung zu Gunsten der Behandlungsgruppe mit der kürzeren Beobachtungszeit zu erwarten, da die Patienten kein Risiko haben, dass nach der Verbesserung eine Verschlechterung festgestellt wird.

Die Kommission rät von der Durchführung derartiger vereinfachter Ereigniszeitanalysen bei komplexen Zeitverläufen ab, da sie auf die Zukunft bedingen, was grundsätzlich unterbleiben sollte (siehe etwa Andersen und Keiding, 2012). Wenn man Therapiegruppen bzgl. ihres zeitlichen Verlaufs in bestimmten verschlechterten oder verbesserten Zuständen vergleichen möchte, sollte man nach Ansicht der Kommission geeignete Mehrstadienmodelle verwenden. Diese würden die Zeiten unter Risiko adäquat berücksichtigen, und die in Abschnitt 9.3.12 erwähnten Probleme bei unterschiedlich langen Beobachtungszeiten würden somit nicht bestehen. Ein Beispiel für solche Analysen (in einem anderen inhaltlichen Kontext, aber genauso auf PRO-Zustände übertragbar) findet sich in Szenario 4 in Schmoor et al (2013) (Referenz [638] im Methodenpapier Version 7.0). Inferenzstatistische Vergleiche von Therapiegruppen, wenn notwendig, sollten nach Ansicht der Kommission entweder bezüglich der *transition hazards* oder der *probabilities* über die Zeit durchgeführt werden; siehe hierzu z.B. Bluhmki et al (2018, 2020).

Es sollte im gleichen Zuge auch die Diskussion zur sogenannten "bestätigten" Veränderung aufgenommen werden, wie es auch auf der FAQ-Seite des G-BA beschrieben ist (www.g-ba.de/themen/arzneimittel/arzneimittel-richtlinie-anlagen/nutzenbewertung-35a/faqs). Die "bestätigte" Veränderung ist dabei eine Veränderung, bei der der Schwellenwert der Response für zwei oder mehrere aufeinanderfolgende Erhebungen überschritten wird. Hierbei ist zu berücksichtigen, dass eine erstmalige Veränderung zum letzten Beobachtungszeitpunkt nicht als "bestätigte" Veränderung gewertet werden sollte, um eine Verzerrung zugunsten der Behandlungsgruppe mit der längeren Beobachtungszeit zu vermeiden.

Referenzen

Andersen PK, Keiding N (2012) Interpretability and importance of functionals in competing risks and multistate models. *Statistics in Medicine* 31: 1074 – 1088.

Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, RECORD Working Committee (2015) The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Medicine*. DOI:10.1371/journal.pmed.1001885

Bluhmki T, Schmoor C, Dobler D, Pauly M, Finke J, Schumacher M, Beyersmann J (2018) A wild bootstrap approach for the Aalen–Johansen estimator. *Biometrics* 74: 977–985.

Bluhmki T, Schmoor C, Finke J, Schumacher M, Beyersmann J (2020) Relapse- and immunosuppression-free survival after hematopoietic stem-cell transplantation: How can we assess treatment success for complex time-to-event endpoints? *Biology of Blood and Marrow Transplantation* 26: 992-997.

EMA (2022) Qualification Opinion of IMI PREFER. https://www.ema.europa.eu/documents/regulatory-procedural-guideline/qualification-opinion-imi-prefer_en.pdf (Zugriff: 18. Januar 2023).

Friede T, Röver C, Wandel S, Neuenschwander B (2017) Meta-analysis of few small studies in orphan diseases. *Research Synthesis Methods* 8: 79–91.

Friedrich S, Friede T (2020) Causal inference methods for small non-randomized studies: Methods and an application in COVID-19. *Contemporary Clinical Trials* 99: 106213.

Hernán MA, Robins JM (2016) Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology* 183(8): 758-764.

Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I (2016) Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *Journal of Clinical Epidemiology* 79: 70–75.

Hernán MA, Wang W, Leaf DE (2022) Target Trial Emulation: A framework for causal inference from observational data. *JAMA* 328: 2446–2447.

Hot A, Bossuyt PM, Gerke O, Wahl S, Vach W, Zapf A (2021) Randomized test-Treatment studies with an outlook on adaptive designs. *BMC Medical Research Methodology* 21(1): 110.

Huang A, Komukai S, Friede T, Hattori S (2021) Using clinical trial registries to inform Copas selection model for publication bias in meta-analysis. *Research Synthesis Methods* 12: 658–673.

Huang A, Morikawa K, Friede T, Hattori S (2022) Adjusting for publication bias in meta-analysis via inverse probability weighting using clinical trial registries. *Biometrics* (in press). <https://doi.org/10.1111/biom.13822>

IQWiG. Überprüfung der Altersgrenzen im Mammografie-Screening-Programmen. IQWiG Abschlussbericht Auftrag S21-01, Version 1.1, Stand: 16.08.2022. IQWiG-Berichte – Nr. 1383. Internet: https://www.iqwig.de/download/s21-01_altersgrenzen-im-mammografie-screening-programm_abschlussbericht_v1-1.pdf

Jahn B, Friedrich S, Behnke J, Engel J, Garczarek U, Münnich R, Pauly M, Wilhelm A, Wolkenhauer O, Zwick M, Siebert U, Friede T (2022) On the role of data, statistics and decisions in a pandemic. *Advances in Statistical Analysis* 106(3): 349-382.

Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J (2012) Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Medical Decision Making* 32(5):701-11.

Kühne F, Schomaker M, Stojkov I, Jahn B, Conrads-Frank A, Siebert S, Sroczynski G, Puntcher S, Schmid D, Schnell-Inderst P, Siebert U (2022) Causal Evidence in Health Decision Making: Methodological Approaches of Causal Inference and Health Decision Science. *HTA Report Vol. 509-1. GMS German Medical Science — an Interdisciplinary Journal* 20:Doc12. DOI: 10.3205/000314.

Langan SM,¹ Schmidt SAJ, Wing K, Ehrenstein V, Nicholls SG, Fillion KB, Klungel O, Petersen I, Sorensen HT, Dixon WG, Guttmann A, Harron K, Hemkens LG, Moher D, Schneeweiss S, Smeeth L, Sturkenboom M, von Elm E, Wang SV, Benchimol EI (2018) The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ* 363:k3532. DOI:10.1136/bmj.k3532

National Institute for Health and Care Excellence (2022) NICE real-world evidence framework. Corporate document [ECD9]. Published: 23 June 2022. Internet: <https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837>

Pepe MS (2003) The statistical evaluation of medical tests for classification and prediction. Oxford University Press, USA, Kapitel 8.

Pitman R, Fisman D, Zaric GS, Postma M, Kretschmar M, Edmunds J, Brisson M; ISPOR-SMDM Modeling Good Research Practices Task Force (2012) Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-5. *Medical Decision Making* 32(5): 712-21.

Robins JM, Hernán MA, Siebert U (2004) Effects of Multiple Interventions. In: Ezzati M, Lopez AD, Rodgers A, Murray CJL (eds.). *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Vol. 1. World Health Organization: Geneva, chapter 28, pages 2191-2230. Volltext: https://apps.who.int/iris/bitstream/handle/10665/42770/9241580313_eng.pdf?sequence=1&isAllo wed=y

Röver C, Bender R, Dias S, Schmid C, Schmidli H, Sturtz S, Weber S, Friede T (2021) On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods* 12: 448–474.

Siebert U (2003) When should decision-analytic modeling be used in the economic evaluation of health care? *The European Journal of Health Economics* 4(3): 143-50.

Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, Kuntz KM (2012) State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Medical Decision Making* 32(5): 690-700.

Stojkov I, Conrads-Frank A, Rochau U, Koinig KA, Arvandi M, Puntcher S, van Marrewijk C, Fenaux P, Symeonidis A, Chermat F, Garelius H, Bowen D, Mittelman M, Mora E, de Witte T, Efficace F, Siebert U, Stauder R (2021) Core set of patient-reported outcomes for myelodysplastic syndromes: an EUMDS Delphi study involving patients and hematologists. *Blood Advances* 6(1): 1-12. DOI 10.1182/bloodadvances.2021004568

The PREFER consortium (2022) PREFER Recommendations: Why, when and how to assess and use patient preferences in medical product decision-making. Zenodo. DOI: 10.5281/zenodo.6470922

Über die GMDS e.V.

Die Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. ist mit derzeit ca. 2.000 Mitgliedern die einzige wissenschaftliche Fachgesellschaft in Deutschland, Österreich und der Schweiz, die die fünf Disziplinen Medizinische Informatik, Medizinische Biometrie, Epidemiologie, Medizinische Dokumentation und Medizinische Bioinformatik und Systembiologie gemeinsam vertritt. Sie kooperiert mit einer Reihe benachbarter Fachgesellschaften und Verbänden. Zudem entwickelt sie die Fachgebiete weiter durch sachverständige Repräsentation u. a. bei der Planung von Förderungsmaßnahmen der Öffentlichen Hand, bei Fragen der Standardisierung und Normung, bei der Errichtung von Lehrinstitutionen, bei Ausbildungs-, Weiter- und Fortbildungsfragen und bei gesetzgebenden Maßnahmen.

Deutsche Gesellschaft für Medizinische Informatik,
Biometrie und Epidemiologie (GMDS) e.V.

██████████

██████████

www.gmds.de

Über die IBS-DR

Die Deutsche Region der Internationalen Biometrischen Gesellschaft (IBS-DR) umfasst die in der Bundesrepublik Deutschland ansässigen Mitglieder der Internationalen Biometrischen Gesellschaft. Die Internationale Biometrischen Gesellschaft fördert die Entwicklung von statistischer und mathematischer Theorie und die Anwendung ihrer Methoden in den Biowissenschaften. Anwendungsgebiete sind unter anderem biomedizinische Wissenschaften und Öffentliches Gesundheitswesen, Ökologie, Umweltwissenschaft, Forstwissenschaft und Landwirtschaft.

Deutsche Region der Internationalen Biometrischen Gesellschaft
c/o Institut für Biometrie, Epidemiologie und Informationsverarbeitung
Tierärztliche Hochschule Hannover

██████████

██████████

<http://biometrische-gesellschaft.de/>

A.1.22 Ecker + Ecker GmbH

Autorinnen und Autoren

- Ecker, Christof
- Ecker, Thomas
- Pütz, Claudia
- Beuermann, Janik
- Musiolik, Katharina
- Aranyossy, Tim
- Beenk, Katharina
- Groneberg, Marie

Stellungnahme der Ecker + Ecker GmbH zum Entwurf der Allgemeinen Methoden, Version 7.0 vom 06.12.22

Autoren:

Dr. Christof Ecker

Dr. Thomas Ecker

Dr. Claudia Pütz

Janik Beuermann

Katharina Musiolik

Dr. Tim Aranyossy

Dr. Katharina Beenk

Dr. Marie Groneberg

Einleitung

Als Fachberatung bietet die Ecker + Ecker GmbH seit mehr als 10 Jahren Lösungen für den Marktzugang von Arzneimitteln, Diagnostika und Medizinprodukten in Deutschland. In diesem Kontext unterstützt die Ecker + Ecker GmbH regelmäßig pharmazeutische Unternehmen zu verschiedenen Fragestellungen im Rahmen der frühen Nutzenbewertung.

Grundsätzlich werden die Überarbeitungen des Methodenpapiers im Rahmen der Version 7.0 begrüßt, da die Methodik an vielen Stellen geschärft und die Rahmenbedingungen damit klarer definiert bzw. ausformuliert werden. Im Folgenden sollen jedoch einige spezifische Anpassungen kommentiert werden:

Kompetenzbereich Biostatistik

1. Methodik zur Identifikation von Confoundern
2. Nicht-randomisierte Studien in der frühen Nutzenbewertung
3. Umgang mit variablen Beobachtungszeiten
4. Relevanzschwelle von 15 % der Skalenspannweite
5. Ermittlung der Effektstärke bei stetigen Zielgrößen

Kompetenzbereich Gesundheitsökonomie

6. Kostendarstellung
7. Relevanz des Ergebnisses der Kosten-Nutzen-Bewertung nach § 35b SGB V für die Verhandlung des Erstattungsbetrags im Fall eines nicht nachgewiesenen Zusatznutzens in der Nutzenbewertung gemäß § 35a SGB V

Kompetenzbereich Biostatistik

1 Methodik zur Identifikation von Confoundern

Bezug: Abschnitt 9.3.6 der Allgemeinen Methoden, Entwurf der Version 7.0

Problembeschreibung:

In Abschnitt 9.3.6 beschreibt das IQWiG Methoden der Adjustierung bei nicht-randomisierten Studiendaten. Insbesondere wird beschrieben, dass im Rahmen der anwendungsbegleitenden Datenerhebung (abD) die systematische Identifikation von Confoundern von großer Bedeutung sei. Als gewünschtes Vorgehen wird hier die in der Arbeit von Pufulete et al.¹ beschriebene Methodik zur Identifikation von Confoundern genannt. In dieser Arbeit werden Confounder in einem Zwei-Schritt-Verfahren, bestehend aus systematischer Literaturrecherche und Expertenbefragungen identifiziert. Auch wenn dieses Vorgehen der Kombination von Literatur und klinischen Wissen prinzipiell als sinnvoll zu beurteilen ist, so ist die explizite Ausgestaltung als problematisch einzuordnen. Insbesondere betrifft dies die im Folgenden weiter ausgeführten Aspekte hinsichtlich:

- Identifikation von Confoundern
- Durchführung der systematischer Literaturrecherche
- Expertenbefragungen hinsichtlich relevanter Confounder

Identifikation von Confoundern

Nach Methodik von Pufulete et al. werden RCTs und Beobachtungsstudien (retrospektiv/prospektiv) im relevanten Indikationsgebiet in die Literaturrecherche eingeschlossen, was grundsätzlich richtig und zu begrüßen ist. Als Confounder werden bei RCTs dabei alle in den Baseline-Charakteristika aufgeführte Größen interpretiert. Ob diese Größen allesamt einen Einfluss auf Therapiewahl und/oder Therapieerfolg haben, ist fragwürdig, da dies in der Regel nicht der alleinige Entscheidungsgrund zur Darstellung von Baseline-Charakteristika sein muss. Dies birgt die Gefahr, dass eine Vielzahl von Größen fälschlicherweise als Confounder identifiziert werden und die spätere Berücksichtigung im statistischen Modell zur verzerrten Schätzung des Behandlungseffekts führt.

Abweichend zu den RCT werden nach Pufulete et al. bei Beobachtungsstudien alle in die statistische Analyse eingeschlossene Größen als Confounder interpretiert – unabhängig der statistischen Signifikanz bzw. klinischen Relevanz. Ob diese Größen tatsächlich Confounder darstellen ist mit Unsicherheiten behaftet, da die Autoren nicht-randomisierter Studien regelhaft keine Rationale zur statistischen Adjustierung liefern. Dies benennen auch Pufulete et al. explizit als Limitation ihres eigenen Vorgehens. Hinzu kommt, dass nur solche Größen zur statistischen Adjustierung herangezogen werden können, welche regelhaft in Studien mit hinreichender Datenqualität erhoben wurden.

Beim vorgeschlagenen Vorgehen zur Identifizierung von Confoundern in RCTs als auch in Beobachtungsstudien ist weiterhin anzumerken, dass relevante prognostischen Größen, welche bereits als Ein- und Ausschlusskriterien für die Studie gewählt wurden, nicht mehr als Confounder benannt werden können.

Ein wichtiger Bestandteil einer Confounderrecherche sollte auch die Diskussion von möglichen Operationalisierungen der Confounder (z. B. kategorische vs. stetige Darstellung des Alters) sein, um den Vorwurf einer ergebnisgesteuerten Definition der Confounder und der damit verringerten Ergebnissicherheit entgegenzuwirken. Hierzu findet sich in der Arbeit von Pufulete et al. keinerlei Ansatz.

¹ Pufulete M, Mahadevan K, Johnson TW et al. Confounders and co-interventions identified in non-randomized studies of interventions. J Clin Epidemiol 2022; 148(148): 115-123. <https://dx.doi.org/10.1016/j.jclinepi.2022.03.018>.

Durchführung der systematischen Literaturrecherche

Einen Kritikpunkt an der Methodik von Pufulete et al. stellt der uneingeschränkte Suchzeitraum der Literaturrecherche dar. So kann es insbesondere bei Indikationsgebieten mit schnell voranschreitendem Kenntnisstand zu Confoundern der Fall sein, dass Größen in älteren Publikationen zu finden sind, deren prognostischer Wert aus heutiger Sicht nicht mehr als gegeben anzunehmen ist.

Zudem resultiert diese sehr breit angelegte Suchstrategie in einer hohen Anzahl von Treffern und damit verbundenem, großem Aufwand bei der Auswertung der Ergebnisse. Als Antwort darauf entschieden sich die Autoren von Pufulete et al., nur eine Zufallsstichprobe aller relevanten Volltexte (70 von 322) zu screenen und die Suche gemäß des „Saturation criterion“ nach 10 Treffern ohne neu identifizierten Confounder zu beenden. Dies macht die Recherche zwar praktikabel, aber die Suche wird nun an zwei Stellen (Auswahl der Stichprobe und Reihenfolge der Treffer) durch Zufall beeinflusst. Dies erschwert die Replizierbarkeit der Auswertung und kann ggf. sogar zu selektiven Ergebnissen führen.

Unklar bleibt zudem, mit wie viel zeitlichem Vorlauf eine Confounderrecherche durchgeführt werden kann, bevor die Ergebnisse z. B. zur Adjustierung von nicht-randomisierten Daten im Rahmen der abD genutzt werden. Eine Differenz von sechs Jahren zwischen Recherche und Veröffentlichung der Resultate wie bei Pufulete et al. scheint als nicht mehr sachgerecht und wird folgerichtig auch als Limitation ihres Vorgehens von den Autoren aufgeführt.

Expertenbefragungen hinsichtlich relevanter Confounder

In der Arbeit von Pufulete et al. werden klinische Experten auf zwei unterschiedlichen Wegen zu Confoundern befragt. So werden zum einen (telefonische) Interviews und zum anderen Online-Befragungen mit einer deutlich größeren Anzahl Experten durchgeführt. Inwiefern wirklich beide Maßnahmen für eine adäquate Confounderrecherche notwendig sind, ist zweifelhaft. Da bereits die systematische Literaturrecherche sehr umfassende Resultate liefern kann und die Expertenbefragungen dann lediglich ergänzenden Charakter haben, scheint eine Expertenbefragung beispielsweise nur durch Interviews als hinreichend. Außerdem liefern gemäß Pufulete et al. Expertenbefragungen häufig Confounder, welche in der Kategorie „hard-to-measure“ einzuordnen sind, weil sie von Patient zu Patient unterschiedlich sind und stark auf persönlichen Erfahrungen/Meinungen der klinischen Experten beruhen. Es erscheint daher zielführender, die Expertenbefragungen eher einfach zu gestalten und den Fokus stattdessen auf die systematischen Ergebnisse der Literaturrecherche zu legen.

Handlungsbedarf:

Der aktuelle Entwurf des IQWiG empfiehlt eine Methodik zur Identifizierung von Confoundern, welche zahlreiche Limitationen aufweist. Außerdem verbleiben Unsicherheiten, ob auf diese Art und Weise alle relevanten Confounder identifiziert werden können.

Lösungsvorschlag:

Im Rahmen dieser Stellungnahme wird vorgeschlagen, dass das IQWiG im Rahmen des Methodenpapiers klar benennt, wie mit den offensichtlichen Limitationen der vorgeschlagenen Methodik von Pufulete et al. umzugehen ist. Alternativ könnte vom Institut eine Methodik entwickelt bzw. vorgeschlagen werden, in der den aktuell vorhandenen Widersprüche und Limitationen Rechnung getragen wird. Ziel sollte die Definition einer eindeutigen und praktikablen Methodik sein, bei der die Ausgestaltung der systematischen Literaturrecherche im Vordergrund steht, so dass Confounder unabhängig von deren Abbildung in Studien identifiziert werden können und deren prognostische Aussagekraft als sehr sicher einzuschätzen ist. Hierfür könnte beispielsweise eine

Recherche nach Confoundern in Leitlinien und Metaanalysen bzw. systematischen Übersichten geeignet sein. Außerdem sollte eine zeitliche Einschränkung der Suche und die Operationalisierung der Confounder diskutiert werden. Grundsätzlich muss bei Entwicklung einer solche Methodik die praktische Anwendbarkeit auch in Anwendungsgebieten mit schnellem Wissensgewinn und hohen Trefferzahlen berücksichtigt werden.

2 Nicht-randomisierte Studien in der frühen Nutzenbewertung

Bezug: Abschnitt 9.3.6 und Abschnitt 9.3.8 der Allgemeinen Methoden, Entwurf der Version 7.0

Problembeschreibung:

In Abschnitt 9.3.6 erläutert das IQWiG das Vorgehen zur gemeinsamen Analyse von zwei nicht-randomisierten Kohorten im Rahmen der abD. Diese Analyse besteht grob gesagt aus einer Confounderrecherche, der Erstellung eines SAPs, der Durchführung einer statistischen Methodik (be ruhend beispielsweise auf Propensity Score Methoden) und der Überprüfung der Annahmen der gewählten Methodik. Eine vergleichbare Methodik zur Auswertung von nicht-randomisierten Studiendaten mittels eines unadjustierten indirekten Vergleichs wird vom Institut in Abschnitt 9.3.8 diskutiert. In diesem Abschnitt 9.3.8. wird diese Methodik jedoch für sämtliche Szenarien abgelehnt und darauf verwiesen, dass nur indirekte Vergleiche über einen Brückenkomparator akzeptiert werden.

Es ist nicht nachvollziehbar, warum solch ein Vorgehen im Kontext der anwendungsbegleitenden Datenerhebung akzeptabel erscheint, die gleiche Methode aber im in anderen Szenarien der frühen Nutzenbewertung kategorisch abgelehnt wird.

Handlungsbedarf:

Im aktuellen Entwurf des Methodenpapier widerspricht Abschnitt 9.3.8 mit seiner vollständigen Ablehnung von indirekten Vergleichen ohne Brückenkomparator dem Abschnitt 9.3.6, in dem erläutert wird, inwiefern Propensity Score Methoden im Rahmen der abD in der frühen Nutzenbewertung Anwendung finden können. Außerdem bleibt unklar, warum solche Methoden nicht außerhalb einer abD in der frühen Nutzenbewertung eingesetzt werden sollten.

Lösungsvorschlag:

Es wird geraten den Abschnitt 9.3.8 zu überarbeiten und zu erläutern, in welchen Szenarien auch indirekte Vergleiche ohne Brückenkomparator eingesetzt werden können. Außerdem sollte in Abschnitt 9.3.6. darauf verzichtet werden, die dort beschriebenen Methodiken zur Analyse von nicht randomisierten Daten lediglich auf die abD einzuschränken.

3 Umgang mit variablen Beobachtungszeiten

Bezug: Abschnitt 9.3.12 der Allgemeinen Methoden, Entwurf der Version 7.0

Problembeschreibung:

In Abschnitt 9.3.12 wird der Umgang mit variablen Beobachtungszeiten diskutiert. Dabei wird festgestellt, dass insbesondere patientenberichtete Endpunkte in onkologischen Studien oft nur bis zur Krankheitsprogression bzw. Abbruch der Behandlung untersucht/beobachtet werden, wodurch sich die mittleren Beobachtungsdauern in den Behandlungsarmen unterscheiden können. Das IQWiG gibt in seinen Methoden vor, dass alle Daten zu allen bewertungsrelevanten Endpunkten inklusive PROs und unerwünschten Ereignissen vollständig auch nach einem Behandlungsende oder Behandlungswechsel erfasst werden sollen um Verzerrungen zu vermeiden und die Interpretierbarkeit zu gewährleisten.

Handlungsbedarf:

Bei unterschiedlichen mittleren Beobachtungszeiten scheint die Forderung nach vollständiger Datenerfassung aller patientenrelevanten Endpunkte aus statistischer Sicht sinnvoll. Es müsste jedoch darauf hingewiesen werden, dass auch immer die klinische Realität zu berücksichtigen ist. Ein häufiges Problem ist, dass eine Krankheitsprogression bzw. Behandlungsabbruch im Rahmen onkologischer Studien oft die Behandlung im Rahmen einer anderen klinischen Studie und/oder deutliche Einschnitte sowie Veränderungen der Prioritäten im Leben des Patienten zur Folge hat. In der Realität gestaltet sich die, aus methodischer Sicht berechnete, Forderung zur weiteren Erhebung von Daten, insbesondere Fragebögen in entsprechender Qualität und mit hohen Rücklaufquoten entsprechend schwierig.

Lösungsvorschlag:

Im Rahmen dieser Stellungnahme wird vorgeschlagen, Abschnitt 9.3.12 um einen Absatz zu ergänzen, der erläutert, dass eine vollständige Datenerfassung aller patientenrelevanten Endpunkte inklusive unerwünschter Ereignisse und PROs immer im Kontext der klinischen Realität erfolgen sollte. Der alleinige Fokus auf den methodisch gewünschten, in der Realität aber kaum erreichbaren, Idealzustand erscheint nicht angemessen.

4 Relevanzschwelle von 15 % der Skalenspannweite

Bezug: Abschnitt 9.3.3 der Allgemeinen Methoden, Entwurf der Version 7.0, Unterkapitel „Relevanzbewertung bei Skalen“

Problembeschreibung:

Bei Aktualisierung der Allgemeinen Methoden zu Version 6.0 im November 2020 wurde vom Institut ein allgemeingültiger Schwellenwert von 15 % der Skalenspannweite aufgenommen, der die bisher genutzten, instrumentenindividuellen Minimal Important Differences (MID) bei Responderanalysen ersetzt hat. Im Entwurf 7.0 der allgemeinen Methoden findet in diesem Unterkapitel keine Aktualisierung gegenüber Version 6.1 statt. Dies ist vor allem aus folgenden zwei Gründen bedauernd:

Die systematisch recherchierten Übersichtsarbeiten zu MIDs, die als Grundlage der 15-%-Schwelle dienen, stammen aus den Jahren 2016 bis 2019, die Suche selbst aus dem Jahr 2020. Es wäre wünschenswert an dieser Stelle eine Aktualisierung der Recherche vorzunehmen und die Konsistenz des Schwellenwerts mit den aktuellsten Publikationen zu validieren.

Im Entwurf der 7.0. Allgemeinen Methoden findet sich die folgende Aussage: „Ein erster Vorschlag für ein Instrument zur Qualitätsbewertung von Studien zur Ermittlung von MIDs und zur Aussagekraft der MIDs wurde im Jahr 2020 veröffentlicht [190]. Umfassende Ergebnisse zur Anwendbarkeit stehen noch aus.“². Hier wäre ebenfalls eine Aktualisierung wünschenswert um zu erfahren in den vergangenen Jahren IQWiG in dieser Hinsicht weitere Erkenntnisse gewinnen konnte um die Methodik zu validieren.

Handlungsbedarf:

Die vom IQWiG definierte Methodik entspricht nicht den Ansprüchen der evidenzbasierten Medizin. Zudem wäre es angebracht, eine Aktualisierung und Evaluierung der entsprechenden Passagen vorzunehmen.

Lösungsvorschlag:

An den genannten Stellen sollte im Rahmen der Allgemeinen Methoden 7.0 eine Aktualisierung vorgenommen werden, um die Validität der Methodik mit neuen wissenschaftlichen Erkenntnissen zu untermauern oder entsprechend anzupassen.

² Seite 172 der Allgemeinen Methoden, Entwurf der Version 7.0

5 Ermittlung der Effektstärke bei stetigen Zielgrößen

Bezug: Abschnitte 3.3.3 und 9.3.3 der Allgemeinen Methoden, Entwurf der Version 7.0

Problembeschreibung:

Im Rahmen der Nutzenbewertung kann das Ausmaß des Zusatznutzens im Kontext von stetigen Zielgrößen patientenrelevanter Endpunkte entweder über Responderanalysen oder auf Basis von standardisierte Mittelwertdifferenzen bewertet werden. Als Grundlage kontinuierlicher Auswertungen werden dabei Auswertungen mittels Mixed Model for Repeated Measures (MMRM) gefordert, die „über den Studienverlauf ermittelten Informationen vollständig berücksichtig[en]“³.

Laut der Allgemeinen Methoden⁴ soll bei kontinuierlichen Auswertungen für die Relevanzbewertung dabei auf die standardisierten Mittelwertdifferenzen in Form von Hedges' g zurückgegriffen werden. Die Berechnung der für Hedges' g benötigten, gepoolten Standardabweichung aus einem MMRM ist jedoch nicht ohne weiteres möglich, auch nach Aussage des IQWiG liegt derzeit keine adäquate Methode der Berechnung vor⁵. Im Rahmen der Allgemeinen Methoden wird dieses Thema leider nicht aufgegriffen, so dass hier weiterhin Unsicherheiten bezüglich des adäquaten Vorgehens bestehen.

Handlungsbedarf:

Es sollte klargestellt werden, mit welcher Methodik ein Hedges' g auf Basis eines MMRM ermittelt werden soll und ob ein Hedges' g in jedem Fall bei der Analyse kontinuierlicher Endpunkte anzugeben ist, insbesondere wenn dies nicht sinnvoll ist.

Lösungsvorschlag:

Im Zuge dieser Stellungnahme wird vorgeschlagen, dass das IQWiG im Rahmen des Methodenpapiers klar benennt, in welchen Situationen die Berechnung und Darstellung eines Hedges' g notwendig ist. Zudem sollte beschrieben werden, welche Methodik zur Schätzung der gepoolten Standardabweichung bei einer Analyse mittels MMRM angewendet werden soll.

³ Vorlage des G-BA zu Modul 4, https://www.g-ba.de/downloads/17-98-4825/2021-12-16_Anl2_6_Modul4.pdf, Seite 32

⁴ Unterkapitel „Relevanzbewertung bei Skalen“, Seite 173

⁵ Mitarbeiter des IQWiG stellten im Rahmen einer Arbeit beim biometrischen Kolloquium 2021 fest, dass „[d]ie Schätzung einer SMD mittels der Standardfehler aus dem MMRM [...] nicht angemessen [ist].. Stattdessen seien „Verfahren notwendig, aus denen die gepoolte Standardabweichung der Änderung zum Ausgangswert mit [individuellen Patientendaten] geschätzt werden kann“; Vortrag des IQWiG mit dem Titel „Standardisierte Mittelwertdifferenzen aus Mixed Model Repeated Measures – Analysen“; <https://www.biometrisches-kolloquium2021.de/sessions/evidence-based-medicine-and-meta-analysis-ii/>, Zugriff 11.01.2023

Kompetenzbereich Gesundheitsökonomie

6 Kostendarstellung

Bezug: Abschnitt 3.3.3 der Allgemeinen Methoden, Entwurf der Version 7.0

Problembeschreibung:

In diesem Abschnitt erläutert das IQWiG, dass im Rahmen der Nutzenbewertung eines Arzneimittels gemäß § 35a SGB V der pharmazeutische Unternehmer u. a. Angaben zu den Kosten der Therapie für die gesetzliche Krankenversicherung machen muss. In Abschnitt 3.3.3 finden sich jedoch keine methodischen Anmerkungen, wie das Institut die Angaben zu den Kosten der Therapie für die gesetzliche Krankenversicherung überprüft. Dies wurde bereits in den Stellungnahmeverfahren zu Allgemeine Methoden, Entwurf für die Version 4.2 vom 18.06.2014, Entwurf für die Version 5.0 vom 07.12.2016, Entwurf für die Version 6.0 vom 05.12.2019 und in dem Entwurf für die Version 6.1 vom 23.08.2021 angeführt.

Handlungsbedarf:

Die Bewertung der Angaben zu den Kosten der Therapie für die gesetzliche Krankenversicherung durch das IQWiG weist Inkonsistenzen auf:

- Kostendarstellungen in vergleichbaren therapeutischen Situationen werden ungleich bewertet.
- Die Bewertungen durch das IQWiG weichen nicht unerheblich vom G-BA in den Beschlüssen ab.

Diese Inkonsistenzen sind wissenschaftlich nicht nachvollziehbar.

Lösungsvorschlag:

Im Rahmen dieser Stellungnahme wird die Einführung eines neuen Abschnitts „3.4 Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V“ vorgeschlagen. In diesem Abschnitt sollte ein Unterabschnitt „3.4.5 Kosten“ eingeführt werden. In diesem Abschnitt wäre u. a. zu beschreiben,

- welche Packungen für eine Kostenkalkulation herangezogen werden,
- welche Abzüge vom Apothekenverkaufspreis abgezogen werden,
- wann Kostenspannen angesetzt werden,
- welche Kostenpositionen saldiert werden,
- inwieweit Companion Diagnostics als zusätzliche GKV-Leistung berücksichtigt werden,
- welche Sonderregeln für die Kostendarstellung bei Orphan Drugs bestehen.

7 Relevanz des Ergebnisses der Kosten-Nutzen-Bewertung nach § 35b SGB V für die Verhandlung des Erstattungsbetrags im Fall eines nicht nachgewiesenen Zusatznutzens in der Nutzenbewertung gemäß § 35a SGB V

Bezug: Abschnitt 4 der Allgemeinen Methoden, Entwurf der Version 7.0

Problembeschreibung:

Das in seiner wissenschaftlichen Arbeit unabhängige IQWiG beschreibt in den Allgemeinen Methoden seine grundlegende Arbeitsweise. Entsprechend liefert das IQWiG in Abschnitt 4 der Allgemeinen Methoden eine Beschreibung seines Vorgehens bei Kosten-Nutzen-Bewertungen (KNB) von medizinischen Interventionen. Sinn und Zweck der Allgemeinen Methoden ist es dabei, transparent und detailliert darzustellen, welchen wissenschaftlichen Grundsätzen das IQWiG folgt, wie die Abfolge der einzelnen Schritte beim Erstellen der Arbeitsergebnisse aussehen und wie die Ergebnisse veröffentlicht werden.

In Unterabschnitt 4.5 bewertet das IQWiG jedoch darüber hinaus die Relevanz des Ergebnisses der KNB gegenüber dem Ergebnis der Nutzenbewertung gemäß § 35a SGB V: Das IQWiG beschreibt, dass das Ergebnis der KNB das Ergebnis der vorangegangenen Nutzenbewertung gemäß § 35a SGB V nicht infrage stellt und führt konkretisierend aus, dass in Situationen, in denen im Rahmen des Verfahrens der Nutzenbewertung kein Zusatznutzen nachgewiesen wurde, durch die KNB auch kein Zusatznutzen suggeriert wird. Diese wertenden Ausführungen obliegen dem IQWiG nicht und greifen der Kompetenz des G-BA insbesondere vor dem Hintergrund von § 35b (3) Satz 3 SGB V vor.

Der Gesetzgeber hat für den Fall, dass der G-BA in seinem Beschluss nach § 35a (3) oder § 35a (4) SGB V keinen Zusatznutzen oder keine therapeutische Verbesserung feststellt, nach § 35a (5a) SGB V für den pharmazeutischen Unternehmer (pU) bewusst die Möglichkeit der KNB eröffnet. Auf Grundlage des Ergebnisses der KNB ist der Erstattungsbetrag nach § 130b (8) SGB V neu zu vereinbaren.

Die Neuverhandlung des Erstattungsbetrags auf Grundlage des neu vorliegenden Ergebnisses der KNB hat vielmehr losgelöst vom Ergebnis der Nutzenbewertung gemäß § 35a SGB V zu erfolgen, da der G-BA gemäß § 35b (3) Satz 3 SGB V neu über den Zusatznutzen entscheidet. Würde hingegen das Ergebnis der Nutzenbewertung gemäß § 35a SGB V weiterhin uneingeschränkt herangezogen, wäre die Konsequenz, dass die Leitplanken für die Erstattungsbetragsfindung für ein Arzneimittel ohne Zusatznutzen nach § 130b (3) SGB V unabhängig von dem Ergebnis der KNB bestehen bleiben und der pU keine Möglichkeit hätte, bei einer Neuverhandlung des Erstattungsbetrags einen Aufschlag auf den vorherigen Erstattungsbetrag zu verhandeln. Folglich bestünden für den pU im Fall eines nicht belegten Zusatznutzens keine Anreize, eine KNB zu beantragen. Damit wäre das Instrument der KNB für den pU nutzlos.

Lösungsvorschlag:

Streichen der Feststellungen des IQWiG in Unterabschnitt 4.5.

A.1.23 GlaxoSmithKline GmbH & Co. KG (GSK)

Autorinnen und Autoren

- Westermayer, Bernd

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Westermayer, Bernd
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: GSK
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>
In den Änderungen im Entwurf der Allgemeinen Methoden 7.0 ergeben sich an mehreren Stellen wichtige Änderungen für forschenden Pharmaunternehmen. Darunter sind Anpassungen, die sich auf die Nutzenbewertung von Arzneimittel auswirken.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 9.3.2 zu statistischer Signifikanz und Konfidenzintervallen, S. 170	Anmerkung: Die Verwendung von Adjustierung sollte genauer spezifiziert werden. Vorgeschlagene Änderung: Es sollte präzisiert werden, ob Methoden zur Adjustierung für multiples Testen regelhaft beim Vorliegen von mehreren Operationalisierungen eines Endpunktes angewandt werden sollen. Das Vorgehen einer möglichen Änderung der üblichen Signifikanzschwelle sollte genauer beschrieben und an Beispielen präzisiert werden.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

**A.1.24 Institute of Public Health, Medical Decision Making and Health Technology
Assessment, UM-IT Tirol**

Autorinnen und Autoren

- Siebert, Uwe
- Sroczynski, Gaby

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigelegten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Prof. Dr. Uwe Siebert, MPH, MSc (Institutsleiter)
Assoc.-Prof. PD Dr. Gaby Sroczynski, MPH (Programmkordinatorin)
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Program on Decision Analysis, Institute of Public Health, Medical Decision Making and Health Technology Assessment, UMIT TIROL – University for Health Sciences and Health Technology, Hall in Tirol, Österreich
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Diese Stellungnahme bezieht sich vorwiegend auf die Inhalte von Abschnitt 4 Kosten-Nutzen-Bewertung medizinischer Interventionen. Vereinzelt werden auch andere Passagen kommentiert, die einen Bezug zu ökonomischen Evaluationen oder entscheidungsanalytischen Modellierungen besitzen wie z.B. Abschnitt 3 Nutzenbewertung medizinischer Interventionen, da die Ergebnisse aus der Nutzenbewertung wesentliche Parameter für die gesundheitsökonomischen Evaluationen darstellen.

Insgesamt wurden in der Version 7.0 umfangreiche Änderungen im Kapitel 4 Kosten-Nutzen-Bewertung medizinischer Interventionen vorgenommen. Die Konzepte, Begriffe und methodischen Ansätze entsprechen in dieser Version deutlich mehr den Inhalten internationaler Lehrbücher und methodischer Standardwerke, die auch zum großen Teil von anderen HTA-Organisationen verwendet werden, sofern sie ausführliche methodische Guidelines besitzen. Diese Änderung in Version 7.0 wird aus wissenschaftlicher Perspektive sehr begrüßt.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung
	<i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
3.3.4. und 9.3.6	Anmerkung: In diesen Abschnitten wird auf das Verzerrungspotenzial durch Confounding in Beobachtungsstudien eingegangen. Diese Hinweise werden sehr begrüßt, da dies eines der größten Probleme bei Beobachtungsstudien ist. Dies ist insbesondere vor der derzeit teils relativ unkritischen Diskussion zu „Real-World-Evidence“ sehr relevant. Die beschriebenen Adjustierungsverfahren (z. B. multifaktorielle Regressionsmodelle, Instrumentenvariablen, Propensity Scores) beschränken sich auf zeitunabhängiges (baseline) Confounding. Die Nennung der Verfahren für zeitabhängiges Confounding fehlen.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Änderungsvorschlag:</u> An dieser Stelle sollte darauf hingewiesen werden, dass zur Kontrolle von zeitabhängigem Confounding ggf. spezifische Kausalinferenz-Verfahren zur Kontrolle dieses Confounding-Typs erforderlich sind wie u.a. die parametrische g-Formula, marginal structural models mit inverse probability weighting, oder structural nested models mit g-estimation (Robins et al. 2004, Jahn et al. 2022, Kuehne 2022b). <u>Anmerkung:</u> Die Bemerkungen zu Variablenselektionsverfahren (Backward-Elimination, Forward-Selektion) stammen aus der Modellbildung für Prädiktions- oder prognostische Modelle, nicht aus dem Repertoire der Kausalmodelle. Hier werden Kausaldiagramme und andere Verfahren zur Variablenauswahl verwendet. <u>Änderungsvorschlag:</u> „Wichtig ist hierbei eine sorgfältige und systematische Vorauswahl der potenziellen erklärenden Variablen unter sach-wissenschaftlicher Kenntnis [176], z.B. unter Verwendung von Kausaldiagrammen (Directed acyclic graphs, DAGs) (Jahn et al 2022, Kühne et al. 2022b)“. <u>Anmerkung:</u> Im Gegensatz zu Confounding werden anderen „self-inflicted biases“, also durch korrektes Studiendesign und adäquate Datenaufbereitung und -analyse vermeidbare Verzerrungen, weniger Raum gegeben. Beispielsweise wird das wissenschaftlich anerkannte Konzept des Target-Trial Ansatzes hier nicht explizit erwähnt. <u>Änderungsvorschlag:</u> Ansätze, die u.a. zu einer geeigneten expliziten Nachbildung einer randomisierten Studie gehören, wie die (Target Trial Emulation), sollten genannt und kurz beschrieben werden (Hernan et al., 2016). In der evidenzbasierten Medizin wird das Potential dieser Konzepte zunehmend erkannt (Braitmaier & Didelez 2022, Kuehne 2022a) und die Target Trial Emulation wurden jüngst auch von anderen HTA Organisationen in deren Methodenguidelines aufgenommen (NICE 2022). Diese Gelegenheit sollte auch im Update der Methodenversion 7.0 wahrgenommen werden.
3.5.1.2	<u>Anmerkung:</u> In diesem Abschnitt zu „Linked Evidence“ wird nicht explizit auf Screeningverfahren eingegangen. Hier sollte zumindest

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	kurz angesprochen werden, dass ähnliche Methoden auch zur Nutzen-Schaden-Bewertung von Screeningverfahren eingesetzt werden können, sofern keine aktuelle direkte Evidenz vorliegt. Dies ist insbesondere der Fall, wenn für ein Screeningverfahren der Nutzen bereits als gesichert gilt und ein entsprechendes Screeningprogramm bereits etabliert ist, aber verschiedene Varianten (Startalter, Stopalter, Screeningintervall, Screeningtestkombinationen) des Screenings verglichen werden.
4. Zum Begriff der Modellierung	<u>Übergreifende Anmerkung zum Abschnitt 4 Kosten-Nutzen-Bewertung medizinischer Interventionen:</u> Im Abschnitt 4 wird mehrfach auf „Modellierungen“ eingegangen. Es wird vorgeschlagen, in Abschnitt 4 zu Beginn darzulegen, dass es sich beim Begriff „Modellierungen“ um entscheidungsanalytische Modelle handelt (grundsätzlich Vergleich verschiedener Interventionen und damit Kausalinterpretation), um diese von statistischen Modellen zur reinen Prädiktion/Assoziation abzugrenzen. <u>Beispielhafter Änderungsvorschlag:</u> z.B. S. 84 erste Zeile wie folgt ergänzen: „...werden in der Regel Modellierungen mit Hilfe von entscheidungsanalytischen Modellen durchgeführt (Siebert, 2003)“. <u>Anmerkung:</u> In Abschnitt 4.9 Gesundheitsökonomisches Studiendesign wäre es hilfreich, im Unterabschnitt ‚Auswahl der Modellierungstechnik‘ neben der Referenz 596 auch auf die internationalen Empfehlungen zu den genannten einzelnen Verfahren hinzuweisen. <u>Änderungsvorschlag:</u> „Die Konzeption des Modells ist in der Regel ausschlaggebend dafür, welche Modellierungstechnik (Entscheidungsbaum, Markov-Modell, diskrete Ereignissimulation, etc.) zum Einsatz kommt [596, weitere Referenzen: Karnon et al. 2012, Pitman et al 2012, Siebert et al. 2012]“. Am Ende des einführenden Unterkapitels in Abschnitt 4.9 (S. 99) sollte darauf hingewiesen werden, dass Inputparameter für entscheidungsanalytische Modellierungen, die sich auf Interventionseffekte beziehen, mittels der entsprechenden Verfahren der Kausalinferenz abzuleiten sind (Jahn et al. 2022, Kühne et al. 2022b)“.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
4.1 Einführung	<u>Anmerkung:</u> Die Einordnung der Methode 'gesundheitsökonomische Evaluation' in den Gesamtkontext der umfassenderen Disziplin Gesundheitsökonomie in Abschnitt 1.4 wird begrüßt. Ebenso wird die ausführliche Darstellung der Bedeutung der gesundheitsökonomischen Evaluation einschließlich der Begriffsbestimmungen in diesem Abschnitt begrüßt, die in der Version 7.0 der internationalen Terminologie aus wissenschaftlichen Standardwerken und Lehrbüchern entsprechen.
4.1. „Ziel der KNB ist es, ökonomische Informationen als Zusatz zur Nutzenbewertung insbesondere für die Preisverhandlungen im Sinne einer Informationssynthese zusammenzufassen.“	<u>Anmerkung:</u> Ziel einer Kosten-Effektivitätsanalyse (und somit die angewandten Methode in der KNB gemäß § 35b SGB V bzw. in Kapitel 1.4. wird die KNB definiert als Kosten-Effektivitätsanalyse) ist nicht nur die ökonomische Information in einer Informationssynthese zusammenzufassen, sondern systematisch und transparent die jeweiligen Erwartungswerte der <u>verschiedenen</u> Zielgrößen wie Nutzen, Schaden, Kosten gegenüberzustellen und diese Gegenüberstellung (Abwägung) nachvollziehbar zu berichten, z.B. in Form inkrementeller Verhältnisse. Eine Limitation des Ziels einer KNB auf Preisverhandlungen ist aus wissenschaftlicher Sicht zu kurz gegriffen.
4.1 „Die KNB gemäß § 35b SGB V umfasst neben der Berechnung des inkrementellen Kosten-Effektivitäts-Verhältnisses (ICER) dessen Einordnung sowie eine Ausgaben-Einfluss-Analyse.“	<u>Anmerkungen:</u> Die Beschreibung der Kosten-Effektivitätsanalyse und der Ausgaben-Einfluss-Analyse wird begrüßt. Eine Ausgaben-Einfluss-Analyse hat ein anderes Ziel als eine Kosten-Effektivitätsanalyse. Die KNB wird in diesem Satz nicht wie vorher in Kapitel 1.4. definiert als „Synonym“ für Kosten-Effektivitätsanalysen angewandt (vgl. Kapitel 1.4.). Hier könnten klarer die unterschiedlichen Ziele einer Kosten-Effektivitätsanalyse und einer Ausgaben-Einfluss-Analyse dargestellt werden: • Eine Kosten-Effektivitätsanalysen dient der <u>Entscheidungsfindung</u> bezüglich der Auswahl einer effektiven und effizienten Intervention. • Eine Ausgaben-Einfluss-Analyse dient der Planung innerhalb der Budgetperiode.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Wie genau die „Einordnung des inkrementellen Kosten-Effektivitäts-Verhältnisses (ICER) durchgeführt werden soll, könnte expliziter dargestellt werden. Der offene Umgang mit diesem Thema ist vermutlich (wie in vielen anderen Ländern auch) intendiert.
4.2 Tabelle 6. Gesundheitsökonomische Zielgröße: Kosten pro Endpunkte der Nutzenbewertung, Kosten pro QALY	<u>Anmerkung:</u> Die explizite Nennung eines Referenzfalls, wie dies in anderen Ländern auch der Fall ist, wird sehr begrüßt und dient der Transparenz, Vergleichbarkeit und Nachvollziehbarkeit, aber auch der Planbarkeit bei Studien seitens der Hersteller. <u>Vorgeschlagene Änderungen:</u> • Explizite Nennung des Begriffs inkrementelles Kosten-Effektivitäts-Verhältnis (ICER) in der Zeile „Gesundheitsökonomische Evaluation“. • Hervorhebung des <u>inkrementellen</u> Charakters einer vollständigen gesundheitsökonomischen Evaluation durch „Zusätzliche Kosten pro gewonnenem QALY“ etc. • Hinweis darauf, dass ICERs nur für nicht (stark oder schwach) dominierte Interventionen zu berechnen sind.
4.2 Tabelle 6. Studienform GÖ Evaluation: Kosten-Nutzwert-Analyse oder Kosten-Effektivitätsanalyse jeweils mit einer ergänzenden Ausgaben-Einfluss-Analyse	<u>Anmerkung:</u> Eine Ausgaben-Einfluss-Analyse ist nicht eine Ergänzung einer Kosten-Effektivitäts-Analyse, sondern dies sind 2 verschiedene Studientypen. Eine Ausgaben-Einfluss-Analyse kann allerdings eine Ergänzung im Rahmen einer Kosten-Nutzen-Bewertung des IQWiG sein. <u>Vorgeschlagene Änderung:</u> Kosten-Nutzwert-Analyse oder Kosten-Effektivitätsanalyse und zusätzlich eine Ausgaben-Einfluss-Analyse.
4.2 Tabelle 6. Nutzwerte: Für die Berechnung von QALYs sollen die für die in das Modell einfließenden Nutzwerte auf der Bewertung durch Patientinnen und Patienten basieren.	<u>Anmerkung:</u> Die expliziten Hinweise auf den QALY-Ansatz als integratives Maß für Lebensdauer und Lebensqualität werden begrüßt. <u>Vorgeschlagene Änderung:</u> Ggf. könnte auch auf „Disutilities“ – also eine Verringerung von Nutzwerten (z.B. temporär durch unerwünschte Effekte des Arzneimittels) – hingewiesen werden.
4.2 Tabelle 6. Zeithorizont: In Szenarioanalysen wird ergänzend ein Zeithorizont von 5 Jahren betrachtet.	<u>Anmerkung:</u> Der Sinn einer Szenarioanalysen in einer Kosteneffektivitätsanalyse mit einem Zeithorizont von 5 Jahren ist aus wissenschaftlicher Sicht nicht nachvollziehbar und wird in vielen Fällen nicht konsistent mit den Vorgaben

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	in Kapitel 5.4.5. zur Wahl des Zeithorizonts sein (vgl. Notwendigkeit eines angemessenen Zeithorizonts zur Abbildung aller relevanten Kosten und Nutzenaspekte). Eine grundsätzliche Empfehlung für einen 5-Jahreshorizont sollte der Ausgaben-Einfluss-Analyse vorbehalten sein.
4.3.2 Intervention und Komparator(en) „Der Nachweis des Zusatznutzens des zu bewertenden Arzneimittels erfolgt dabei regelhaft im Vergleich zur vom G-BA bestimmten zweckmäßigen Vergleichstherapie (zVT). Sind mehrere Alternativen für die Vergleichstherapie gleichermaßen zweckmäßig, genügt der Nachweis des Zusatznutzens gegenüber einer dieser Therapien als Grundlage für den Beschluss zum Zusatznutzen gegenüber der Gesamtheit der zVT. ... Im Basisfall der KNB erfolgt daher der Vergleich zwischen dem zu bewertenden Arzneimittel und der für den Nachweis des Zusatznutzens herangezogenen zVT.“	<u>Anmerkung:</u> Die Beschränkung auf einen Komparator kann zu suboptimalen Entscheidungen führen, u.a. ignoriert man evtl. die Möglichkeit erweiterter Dominanz, was zu einer Fehleinschätzung (Unterschätzung) der ICER führen kann. Wissenschaftliche Empfehlungen sprechen sich deshalb für den Einschluss von allen relevanten konkurrierenden alternativen Interventionen ein. Das Vorgehen des IQWiG ist allerdings im Sinne der Vereinfachung durchaus auch in anderen HTA-Organisationen üblich. Evtl. kann hier darauf hingewiesen werden, dass dies eine „Vereinfachung“ ist.
4.3.3 Gesundheitsökonomische Zielgröße „Vor dem Hintergrund dieser Vorgaben ist es naheliegend die KNB in Form einer Kosten-Effektivitäts-Analyse durchzuführen. Hierbei können Kosten in Relation zu einzelnen Endpunkten der Nutzenbewertung als	<u>Anmerkung:</u> Die Empfehlung, die KNB in Form einer Kosten-Effektivitäts-Analyse durchzuführen wird aus wissenschaftlicher Perspektive begrüßt. <u>Vorgeschlagene Änderung:</u> Evtl. könnte im Text das inkrementelle Kosten-Effektivitätsverhältnis (ICER) explizit als Zielgröße genannt werden und noch expliziter auf den <u>inkrementellen</u> Ansatz eingegangen werden.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
gesundheitsökonomische Zielgröße bestimmt werden. Im Fokus sollte dabei die Berechnung des Kosten-Effektivitäts-Verhältnisses bezogen auf ausgewählte möglichst entscheidungsleitende Endpunkte stehen.“	
4.3.3 Gesundheitsökonomische Zielgröße „Das QALY-Konzept stellt dennoch die derzeit am umfassendsten angewandte Methodik dar, die explizit und mit einem theoretisch begründeten Konstrukt die Bewertung der Effektivität von Therapien in gesundheits-ökonomischen Evaluationen unter Berücksichtigung von Präferenzen ermöglicht „	<u>Anmerkung:</u> Die Beschreibung des QALY-Konzept als derzeit am umfassendsten angewandte Methodik wird begrüßt. <u>Vorgeschlagene Änderung:</u> Hier könnte man ggf. ein wesentliches Argument für die Verwendung des QALY ergänzen: die Integration verschiedenster unterschiedlich gewichteter Konsequenzen, inkl. solcher die sich nicht auf die Lebenserwartung auswirken. Es sollte auch darauf hingewiesen werden, dass das Konzept QALY nicht gleichzusetzen ist mit dem Konzept Kosten pro QALY. Dies führt noch immer zu Unklarheiten in wissenschaftlichen Diskussionen.
4.3.5 Zeithorizont „...Daher sollte insbesondere unter Berücksichtigung der Indikation, der Studiendauer, der langfristigen Konsequenzen der betrachteten Therapien, des Zeitraums, über den Kosten entstehen, der Dynamik des Therapiegeschehens und auch der Perspektive des Entscheidungsträgers ein angemessener Zeithorizont gewählt werden [220,245,348,597]. In	<u>Anmerkung:</u> Es erschließt sich nicht, warum Szenarioanalysen regelhaft einen Zeithorizont von 5 Jahren haben sollten (vgl. Kommentar zur Tabelle 6). Es steht im Gegensatz zum vorangehenden Satz, der einen angemessenen Zeithorizont fordert, anstatt eine feste Jahreszahl zu nennen. Ein genereller Zeithorizont von 5 Jahren macht allenfalls für die Ausgaben-Einfluss-Analysen Sinn, da es hierbei um Planungszeiträume geht.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Szenarioanalysen wird regelmäßig ein Zeithorizont von 5 Jahren betrachtet.“	
4.4 „In Abhängigkeit von der gewählten Zielgröße für die gesundheitliche Veränderung unterscheidet man verschiedene Formen vergleichender gesundheitsökonomischer Evaluationen [208]. Aufgrund der Wahl von Kosten pro zusätzlicher Einheit eines Endpunkts aus der Nutzenbewertung oder Kosten pro QALYs als Zielgröße für den Referenzfall wird die KNB daher in Form einer Kosten-Effektivitäts- Analyse oder in Form einer Kosten-Nutzwert-Analyse durchgeführt jeweils mit einer ergänzenden Ausgaben-Einfluss- Analyse.“	<u>Anmerkung:</u> Siehe Kommentare in Tabelle 6. <u>Vorgeschlagene Änderung:</u> Für den Referenzfall wird die KNB in Form einer Kosten- Effektivitäts-Analyse (Zielgröße ICER in Euro pro zusätzlicher Nutzeneinheit) bzw. in Form einer Kosten- Nutzwert-Analyse (Zielgröße ICER in Euro pro gewonnenem QALY) durchgeführt. Ergänzend wird eine Ausgaben- Einfluss-Analyse durchgeführt.
4.6 Nutzwerte (im Fall einer Kosten-Nutzwert- Analyse)	<u>Anmerkung:</u> Ein eigener Abschnitt zu Nutzwerten als ein wesentliches Element einer Kosten-Nutzwert-Analyse wie hier wird begrüßt. Ggf könnten hier weitere Hinweise zum Umgang mit der Lebensqualität in Abhängigkeit von den Populations- Charakteristika wie z.B. Alter und Geschlecht enthalten. Hilfreich wäre auch ein Hinweis zum Umgang mit (temporären) Lebensqualitätsreduktionen (z.B. durch vorübergehende unerwünschte Effekte von Arzneimitteln) (s. Kommentar oben).
4.9 Gesundheitsökonomisches Studiendesign „... Gesundheitsökonomische	<u>Anmerkung:</u> Diese Aussage ist aus wissenschaftlicher Sicht teilweise nicht korrekt. Korrekt ist, dass das Ziel von Modellierungen nicht eine exakte Prädiktion der zukünftigen Realität leisten kann

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Modelle zielen nicht darauf ab, die Realität präzise vorherzusagen, Aus den beschriebenen Gründen eignen sich Modellierungen jedoch nicht als Basis für einen Nachweis des Zusatznutzens im Sinne des § 35a SGB V und haben insbesondere prognostische Bedeutung.“	(schon deshalb weil sich Umgebungsverhältnisse ändern können, wie z.B. die Lebenserwartung aufgrund unvorhersehbaren technischen Entwicklungen). Korrekt ist aber auch, dass Modellierungen das Ziel der Entscheidungsunterstützung für die Auswahl einer Intervention haben. Dazu ist es wichtig, dass Modellierungen unter kausalen Aspekten durchgeführt werden („causal decision-analytic modeling“) und in den Modellen keine reinen (statistischen) Prädiktionen oder Prognosen ermittelt werden, sondern Erwartungswerte der Zielgrößen, die rechnerisch aus kausalen Einflüssen der verglichenen Interventionen folgen (Siebert 2003). Dies hat insb. Konsequenzen für die Auswahl der Modellparameter, die kausalen und nicht rein assoziativen Anforderungen genügen müssen (Kühne et al. 2022b). Dies wird z.B. in dem folgenden Satz auf S. 99 sehr klar: „Es ist wichtig, dass es sich bei entscheidungsanalytischen Verfahren, zu denen auch gesundheitsökonomische Modellierungen gehören, um Instrumente zur Entscheidungsfindung unter Unsicherheit handelt.“ Dies zeigt, dass statt „Modellierungen“ besser der Begriff entscheidungsanalytische Modellierungen verwendet werden sollte (s. Kommentar oben). <u>Vorgeschlagene Änderung:</u> Entweder diesen letzten Satz des Absatzes löschen oder abändern in: „Deshalb sind für Nutzenbewertungen Modellierungen eher dann von besonderer Bedeutung, wenn eine Entscheidung getroffen werden muss, aber empirische Studien aus Gründen der Machbarkeit, Ethik oder Dauer nicht möglich sind. Dies ist z.B. meist bei der Bewertung oder Optimierung von Screeningverfahren der Fall (z.B. Optimierung der Altersgrenzen oder Screeningintervalle).
4.12 Einordnung des ICERs „Die Zumutbarkeit einer Kostenübernahme durch die GKV-Versichertengemeinschaft hängt einerseits von der Angemessenheit des Preises einer Therapie ab, andererseits aber auch von	<u>Anmerkung:</u> Diese Einordnung des ICERs ist grundsätzlich zu begrüßen Es bleibt unklar, wie die Zahlungsbereitschaft der GKV-Versicherten ermittelt werden soll (s. Kommentar oben).

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
den damit verbundenen künftigen Gesamtausgaben in Abhängigkeit der finanziellen Leistungsfähigkeit und Zahlungsbereitschaft der GKV- Versichertengemeinschaft.“	

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

Braitmaier M, Didelez V (2022). Emulierung von „target trials“ mit Real-world-Daten. Präventiv Gesundheitsf 2022. <https://doi.org/10.1007/s11553-022-00967-9>

Hernán MA, Robins JM (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64. doi: 10.1093/aje/kwv254. Volltext: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4832051/>

Jahn B, Friedrich S, Behnke J, Engel J, Garczarek U, Münnich R, Pauly M, Wilhelm A, Wolkenhauer O, Zwick M, Siebert U, Friede T (2022). On the role of data, statistics and decisions in a pandemic. Adv Stat Anal. 2022;106(3):349-382. doi: 10.1007/s10182-022-00439-7. Volltext: <https://pubmed.ncbi.nlm.nih.gov/35432617/>

Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J (2012). Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Med Decis Making. 2012;32(5):701-11. doi: 10.1177/0272989X12455462. Volltext: <https://pubmed.ncbi.nlm.nih.gov/22990085/>

Kuehne F, Arvandi M, Hess LM, Faries DE, Matteucci Gothe R, Gothe H, Beyrer J, Zeimet AG, Stojkov I, Mühlberger N, Oberaigner W, Marth C, Siebert U (2022a). Causal analyses with target trial emulation for real-world evidence removed large self-inflicted biases: systematic bias assessment of ovarian cancer treatment effectiveness. J Clin Epidemiol. 2022;152:269-280. doi: 10.1016/j.jclinepi.2022.10.005. Epub 2022 Oct 15. PMID: 36252741.

Kühne F, Schomaker M, Stojkov I, Jahn B, Conrads-Frank A, Siebert S, Sroczynski G, Puntscher S, Schmid D, Schnell-Inderst P, Siebert U (2022b). Causal Evidence in Health Decision Making: Methodological Approaches of Causal Inference and Health Decision Science. HTA Report Vol. 509-1. GMS German Medical Science — an Interdisciplinary Journal 2022;20:Doc12. DOI: 10.3205/000314, URN: urn:nbn:de:0183-0003147

NICE real-world evidence framework (2022). Corporate document [ECD9]. Published: 23 June 2022. Internet: <https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837>

Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, Brisson M (2012). Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-5. *Med Decis Making*. 2012;32(5):712-21. doi: 10.1177/0272989X12454578. Volltext: <https://pubmed.ncbi.nlm.nih.gov/22990086/>

Robins JM, Hernán MA, Siebert U (2004). Estimations of the Effects of Multiple Interventions. In: Ezzati M, Lopez AD, Rodgers A, Murray CJL (eds.). *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Vol. 1. World Health Organization: Geneva, 2004;2191-2230. Volltext: https://apps.who.int/iris/bitstream/handle/10665/42770/9241580313_eng.pdf?sequence=1&isAllowed=y

Siebert U (2003). When should decision-analytic modeling be used in the economic evaluation of health care? *The European Journal of Health Economics* 2003;4(3):143-50.

Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, Kuntz KM (2012). State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Med Decis Making*. 2012;32(5):690-700. doi: 10.1177/0272989X12455463. Volltext: <https://pubmed.ncbi.nlm.nih.gov/22990084/>

A.1.25 Ipsen Pharma GmbH

Autorinnen und Autoren

- Kucka, Annett

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Kucka, Annett
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Ipsen Pharma GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abs. 2.1.3.2 zu AbD-Konzept (S. 16)	<u>Anmerkung:</u> Das IQWiG führt im Abschnitt 2.1.3.2 das Konzept der anwendungsbegleitenden Datenerhebung (AbD) neu in den Entwurf des Methodenpapiers 7.0 ein. Das AbD-Konzept dient dazu weitere Informationen für die Nutzenbewertung zu erhalten. Hier wäre es essenziell wichtig, dass die pharmazeutische Industrie in den Prozess der AbD frühzeitig eingebunden wird und Stellung nehmen kann. Es sollte im Rahmen einer optimalen Ausgestaltung des Verfahrens der AbD, einen Austausch geben aber mindestens eine Art Erörterung im Verlauf des Vorhabens mit dem pU um die erwartbaren Informationen zu besprechen. <u>Vorgeschlagene Änderung:</u> Die Allgemeinen Methoden des IQWiGs sollten für eine optimale Ausgestaltung einer AbD eine frühzeitige Einbindung des pUs in den Prozess beinhalten.
Abs. 3.3.4 zu Konzept für AbD gemäß § 35a Abs. 3b SGB V (S.66)	<u>Anmerkung:</u> Ebenfalls führt das IQWiG in Abschnitt 3.3.4 das Konzept für AbD gemäß § 35a Abs. 3b SGB V und die dazugehörigen Inhalte und methodischen Grundzüge in den Entwurf des Methodenpapiers 7.0 ein. Das AbD-Konzept dient als Möglichkeit der Evidenzgenerierung, wenn im Falle von besonderen Therapieoptionen weitere Evidenz für die Nutzenbewertung benötigt wird. Jedoch wird aus den

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	entsprechenden Abschnitten nicht ersichtlich, wie das IQWiG die Machbarkeit des AbD-Konzeptes bewertet. So schreibt das IQWiG: <i>„Basierend auf den vorliegenden Erkenntnissen zum Arzneimittel und zur zweckmäßigen Vergleichstherapie wird abgeschätzt, ob innerhalb eines akzeptablen Zeitraums prinzipiell ausreichend Patientinnen und Patienten rekrutiert und beobachtet werden können, um mit einer AbD grundsätzlich aussagekräftige Ergebnisse für eine Nutzenbewertung generieren zu können“</i> Hier beschreibt das IQWiG leider nicht, wie es die Machbarkeit auf Basis einer wissenschaftlichen fundierten Rationalen ermöglicht, wenn z.B. Fallzahlen generell oder akzeptable Zeiträume nicht erreicht werden können. Hier sollte es Kriterien geben, die eine nachvollziehbare Prüfung z.B. für einen akzeptablen Zeitraum einer AbD oder ausreichend hohe Patientenzahlen, ermöglichen. <u>Vorgeschlagene Änderung:</u> Für den entsprechenden Abschnitt sollte nachvollzogen werden können, auf welche Kriterien die Machbarkeit einer AbD durch das IQWiG berücksichtigt und geprüft wird.
Abs. 9.1.1 zu Kriterien für den Einschluss von Studien (S. 159)	<u>Anmerkung:</u> In Abschnitt 9.1.1 führt das IQWiG die folgenden Kriterien für den Einschluss von Studien neu an: <i>„Falls geeignete Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium (Population bzw. Interventionen) vorliegen, muss über die Berücksichtigung der Studie anhand der Stärke der Effektmodifikation und dem Anteil der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen bzw. dem Grad der Abweichung der Interventionen entschieden werden.“</i> Zum einen geht aus dem Entwurf für das Methodenpapier 7.0 nicht hervor, ob im Gegensatz zu Methodenpapier 6.1 auch Studien, bei den das Einschlusskriterium weniger als bei 80% der Population erfüllt ist, eingeschlossen werden können. Hier sollte genauer beschrieben werden, wann Teilpopulationen in die Bewertung eingeschlossen oder ausgeschlossen werden. Auch sollte genauer definiert werden, was eine geeignete Information über eine Effektmodifikation ist. <u>Vorgeschlagene Änderung:</u> Die Kriterien zum Ausschluss von Studien sollten präzisiert werden.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abs. 9.3.2 zu statistischer Signifikanz und Konfidenzintervallen (S. 170)	<u>Anmerkung:</u> Das IQWiG ergänzt im Abschnitt 9.3.2 zur statistischen Signifikanz und Konfidenzintervallen folgendes: „Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an [59], z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.“ Für die verschiedene Operationalisierungen von Endpunkten werden in der Nutzenbewertung die Primäranalysen und die Sensitivitätsanalysen in der Operationalisierung der Endpunkte festgelegt, sodass hier die Notwendigkeit einer Adjustierung für multiples Testen entfällt. <u>Vorgeschlagene Änderung:</u> Die Herleitung wann die Adjustierung für multiples Testen z.B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht wird, sollte nachvollziehbar und wissenschaftlich begründet werden. Zudem sollte die Anforderung internationalen Standards und Kriterien der evidenzbasierten Medizin genügen. Anderenfalls sollte diese Neuerung gestrichen werden.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

A.1.26 Janssen-Cilag GmbH

Autorinnen und Autoren

- Mehnert, Angelika
- Eisele, Lewin

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Mehnert, Angelika, Dr.
Eisele, Lewin, Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Janssen-Cilag GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) (S. 66)	Folgende Passagen wurden in Version 7.0 angepasst: <i>„Aus einer nicht randomisierten Studie kann aufgrund potenziell unbekannter Confounder aus den in der Studie beobachteten Effekten erst ab einer bestimmten Effektstärke eine Aussage zum Nutzen oder Schaden einer Intervention abgeleitet werden (siehe Abschnitte 3.1.4 und 3.2.2), z. B. in einem Bereich von 2 bis 5 für das relative Risiko (bzw. 0,2 bis 0,5 bei mortalitätssenkenden Interventionen) bezogen auf die untere bzw. obere Konfidenzintervallgrenze [393].“ (S. 68, Z. 6ff)</i> <u>Anmerkung:</u> Aufgrund potenziell unbekannter Confounder kann laut den Ausführungen in der Version 7.0 des Methodenpapiers erst ab einer bestimmten Effektstärke bei nicht randomisierten Studien eine Aussage zur Intervention gemacht werden. Für diese Festlegung der benötigten Effektstärke für den Bereich von 2 bis 5 für das relative Risiko (bzw. 0,2 bis 0,5 bei mortalitätssenkenden Interventionen) bezogen auf die untere bzw. obere Konfidenzintervallgrenze fehlen jedoch nähere Informationen zur Herleitung der Grenzen. Zudem wird darauf verwiesen, dass für die benötigte Effektstärke keinerlei Limitationen hinsichtlich der Datenqualität, wie beispielsweise das Fehlen eines einzelnen Confounders, unabhängig von seiner Relevanz für die Adjustierung, bestehen dürfen. Aus methodischer Sicht sollte das Fehlen von Daten für einzelne Confounder nicht als alleiniges Kriterium zu einer pauschalen Verneinung eines Effekts führen. Aus Sicht der Janssen-Cilag GmbH können bei der Beurteilung eines Effektes im Rahmen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	einer Indikation in der Nutzenbewertung Datenqualität und Effektgröße nicht anhand generischer Anforderungen gesondert voneinander betrachtet werden, sondern müssen in der Gesamtschau aus Datenqualität und Effektgröße in einer konkret vorliegenden Nutzenbewertung beurteilt werden. Bei dieser Betrachtung sollten zudem kontextuale Aspekte wie das Vorliegen einer besonderen Therapiesituation unter Berücksichtigung der Schwere und Seltenheit der Erkrankung, der Schutzwürdigkeit der Patientenpopulation und der (fehlenden) Verfügbarkeit von Behandlungsalternativen Berücksichtigung finden. Erst nach differenzierter Betrachtung und Bewertung der zur Nutzenbewertung vorgelegten Evidenz und kontextueller Aspekte kann beurteilt werden, ob ein beobachteter Effektunterschied für patientenrelevante Zielgrößen in einer konkreten Nutzenbewertung hinreichend ist. <u>Vorgeschlagene Änderung:</u> Für die Beurteilung, ob ein beobachteter Effektunterschied für patientenrelevante Zielgrößen in einer konkreten Nutzenbewertung hinreichend ist, sollten außer Datenqualität und Effektgröße auch kontextuale Aspekte wie das Vorliegen einer besonderen Therapiesituation unter Berücksichtigung der Schwere und Seltenheit der Erkrankung, der Schutzwürdigkeit der Patientenpopulation und der (fehlenden) Verfügbarkeit von Behandlungsalternativen Berücksichtigung finden.
Abschnitt 9.1.1 Kriterien für den Einschluss von Studien (S. 159)	Folgende Passage wurde in Version 7.0 angepasst: <i>„Liegen für eine Studie Analysen der relevanten Teilpopulation bzw. der relevanten Prüf- und Vergleichsinterventionen vor, wird auf diese Analysen zurückgegriffen. Liegen solche Analysen nicht vor, wird die Studie dann eingeschlossen, falls auf mindestens 80 % der eingeschlossenen Patientinnen und Patienten das Kriterium (Population) zutrifft bzw. falls das Einschlusskriterium bezüglich der Prüfintervention (Interventionsgruppe der Studie) und der Vergleichsintervention (Vergleichsgruppe der Studie) jeweils zu mindestens 80 % erfüllt ist. Falls geeignete Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium (Population bzw. Interventionen) vorliegen, muss über die Berücksichtigung der Studie anhand der Stärke der Effektmodifikation und dem Anteil der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen bzw. dem Grad der Abweichung der Interventionen entschieden werden.“</i> <u>Anmerkung:</u> Der Abschnitt zu den Kriterien für den Einschluss von Studien wurden vom IQWiG in Version 7.0 des Methodenpapiers gekürzt und verändert. Insbesondere der erwähnte

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Zusammenhang zwischen dem Einschluss von Studien und Effektmodifikationen wurde neu hinzugefügt. Im Entwurf für Version 7.0 des Methodenpapiers führt das IQWiG an, dass beim Vorliegen von geeigneten Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium anhand der Stärke der Effektmodifikation und anhand der betroffenen Anzahl an Patienten über den Einschluss der Studie entschieden wird. Aus dem Absatz des Methodenpapiers zu Version 7.0 geht nicht hervor, wann die Prüfung auf eine Effektmodifikation konkret erfolgen soll und in welchem Bezug die Prüfung zu dem genannten 80 % Schwellenwert steht. Es bleibt unklar, welche Informationen als „geeignete Informationen über eine Effektmodifikation“ gelten. Wie die Informationen über eine Effektmodifikation in die Entscheidung zum Heranziehen von Studiendaten einfließen soll, beispielsweise endpunktspezifisch oder ob es zu einem pauschalen Studienausschluss kommt, stellt das IQWiG ebenfalls nicht klar. Aus Sicht der Janssen-Cilag GmbH muss zudem für die Annahme einer Effektmodifikation neben einer Interaktion auch eine biologische Rationale vorhanden sein (1, 2). Somit stellt nicht jede Subgruppeninteraktion notwendigerweise eine Effektmodifikation dar. Die Janssen-Cilag GmbH bittet um weitere Konkretisierung bezüglich der genannten Unklarheiten.

Abschnitt 9.3.2 Statistische Signifikanztests und Konfidenzintervalle (S. 170)	Folgender Halbsatz wurde in der Version 7.0 hinzugefügt: <i>Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an [59], z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.</i> <u>Anmerkung:</u> Im Vergleich zu der Version 6.1 wird ein Halbsatz eingefügt, der einer Klarstellung bedarf. Es wird in dem Absatz weder nachvollziehbar gemacht, welche Konstellationen mit „verschiedenen Operationalisierungen“ von Endpunkten gemeint sind, noch, welche möglichen „Methoden zur Adjustierung“ unter welche Voraussetzungen oder Konstellationen vom Institut in diesem (Endpunkt-)Kontext als „sinnvoll und möglich“ erachtet werden sollen. Eine inhaltliche Stellungnahme zu dieser Absichtserklärung ist dadurch nicht möglich. Denkbare Konstellationen für verschiedene Konstellationen könnten sich aus vielfältigen Situationen ergeben, z. B. • für einen Endpunkt verschiedene Effektschätzer berechnet werden (z. B. Relatives Risiko und Hazard Ratio) oder • für einen Endpunkt verschiedene Schwellenwerte herangezogen werden (z. B. Schwellenwerte bei der MRD-Negativität) oder • ein Endpunkt sowohl stetig als auch als Responderanalyse ausgewertet wird oder • ein Endpunkt mit verschiedenen Operationalisierungen der Ereigniszeitanalysen analysiert wird (Zeit bis zur ersten Veränderung, Zeit bis zur anhaltenden Veränderung) • Messung zu verschiedenen Zeitpunkten • Messung desselben Endpunkts mit verschiedenen Instrumenten (z. B. Symptomatik) Diese Aufzählung umfasst lediglich alltägliche Beispiele und lässt sich durch weitere Konstellationen fortführen. Es ist zu bedenken, dass denkbare Konstellationen (s. o.), die eventuell unter den Begriff „verschiedener Operationalisierungen“ von Endpunkten fallen könnten, unmittelbare Folge von Vorgaben gemäß der Modulvorlage des G-BA oder den FAQ auf der G-BA-Website zur Darstellung von Endpunkten sind, die in Herstellerdossiers verpflichtend präsentiert werden müssen. Dieses betrifft beispielsweise die Präsentation der Effektschätzer als Relatives Risiko und als Hazard Ratio bei Vorliegen unterschiedlicher Behandlungs- und/oder Beobachtungsdauer (vgl. Modulvorlage M4 S. 32 (3)). Gleichmaßen gilt dieses auch für die Darstellung von
---	--

	Endpunkten mit verschiedenen Operationalisierungen. Wie auf der Website des G-BA dargelegt ist, stehen für die patientenberichteten Endpunkte für die Zeit bis zur Veränderung folgende Möglichkeiten zur Verfügung (4): Zeit bis zur erstmaligen Veränderung Zeit bis zur bestätigten Veränderung Zeit bis zur dauerhaften Veränderung Liegen unterschiedliche Beobachtungszeiten zwischen den Studienarmen vor, muss die Zeit bis zur erstmaligen Veränderung berichtet werden, da die Ergebnisse der beiden anderen Auswertungsmethoden in der Regel nicht sinnvoll interpretierbar sind (4). Abgesehen davon beschreiben die unterschiedlichen Ereigniszeitanalysen aus Patientensicht verschiedene patientenrelevante Endpunkte. Die genannten Beispiele, die den Vorgaben zu multiplen Auswertungen im Nutzendossier geschuldet sind, könnten dem formalen Verständnis nach unter multiples Testen fallen, mit dem paradoxen Ergebnis, dass eine Erhöhung der Anforderungen an darzulegende Endpunktanalysen mit einer Erhöhung der Signifikanzschwelle durch multiples Testen einhergehen würde. <u>Vorgeschlagene Änderung:</u> Da der eingefügte Halbsatz ohne erklärende Erläuterung nicht sinnvoll interpretiert werden kann, schlägt die Janssen-Cilag GmbH vor, den Halbsatz zu streichen und den Absatz wie in der Version 6.1 zu belassen. Andernfalls bittet die Janssen-Cilag GmbH um Klarstellung, dass die Darstellung von vorgegebenen sowie ergänzenden Analysen oder Sensitivitätsanalysen nicht zur Adjustierung für multiples Testen führt.
Abschnitt 9.3.6 Adjustierung und multifaktorielle Verfahren (S. 176)	Folgende Passagen wurden in Version 7.0 angepasst (S. 176): <u>„[...] eine eindeutige Beschreibung und A-priori-Festlegung der Zielvariablen und aller potenziell erklärenden Variablen“ (S. 176, Z. 4f)</u> <u>Anmerkung:</u> In Version 7.0 des Methodenpapiers ergeben sich hinsichtlich der Anforderungen an den Zeitpunkt der Festlegung der Zielvariablen und potenziellen Störgrößen Fragen. Unter Berücksichtigung des frühen Zeitpunktes, zu welchem die Ergebnisse der Confounderanalyse für die Studienplanung bzw. in der Vorbereitung und Durchführung vergleichender Analysen zum Zwecke der Nutzenbewertung benötigt werden und des aufgeführten Umfangs einer adäquaten Confounderanalyse und Validierung, verbleiben aus Sicht von Janssen-Cilag GmbH Unklarheiten. Das betrifft den Zeitpunkt der Aktualisierung der Confounderanalyse, die Teile des Identifizierungsprozesses und wie mit ggf. neu d.h. im

	Nachgang identifizierten, potenziellen Confoundern umgegangen werden kann. <u>Vorgeschlagene Änderung:</u> Aus Sicht der Janssen-Cilag GmbH ist eine Präzisierung der genannten Unklarheiten erforderlich. Folgende Passagen wurden in Version 7.0 angepasst (S. 176f): <i>„Weiterhin ist es erforderlich, dass die beiden zu vergleichenden Gruppen bezüglich aller relevanten Confounder ausreichend balanciert sind.“ (S. 177, Z. 28f)</i> <i>„Hierbei ist es insbesondere erforderlich</i> <i>• das kausale Modell exakt zu beschreiben, z. B. mithilfe kausaler Grafiken,</i> <i>• die Annahmen des kausalen Modells darzustellen,</i> <i>• fundiert zu begründen, warum diese Annahmen im konkreten Anwendungsfall zu rechtfertigen sind,</i> <i>• darzustellen, welche Verfahren und Kriterien bei der Modellanpassung verwendet werden sowie</i> <i>• zu erläutern, mit welchen Sensitivitätsanalysen die Robustheit der Ergebnisse überprüft werden.“ (S. 177, Z. 6ff)</i> <i>• „der Umgang mit fehlenden Daten,</i> <i>• der Umgang mit unplausiblen Daten und Ausreißern“ (S. 176, Z. 8f)</i> <u>Anmerkung:</u> Bei Variablen, die einen Einfluss auf den untersuchten Endpunkt haben und mit der Behandlung in Verbindung stehen kann es sich <i>potenziell</i> um Confounder handeln. Für eine Beurteilung, ob diese potenziellen Confounder auch relevant sind, d.h. ob sie geeignet sind in einer Analyse Verzerrung zu minimieren, wird in der Literatur die Berücksichtigung von kausalen Graphen vorgeschlagen (5, 6). Die Hinzunahme von Information zu kausalen Zusammenhängen kann dann dazu führen, dass als potenzielle Confounder identifizierte Variablen nicht in der Analyse berücksichtigt werden müssen, ggf. sogar nicht berücksichtigt werden sollten. Für die Darstellung von kausalen Zusammenhängen werden in der Literatur z.B. direkte azyklische Graphen (DAG) vorgeschlagen. (5, 6) Das IQWiG führt im Entwurf der Methodenpapiers die Berücksichtigung kausaler Zusammenhänge auf. Es bleibt jedoch unklar, wie dies erfolgen soll, ob also beispielsweise eine Darstellung von kausalen Zusammenhängen und die darin enthaltenen Annahmen mittels eines DAG erfolgen kann oder soll. In diesem Zusammenhang nennt das IQWiG im Entwurf
--	---

	auch die Durchführung von Sensitivitätsanalysen. Hierbei wäre eine Präzisierung wünschenswert, welchen Stellenwert das IQWiG Analysen zur Beurteilung von residuellem Confounding (7, 8) bei der Bewertung von Behandlungseffekten beimisst. <u>Vorgeschlagene Änderung:</u> Die Aussagen zu kausalen Grafiken und deren Verwendung sollten präzisiert werden. Eine Präzisierung, inwieweit Analysen zur Beurteilung von residuellem Confounding bei der Bewertung von Behandlungseffekten einfließen können, wäre wünschenswert.
Abschnitt 9.3.12 (Seite 191)	Folgende Textpassage wurde in der Version 7.0 hinzugefügt: <i>Für PROs ist in einer solchen Situation durch die systematisch verkürzte Datenerhebung die Interpretierbarkeit stark eingeschränkt. Nicht sinnvoll interpretierbar sind bei sich deutlich zwischen den Behandlungsarmen unterscheidenden Beobachtungsdauern sogenannte „dauerhafte“ Veränderungen (Verbesserungen bzw. Verschlechterungen). Denn eine „dauerhafte“ Veränderung ist in dem länger beobachteten Behandlungsarm aufgrund der dafür notwendigen ununterbrochenen Über- bzw. Unterschreitung des Schwellenwerts (Responsekriterium) potenziell schwerer zu erreichen. Darüber hinaus ist es in einem solchen Fall auch nicht sachgerecht, von einer „dauerhaften“ Veränderung zu sprechen, da die Beobachtungszeit nur einen verkürzten und ggf. sehr kleinen Anteil der gesamten Studiendauer abdeckt. Solche Auswertungen werden im Falle einer verkürzten oder sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauer daher in der Regel nicht für die Nutzenbewertung herangezogen. Insgesamt sind bei unterschiedlich langer Beobachtungsdauer zwischen den zu vergleichenden Gruppen Datenanalysen, die dies nicht adäquat berücksichtigen, nicht für die Nutzenbewertung verwertbar.</i> <u>Anmerkung:</u> In den neu hinzugefügten Absätzen bleibt wie auch schon in dem Methodenpapier 6.1 unklar, welcher Unterschied in den Beobachtungsdauern zwischen den vergleichenden Gruppen vorliegen muss, um vom IQWiG für die Nutzenbewertung als nicht verwertbar zu gelten und daher nicht für die Nutzenbewertung herangezogen zu werden. Ebenso unklar ist, ab welchem Erhebungszeitpunkt eine PRO-Erhebung als „verkürzt“ zu interpretieren ist. In den zulassungsrelevanten klinischen Studien werden die Erhebungszeitpunkte in Abhängigkeit von der zu

	untersuchenden Patientenpopulation gewählt. Dabei spielen neben der Erkrankung auch der Krankheitszustand und das Behandlungsregime eine Rolle. Dieses ist zum Beispiel gemäß Appendix 2 der Leitlinie der Europäischen Arzneimittelbehörde (European Medicines Agency, EMA) zur Bewertung onkologischer Arzneimittel bei Menschen verankert (9): <i>" PRO measures should be administered to study subjects at time points when there is a clear and hypothesis driven rationale for their use and when it is feasible to expect high levels of completion by the individual patient;"</i> Demnach sind die Erhebungen so gewählt, dass Rücklaufquoten für die eingesetzten Erhebungsinstrumente erhalten werden, die Ergebnisse mit ausreichender Sicherheit und möglichst geringer Verzerrung erwarten lassen. Eine zu häufige Erhebung geht häufig mit einer Abnahme der Compliance einher, was wiederum zu nur eingeschränkt bewertbaren Ergebnissen führt. Unterschiedliche Beobachtungsdauern entstehen, wenn in einem Arm die PRO-Instrumente länger erhoben werden als in dem anderen Arm. Auch wenn ein PRO bis zum Tod der Studienteilnehmer erhoben würde, kann eine Ursache für unterschiedliche Beobachtungsdauern in der zwischen den Armen unterschiedlichen Mortalität begründet sein. Eine solche Situation ist in der Onkologie insbesondere in den hinteren Linien anzutreffen. Es stellt sich die Frage, wie mit PRO-Daten umzugehen ist, die mortalitätsbedingt unterschiedlich lang erhoben und beobachtet worden sind. Wenn das Institut die Ergebnisse für eine dauerhafte Veränderung als nicht sinnvoll interpretierbar erachtet, wäre ein Hinweis hilfreich, welche Analyse als „adäquat“ und „verwertbar“ betrachtet würde. Insgesamt wäre ein wissenschaftlicher Diskurs darüber oder Hinweise dazu, welche Analysen in welchem Kontext als sinnvoll erachtet und präsentiert werden sollen, wünschenswert. Die alleinige Aussage, dass <i>„Datenanalysen, die dies nicht adäquat berücksichtigen, für die Nutzenbewertung nicht verwertbar“</i> seien, ist nicht hilfreich. <u>Vorgeschlagene Änderung:</u> Neben einer Erläuterung, wann bzw. ab welchem Erhebungszeitpunkt eine PRO-Erhebung als „verkürzt“ und ihre Auswertung eingeschränkt interpretierbar betrachtet würde, und ab welcher Differenz eine Beobachtung als unterschiedlich betrachtet wird, wäre ein wissenschaftlicher Diskurs darüber oder Hinweise dazu, welche Analysen in welchem Kontext als sinnvoll erachtet und präsentiert werden sollen, wünschenswert.
--	--

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. EUnetHTA. *EUnetHTA 21 – Individual Practical Guideline Document. D4.5 – APPLICABILITY OF EVIDENCE – PRACTICAL GUIDELINE ON MULTIPLICITY, SUBGROUP, SENSITIVITY AND POST HOC ANALYSES. Version 0.3, 04/07/2022.* 2022 [abgerufen am: 21.02.2023]. Verfügbar unter: <https://www.eunetha.eu/wp-content/uploads/2022/07/D4.5-Applicability-of-Evidence-For-Public-Consultation-v1.0.pdf>.
2. Gemeinsamer Bundesausschuss (G-BA). *Mündliche Anhörung gemäß § 35 a Abs. 3 Satz 2 SGB V des Gemeinsamen Bundesausschusses; hier: Daratumumab (D-812).* 2022 [abgerufen am: 23.01.2022]. Verfügbar unter: https://www.g-ba.de/downloads/91-1031-809/2022-0808_Wortprotokoll_Daratumumab_D-812.pdf.
3. Gemeinsamer Bundesausschuss (G-BA). *Anlage II.6: Modul 4 Dokumentvorlage, Version vom 16.12.2021 - Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen.* . 2021 [abgerufen am: 13.01.2022]. Verfügbar unter: https://www.g-ba.de/downloads/17-98-4831/2021-12-16_Anl2_6_Modul4.docx.
4. Gemeinsamer Bundesausschuss (G-BA). *Antworten auf häufig gestellte Fragen zum Verfahren der Nutzenbewertung. "Was muss bei der Dossiererstellung für die Auswertung von Responderanalysen als Ereigniszeitanalyse beachtet werden?".* 2023 [abgerufen am: 13.01.2023]. Verfügbar unter: <https://www.g-ba.de/themen/arzneimittel/arzneimittel-richtlinie-anlagen/nutzenbewertung-35a/fags/#was-muss-bei-der-dossiererstellung-fur-die-auswertung-von-responderanalysen-als-ereigniszeitanalyse-beachtet-werden>.
5. Rothman KJ, Lash TL, VanderWeele TJ, Haneuse S. *Confounding and confounders.* In: *Modern epidemiology.* Fourth edition. ed. Philadelphia: Wolters Kluwer / Lippincott Williams & Wilkins; 2021.
6. Hernán MA, Robins JM. *Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available.* *Am J Epidemiol.* 2016;183(8):758-764.
7. Lipsitch M, Tchetgen Tchetgen E, Cohen T. *Negative controls: a tool for detecting confounding and bias in observational studies.* *Epidemiology.* 2010;21(3):383-388.
8. VanderWeele TJ, Ding P. *Sensitivity Analysis in Observational Research: Introducing the E-Value.* *Ann Intern Med.* 2017;167(4):268-274.
9. European Medicines Agency (EMA). *Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man.* 2016 [abgerufen am: 18.01.2023]. Verfügbar unter: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf.

A.1.27 Kassenzahnärztliche Bundesvereinigung (KZBV)

Autorinnen und Autoren

- Beck, Jörg
- Hussein, Jameel

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Dr. Beck, Jörg MHA
Dr. Jameel Hussein, Rugzan MPH
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Kassenzahnärztliche Bundesvereinigung (KZBV), Universitätsstraße 73, 50931 Köln
☐ als Privatperson(en)

**Stellungnahme der KZBV zum
Entwurf der Allgemeinen Methoden
des IQWiG
Version 7.0, Stand: 06.12.2022**

Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) hat den Entwurf der Version 7.0 seiner Allgemeinen Methoden am 6. Dezember 2022 öffentlich zur Stellungnahme gestellt. Die regelmäßige Überarbeitung und die Straffung der Allgemeinen Methoden, das Literaturupdate sowie die Aktualisierungen der Methoden werden aus Sicht der KZBV als sinnvoll erachtet und begrüßt.

Die KZBV nimmt gern die Möglichkeit wahr, zu dem Entwurf die nachstehende Anmerkung zu übermitteln.

Anmerkung der KZBV

Abschnitt 1.3.1 Praktische evidenzbasierte Medizin, S. 6:

„Ärztinnen und Ärzte sowie Patientinnen und Patienten können dann besprechen, wie man unter Unsicherheit zu einer Entscheidung kommt. Gerade in unsicheren Situationen kommt es auf die persönlichen Präferenzen an, die darüber entscheiden, welche Option Patientinnen und Patienten wählen. Evidenzbasiert ist eine Entscheidung also, wenn sie 3 Elemente berücksichtigt: die Evidenz, den klinischen Zustand und die Umstände, in denen sich eine Person befindet, und ihre Präferenzen und Einstellungen [350].“

Laut IQWiG sollen evidenzbasierte Entscheidungen drei Elemente berücksichtigen: 1) die Evidenz, 2) den klinischen Zustand und die Umstände sowie 3) die Präferenzen und Einstellungen des Patienten. Diese Darstellung listet die klinische Expertise nicht auf. Die an dieser Stelle zitierte Publikation Haynes et al. (2002) betrachtet die klinische Expertise als integrierendes Element für die drei anderen, auch von IQWiG genannten Elemente. Die klinische Expertise ist in der Publikation von Haynes et al. als viertes Element in der Abb. 2 sowie unter „A new, prescriptive model for evidence-based clinical decisions“ zu sehen.

Auch in aktuelleren Publikationen wie z.B. Satterfield et al. (2009) und Cierpiakowska et al. (2017) wird die klinische Expertise nach wie vor als integrales Element des EbM-Modells betrachtet.

- **Die KZBV regt an, den Sachverhalt zu erläutern und die Begründung für das Nichtberücksichtigen der klinischen Expertise im EbM-Modell anzugeben.**

Literatur

1. Cierpiąłkowska, L.; Soroko, E.; Górska, D.; Sęk, H. (2017): Evidence-based assessment in a transtheoretical and paradigmatic approach. In: *Current Issues in Personality Psychology* 5 (3), S. 153–162. DOI: 10.5114/cipp.2017.70141.
2. Haynes, R. B.; Devereaux, P. J.; Guyatt, G. H. (2002): Clinical expertise in the era of evidence-based medicine and patient choice. In: *Evidence-Based Medicine* 7 (2), S. 36–38. DOI: 10.1136/ebm.7.2.36.
3. Satterfield, J. M.; Spring, B.; Brownson, R. C.; Mullen, E. J.; Newhouse, R. P.; Walker, B. B.; Whitlock, E. P. (2009): Toward a transdisciplinary model of evidence-based practice. In: *The Milbank quarterly* 87 (2), S. 368–390. DOI: 10.1111/j.1468-0009.2009.00561.x.

A.1.28 Maimon Research LLC

Autorinnen und Autoren

- Langle, Paul

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Langley, Paul; Prof.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Maimon Reseach LLC
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

EINLEITUNG

Der Innovationsbericht 2021 der Universität Bremen analysiert die Preisgestaltung und die Notwendigkeit von Preisbeschränkungen auf dem deutschen Arzneimittelmarkt [1]. Dieser Schwerpunkt ist nicht neu; die Preispolitik der Pharmahersteller, die ihre monopolistische, patentgeschützte Position ausnutzen, hat uns immer wieder beschäftigt. Als Reaktion darauf wurde eine Reihe von Maßnahmen ergriffen, die von negativer Publizität bis hin zu willkürlichen Referenzpreisen oder Preisobergrenzen reichen. Dies ist zum großen Teil auf einen Mangel an Informationen bei der Marktzulassung zurückzuführen. Eine Situation, die eindeutig im Interesse der Hersteller liegt und durch die mangelnde Bereitschaft der Gesundheitssysteme verstärkt wird, in Nachweise zu investieren oder solche Nachweise zu verlangen, die die Behauptungen über die Auswirkungen vergleichbarer Therapien während der Patentlaufzeit des Produkts belegen.

Der Zweck dieses Vermerks ist es, die These zu unterstützen, dass Preisentscheidungen nicht nur evidenzbasiert sein müssen, sondern dass jede Preisentscheidung vorläufig ist, vorbehaltlich der erfolgreichen Lösung von empirischen Wertansprüchen, die über den Lebenszyklus des Produkts oder Geräts repliziert und bewertet werden [2]. Dies stellt einen Neuanfang in der Bewertung von Gesundheitstechnologien dar; ein neues Paradigma, das sich auf die realen Erfahrungen mit konkurrierenden Therapien und deren Auswirkungen stützt, aber auch die Standards der normalen Wissenschaft und der grundlegenden Messung unterstützt. Dies ist natürlich eine natürliche Erweiterung der in der Produktentwicklung angewandten Standards, die durch einen Übergang von Wirksamkeitsansprüchen zu Wirksamkeitsergebnissen veranschaulicht werden, aber über eng definierte und fragwürdige Placebo-Vergleiche hinausgehen.

Dieser Vorschlag für einen Neuanfang steht im Gegensatz zu dem im Innovationsbericht beschriebenen evidenzbasierten Ansatz. Wir müssen den Glauben an die Nützlichkeit von Preiskalkulatoren und die Konzentration auf Daten aus klinischen Phase-IV-Studien zurückweisen. Genauso wie faire Preiskalkulationen irrelevant sind, so sind auch Behauptungen irrelevant, die auf annahmebasierten Simulationen beruhen, die langfristige und nicht bewertbare inkrementelle Kosten pro Lebensqualität (QALY) für die Kosteneffektivität erzeugen, obwohl das QALY ein mathematisch unmögliches Konstrukt ist [3,4]. Wir müssen auch Effizienzgrenzen und die von AMNOG und IQWiG angewandten Instrumente zur Unterstützung von Preis- und Zugangsempfehlungen ablehnen, ebenso wie generische Messgrößen mit mehreren Attributen, darunter nicht nur etablierte Messgrößen wie der EQ-5D-5L, sondern auch das vorgeschlagene Instrument EQ-Health and Wellbeing (EQ-HWB) [5,6]. Auch hier werden die Standards für normale Wissenschaft und grundlegende Messungen nicht eingehalten.

VOM MEME ZUM PARADIGMA

Ein Paradigmenwechsel in der HTA ist unabdingbar. Seit über 30 Jahren konzentrieren wir uns auf den Glauben an einen analytischen Rahmen, der Anfang 2022 durch die Veröffentlichung des Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement, einem Leitfaden für die Einreichung von simulierten erfundenen Behauptungen bei Fachzeitschriften, veranschaulicht wurde [7,8]. Es versteht sich von selbst, dass der CHEERS 2022-Leitfaden weder die Standards der normalen Wissenschaft noch grundlegende Messungen als Kriterien für die Bewertung von Therapien und Wertangaben erwähnt [2]. Das Versäumnis von CHEERS 2022, das Konzept der Evolution des objektiven Wissens zu unterstützen, ist der Grund, warum HTA am besten als Meme beschrieben werden kann, als ein sich selbst reproduzierendes Glaubenssystem, und nicht als ein Paradigma, das umgestoßen werden kann, wenn sich die Beweise für das Versagen eines Glaubenssystems häufen. Nicht nur CHEERS 2022 ist irrelevant, sondern auch die seit langem angewandte Methodik, die in dem führenden Lehrbuch von Drummond et al. beschrieben wird, ist ebenso überflüssig [9]. Zweifelsohne wird es Widerstand gegen eine solche grundlegende Änderung geben.

Das vorgeschlagene neue Startparadigma steht im Einklang mit der Forderung der deutschen Gesetzgebung, dass sich die Therapieauswahl und der Zugang zu Arzneimitteln an den Rahmenbedingungen des Patientennutzens orientieren: Verkürzung der Krankheitsdauer, Lebensverlängerung, Verringerung unerwünschter Nebenwirkungen und Verbesserung der Lebensqualität. Dies sind keine außergewöhnlichen Ziele, sondern solche, die durch die Bewertung krankheitsspezifischer Nutzensaussagen erreicht werden müssen, einschließlich der Kernnutzensaussagen, die von den Versicherern als relevant für den Zugang zu Arzneimitteln und die vorläufige Preisfestsetzung erachtet werden. Dies bedeutet ein neues Paradigma im HTA, das die Erfordernisse der Standards der normalen Wissenschaft und der grundlegenden Messung, des Fortschritts und der Entwicklung des objektiven Wissens anerkennt [10]. Ein Neuanfang, der darauf abzielt, einen kontinuierlichen Nutzen während der Patentlaufzeit eines Produkts zu gewährleisten, wobei der therapeutische Nutzen bei der Entdeckung neuer Fakten ständig neu bewertet wird. Der Innovationsbericht weist zwar auf den Bedarf an "neuen" klinischen Daten in Phase-IV-Studien hin, doch ist die Perspektive zu eng und zu zeitlich begrenzt; unser Schwerpunkt muss auf dem Lebenszyklus von Arzneimitteln und Produkten liegen.

Der Neubeginn weist auf die zentrale Rolle hin, die die Gesundheitssysteme, die Versicherer in Deutschland, in Zukunft spielen werden, die auf der Lösung der erforderlichen krankheitsspezifischen Kernwerte als Grundlage für vorläufige Preisvereinbarungen bestehen müssen. Ein Paradigma, bei dem die Hersteller verpflichtet sind, diese Forderungen im Rahmen ihrer laufenden Interaktion mit Patienten, Pflegekräften, Leistungserbringern und Versicherern zu erfüllen. Sobald die Lösung akzeptiert wurde, kann ein vorläufiger Preis vereinbart werden, wobei die Möglichkeit besteht, die gelösten Wertansprüche zu überprüfen, wenn sich die Marktbedingungen durch die Einführung neuer Produkte oder unerwünschte Ereignisse ändern. Die Hersteller müssen damit rechnen, dass die Preisgestaltung im Laufe des Produktlebenszyklus immer wieder in Frage gestellt wird, was sowohl zu Preissenkungen als auch zu Preiserhöhungen führen kann, wenn bestimmte Wertvorstellungen, die Preiserhöhungen rechtfertigen, erfüllt werden.

Die Preisfestsetzung ist vereinbart und kann nicht überprüft werden, ohne dass eine Einigung auf der Grundlage der Klärung von Wertansprüchen erzielt wird. Ansprüche, die auf früheren Studien beruhen, wie z. B. indirekte Vergleiche aus placebobasierten Phase-

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

III-Zulassungsstudien, schaffen lediglich die Voraussetzungen für die Replikation der Vergleichspräparate, einschließlich möglicher Protokolländerungen. Es handelt sich hierbei nicht um eine Vereinbarung für Replikationsstudien, die unnötig verzögert werden kann, sondern um eine, die in einem sinnvollen Zeitrahmen erfüllt werden muss. Dieselbe Anforderung gilt für alle Wertangaben, die im Rahmen der laufenden Überprüfung von Krankheitsgebieten und therapeutischen Klassen überarbeitet werden können. Selbstverständlich muss der Hersteller, sofern nichts anderes vereinbart wurde, alle Bewertungen von Wertangaben übernehmen, einschließlich der Zusammenstellung neuer Daten, falls erforderlich, um eine Auflösung der Wertangabe zu ermöglichen.

Es ist wichtig, darauf hinzuweisen, dass die Anforderungen an Wertansprüche sowohl für Standardtherapien als auch für Arzneimittel (für seltene Krankheiten und seltene Leiden) gelten, die im Rahmen einer beschleunigten Zulassung geprüft wurden; es gibt keine Ausnahmen [11]. Die ursprüngliche Entscheidung über den Markteintritt und ein vereinbarter vorläufiger Preis, der in der Regel für einen begrenzten Zeitraum gelten soll, der mit der Beilegung von Wertansprüchen (oder hochwertigen Ansprüchen, die einen Kernsatz umfassen) übereinstimmt, darf nur geklärte Ansprüche widerspiegeln, mit der Maßgabe, dass die klinische Basis nur allzu oft spärlich ist. Die ersten Anträge auf Erledigung von Ansprüchen müssen alle die Standards für die empirische Bewertung und grundlegende Messung erfüllen. Das bedeutet, wie unten erwähnt, dass alle Anträge die Standards für Intervall- (oder Verhältnis-) Messungen erfüllen müssen. Wenn entscheidende Phase-III-PRO-Studiendaten als Wirksamkeitsanspruch eingereicht werden, müssen diese im Hinblick auf die Anforderungen an ein Intervallmaß bewertet werden, wobei die Regel, dass der Anspruch den Standards der eindimensionalen Rasch-Modellierung entspricht, strikt anzuwenden ist [12].

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

MESSUNG

Wenn es ein Element des bestehenden HTA-Memes oder -Glaubenssystems gibt, das die Notwendigkeit unterstreicht, zu einem neuen Paradigma überzugehen, dann ist es das Versäumnis, die durch die Rasch-Messung gesetzten Grenzen für Ergebnisaussagen anzuerkennen: Das Versäumnis, einen seit über 60 Jahren anerkannten Standard anzuerkennen, wonach Beobachtungen immer ordinal sind, während die Messung in Intervallen erfolgen muss [13,14]. Subjektive Antworten oder das Zählen von Beobachtungen mit PRO-Instrumenten (z. B. Likert-Skalen) sind Ordinalskalen; sie können keine Grundlage für Therapieaussagen bilden. Aussagekräftige Messungen beruhen auf Intervallskalen. Die Rasch-Messung bietet die notwendigen und ausreichenden Mittel, um ordinale Zweifel in eindimensionale Intervallmaße sowohl für dichotome als auch für polytome Instrumente zu transformieren [6].

Die Beurteilung der Glaubwürdigkeit von Wertaussagen muss den Rasch-Intervallstandards entsprechen; andernfalls muss die Aussage zurückgewiesen werden. Die Rasch-Messung ist keine Option, sondern ein Muss. Dies gilt nicht nur für vorgeschlagene Wertaussagen, sondern auch für erste Zulassungsanträge, die auf bereits vorliegenden Phase-III-Ergebnissen beruhen. Ein Phase-III-Protokoll, das Outcome-Claims liefert, die nicht den Standards der Intervall- oder Rasch-Messung entsprechen, deren Ergebnisse aber vom IQWiG und AMNOG akzeptiert werden könnten, muss abgelehnt werden. Aus der Sicht des deutschen Arzneimittelzulassungsverfahrens ist diese Verkennung grundlegender messtechnischer Grenzen ein schwerwiegendes und unentschuldbares Versäumnis. Natürlich müssen wir mit Widerständen gegen eine solche grundlegende Änderung rechnen, obwohl die Rasch-Messung in der Pädagogik, der Psychologie und in einigen Bereichen des HTA seit vielen Jahrzehnten akzeptiert ist [15,16].

DER NEUSTART DER GESUNDHEITSTECHNOLOGIE-BEWERTUNG: GRUNDLEGENDE PRÄMISSEN

Im Mittelpunkt des vorgeschlagenen Neustarts im HTA steht die Einigung auf zentrale Nutzaussagen in Krankheitsbereichen und für Patientenzielgruppen. Diese Wertaussagen müssen zumindest für den Markteintritt eines neuen Produkts oder die erste Bewertung eines bestehenden Produkts Folgendes umfassen (i) klinische [instrumentengestützte] Aussagen; (ii) vom Patienten berichtete Ergebnisaussagen [Krankheitssymptome sowie Aspekte der Lebensqualität als latentes Konstrukt], (iii) Arzneimittelverwendung [Aufnahme bei Markteintritt und Compliance] und (iv) andere Auswirkungen auf die Ressourcennutzung [als Einheiten, nicht als Kosten]. Die drei Prämissen für Wertaussagen sind:

- Alle Wertaussagen müssen als einzelne Attribute vorgeschlagen werden, die den Standards der normalen Wissenschaft entsprechen: Glaubwürdigkeit, empirische Bewertung und Replikation
- Alle Wertaussagen müssen den Rasch-Standards für grundlegende Messungen entsprechen (Intervall- oder Verhältnismaße).
- Alle Wertaussagen müssen mit einem unterstützenden Protokoll für die Bewertung und Berichterstattung innerhalb eines sinnvollen Zeitrahmens vorgelegt werden.

Es muss klargestellt werden, dass diese drei Prämissen die Grundlage für alle Leitlinien bilden, die den Herstellern für die Einreichung von Arzneimitteln zur Verfügung gestellt werden. Einreichungen, die diese Standards nicht erfüllen, müssen abgelehnt werden. Dies bedeutet, dass eine Reihe von Instrumenten, die von Patienten berichtet werden (patient reported outcome, PRO), multiattributiv und krankheitsspezifisch, abgelehnt werden; sie erfüllen die Rasch-Standards nicht [17]. Ebenfalls ausgeschlossen werden annahmebasierte Simulationsmodelle mit nicht bewertbaren Angaben zur Kosteneffektivität, die Verwendung von Angaben zu qualitätsbereinigten Lebensjahren (QALY) und nicht bewertbare Konstrukte wie Effizienzgrenzen mit mehreren Merkmalen. PRO-Angaben mit einem einzigen Attribut müssen eindimensional sein und mindestens Intervall-Messeigenschaften aufweisen, die entweder durch die dichotome Rasch-Modellierung, das polytome Rasch-Ratingskalenmodell oder das Partial-Credit-Rasch-Modell erfasst werden [5]. Wenn Instrumente die Grundlage der Messung bilden, müssen sie Intervall- oder Verhältnismaße unterstützen. Bei der Inanspruchnahme von Arzneimitteln und anderen Ressourcen wird erwartet, dass sie den Ratio-Standards entsprechen (d. h. eine echte Null). Modelle werden nur dann akzeptiert, wenn sie kurze Laufzeiten haben und evaluierbare Aussagen unterstützen (z. B. Modelle zur Unterstützung von Aussagen über das Compliance-Verhalten). Die allgemeine Absicht besteht darin, die Kriterien anzuwenden, die HTA als Wissenschaft und nicht als Nicht-Wissenschaft oder Pseudowissenschaft rechtfertigen können; sie müssen Abgrenzungskriterien erfüllen [18].

Im Gegensatz zu den CHEERS 2022-Leitlinien für die Einreichung von auf Annahmen beruhenden, nicht auswertbaren Kosten-pro-QALY-Angaben bei Fachzeitschriften wurden kürzlich (fast zeitgleich mit CHEERS 2022) die Rasch Reporting Guidelines for Rehabilitation Research (RULER) veröffentlicht [19,20]. Es handelt sich dabei nicht um eine Neuerung; Studien, die über Rasch-Messungen im Zusammenhang mit

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Rehabilitationsergebnissen berichten, wurden erstmals 1988 veröffentlicht. Der Zweck von RULER ist es, eine von Experten begutachtete, evidenzbasierte und konsistente Anleitung für die Berichterstattung über Studien zu geben, die die Rasch-Messtheorie im Rehabilitationskontext anwenden, so dass es einheitliche Erwartungen gibt, wie die Forschung über die Bewertung von Rehabilitationsergebnissen zu schreiben und zu bewerten ist. Auch in Deutschland sollte die RULER-Vorlage in allen Krankheitsbereichen angewandt werden, wenn wir Rasch-basierte Aussagen zum Therapieerfolg bewerten wollen. Eine solche Verpflichtung der Entscheidungsträger sollte als wesentlicher Bestandteil eines Neubeginns in der Bewertung von Gesundheitstechnologien angesehen werden, der in die Einreichungsrichtlinien für Gesundheitssysteme aufgenommen wird.

PROTOKOLLE

Protokolle sind ein wesentlicher Bestandteil der Produktentwicklung. Sobald jedoch die Marktzulassung gesichert ist, wird die Anwendung von Protokollen auf die lange Bank geschoben. Dies ist ein entscheidender Fehler. Wenn wir die Bewertung von Wertangaben für ein Produkt während seines gesamten Lebenszyklus unterstützen wollen, müssen wir darauf bestehen, dass die Protokolle bei der Bewertung von Wertangaben im Vordergrund stehen. Dabei geht es nicht um die Validierung von Behauptungen, sondern um die Bereitstellung eines Rahmens für die Bewertung der Vorzüge von Wertaussagen. In Anerkennung der Asymmetrie zwischen Beweis und Widerlegung sollte ein Protokoll eine Wertbehauptung unterstützen, indem es diese Behauptung in einem Format präsentiert, das die Rolle der Falsifizierung von Behauptungen anerkennt; streng genommen sollte ein Protokoll die Wahrscheinlichkeit maximieren, dass eine Wertbehauptung falsifiziert wird [21].

In Anerkennung der Bedeutung der Falsifikation als Abgrenzungskriterium zwischen Wissenschaft und Nicht-Wissenschaft ist der Fokus auf Falsifikation die Grundlage für die vorläufige Akzeptanz von Nutzenversprechen über den Produktlebenszyklus. Dies unterscheidet den Neubeginn im HTA von der relativistischen Akzeptanz nicht bewertbarer Rezepturvorschläge und der Forderung nach einer fairen Preismodellierung durch feste Parameter und begrenzte Annahmemöglichkeiten. In allen Phasen der Marktzulassung und der weiteren Genehmigung, Platzierung und Preisgestaltung muss es eine starke und ausdrückliche Begründung oder Transparenz geben, um das Produkt zu unterstützen. Preisgestaltung und Zulassung sind immer vorläufig. Fortlaufende Protokolle über Wertangaben bilden die Grundlage für diese fortlaufende Neubewertung bzw. die Weiterentwicklung des objektiven Wissens. Dies sollte der Rahmen für ein dynamisches, evidenzbasiertes Preismodell sein, wie es in dem Bericht gefordert wird, und nicht für einen dynamischen Bewertungsrahmen, der eine evidenzbasierte Preisgestaltung unterstützen kann.

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

ZUGANG ZU DATEN

Ein merkwürdiges Versäumnis in dem Bericht ist die geringe Berücksichtigung des Bedarfs an Daten zur Unterstützung der Bewertung von Wertangaben. In den USA sind die Daten trotz der Fülle von Online-Daten mit umfassenden Datenbanken, die auf der Grundlage der bundesstaatlichen Meldevorschriften erstellt wurden, nach wie vor begrenzt. Die Protokolle müssen daher die Realitäten der verfügbaren und leicht zugänglichen Daten widerspiegeln und gleichzeitig Vorschläge machen, wie die Datenbeschränkungen überwunden werden können. Die Behauptung, es gäbe keine Daten, bedeutet nicht, dass diese nicht ohne weiteres verfügbar sind, sondern dass in der Praxis nicht auf die Datenbanken zugegriffen worden ist. Dennoch sind in Europa im Vergleich zu den USA nur wenige Daten zugänglich; man könnte argumentieren, dass die Berechnung des "fairen Preises" eine Ausrede für die begrenzten Daten zur Unterstützung von Protokollen und Wertangaben ist. Wenn dem so ist, muss dies zurückgewiesen werden. Das Erfinden von Daten, um simulierte, nicht bewertbare Behauptungen zu untermauern, ist keine Grundlage für die Akzeptanz von Wertbehauptungen; die Vorstellung von pauschalen Kostenwirksamkeitsaussagen ist inakzeptabel.

Wenn Krankenversicherer an vorderster Front bei der Einführung von Wertanspruchs-Verfahren stehen, dann haben sie eine Schlüsselrolle bei der Bereitstellung von Daten. Dies kann sowohl für die Nutzung von Arzneimitteln als auch für die Aufzeichnungen anderer direkter medizinischer Ressourcen gelten. Gleichzeitig liegt die letztendliche Verantwortung für die Erstellung von Wertangaben, die mit der Zugänglichkeit von Daten (einschließlich unterstützender neuer Daten) zur Bewertung dieser Angaben vereinbar sind, bei den Herstellern. Dies bedeutet jedoch nicht, dass sich ein Hersteller auf die Ausrede berufen kann, dass eine erforderliche Wertangabe nicht vorgeschlagen werden kann, weil die Daten nicht ohne weiteres verfügbar sind.

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

DATENERFASSUNG: REGISTER

Wenn das Ziel darin besteht, Daten zur Untermauerung von Wertaussagen für (i) klinische Endpunkte, (ii) von Patienten berichtete Ergebnisse, (iii) die Nutzung von Arzneimitteln und (iv) die Nutzung anderer Ressourcen zu erheben, dann ist ein Register nur bedingt geeignet, auch wenn es möglicherweise für (i) und (ii) relevant ist. Es gibt noch ein weiteres Problem: Wenn wir uns auf die Wirksamkeit und die realen Daten konzentrieren wollen, müssen wir das potenzielle Ausmaß der Beeinflussung des Patientenverhaltens durch die Datenerhebung minimieren (z. B. die Einhaltung der Vorschriften nach der Produkteinführung). Die Patienten sind sich dessen bewusst und werden ständig daran erinnert, dass sie Teil eines Registers sind, und stehen unter Druck, im Register zu bleiben, um langfristige Ergebnisse zu verfolgen und die Kosten für die weitere Rekrutierung zu minimieren. Register sind auch mit hohen Kosten verbunden, da die Verwaltung in der Regel akademischen Forschungszentren obliegt, deren Mitglieder sich eher auf die reine Forschung konzentrieren und nicht auf banalere Datenerfassungsaktivitäten zur Unterstützung von Preis- und Zugangsverhandlungen. Das Bestehen auf Registern ist wahrscheinlich ein großes Hindernis für die Akzeptanz des dynamischen Evidenzpreismodells.

In den USA, wo eine Fülle von Daten zur Verfügung steht, um die Nutzung von Arzneimitteln und Ressourcen nach Krankheiten und Therapiebereichen zu verfolgen, lassen sich die Wertansprüche für (iii) und (iv) leicht erfüllen, wobei der Schwerpunkt beim Neustart auf Wertansprüchen liegt, die in Ressourceneinheiten (mit Standardklassifikationen) und nicht in Kosten ausgedrückt werden. Von besonderem Interesse ist das Nutzungsverhalten von Medikamenten, die Regelbefolgung, nach der Produkteinführung. Wenn das durchschnittliche Regelbefolgungs-Profil (wie in den USA) 12-15 Monate beträgt, dann müssen die Wertangaben diesem Umstand Rechnung tragen, wobei Wertangaben für die Therapietreue (möglicherweise modelliert) die Wertangaben für erfolgreiche Ergebnisse ergänzen.

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

KERNWERTFORDERUNGEN

Die Gesundheitssysteme, einschließlich der Versicherer, sollten die Initiative ergreifen und Kernwertangaben für bestimmte Krankheitsbereiche und Patientenpopulationen einführen. Modellerte Aussagen (z. B. zur Kosteneffektivität) sind nicht zulässig; alle Aussagen müssen sich auf einzelne Attribute beziehen, grundlegende Messstandards erfüllen und natürlich in einem sinnvollen Zeitrahmen empirisch evaluierbar sein. Sie sind so konzipiert, dass sie sich mit Fragen der fehlenden Evidenz befassen und für alle Zielpatientenpopulationen spezifiziert werden (z. B. reale Bewertungen von Vergleichstherapien), müssen aber im Kontext der Entdeckung neuer Fakten und der Verfolgung der Auswirkungen der Therapie in einer möglicherweise sehr variablen Patientenpopulation gesehen werden, wobei Therapieabbrüche und neue Marktteilnehmer während der Patentlaufzeit des Produkts mit sich entwickelnden Vergleichstherapien erfasst werden.

Das ergebnisorientierte Contracting als Verhandlungsgrundlage sollte nicht vernachlässigt werden. Die Protokolle müssen so gestaltet werden, dass sie das Contracting mit dem Gesundheitssystem oder dem Versicherer unterstützen, der die Struktur und den Inhalt der Wertangabenprotokolle diktiert; die Protokolle sorgen für Transparenz bei den Angaben als Intervall- (wenn nicht Verhältnis-) Maße, die das Spektrum der klassischen statistischen Verfahren zur Bewertung ihres Wertes für die Formulierungsentscheidungen unterstützen. Es liegt in der Verantwortung des Herstellers, Vorschläge für neue Produkte zu unterbreiten, die mit dieser Anforderung in Einklang stehen. Es besteht die Möglichkeit, ergänzende Angaben zu machen, aber für die Kernwertangaben müssen detaillierte Protokolle für die Bewertung der Angaben und für bestehende Produkte, Krankheitsgebiete und therapeutische Klassen geprüft werden.

Die endgültige Entscheidung über Kernwertforderungen und die zugehörigen Protokolle muss natürlich vom Gesundheitssystem oder von der Krankenkasse getroffen werden und nicht von Agenturen wie dem AMNOG oder dem IQWiG mit den unvermeidlichen bürokratischen Verzögerungen bei dem Versuch, einen nationalen Konsens über Kernwertforderungen für alle Krankheitsgruppen und Patientenzielgruppen zu erzielen, der den erforderlichen Messstandards entspricht. Das International Consortium for Health Outcomes Measurement (ICHOM: www.ichom.org) hat zwar Kernwertangaben für Krankheitsgruppen vorgeschlagen, doch wurden diese vorgeschlagenen Ergebnismessungen nicht auf ihre Messbarkeit hin überprüft [22]. Es ist sinnvoller, wenn das Gesundheitssystem oder der Versicherer, möglicherweise in Verbindung mit den Vorschlägen des ICHOM, die erforderlichen Maßnahmen von Fall zu Fall festlegt, wobei vorrangige Krankheitsbereiche zuerst behandelt werden können.

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

ÜBERGREIFENDE LEITLINIEN

Der erste wichtige Schritt ist die Entwicklung von Leitlinien zur Festlegung von Mindeststandards für die Einreichung von Wertangaben. Dabei sollte es sich um einen "Bottom-up"-Ansatz handeln, der unnötige Eingriffe der nationalen und bundesstaatlichen Behörden ausschließt. Er muss nicht nur die erforderlichen Beweisstandards für Wertangaben klarstellen, sondern auch Angaben, die nicht akzeptiert werden. Dies bedeutet, dass Angaben zu den von den Patienten berichteten Ergebnissen abgelehnt werden, die nicht den Rasch-Messstandards entsprechen (z. B. generische Angaben mit mehreren Attributen auf der Grundlage der EQ-5D-3L-5L-Instrumente), und dass sichergestellt wird, dass instrumentengestützte klinische Angaben sowie Angaben zum Verbrauch von Arzneimitteln und anderen Ressourcen den erforderlichen Intervall- (und Verhältnis-) Messstandards entsprechen.

In den Leitlinien sollte die Grundlage festgelegt werden, auf der Wertangaben vorgeschlagen, bewertet und überwacht werden. Jede Wertbehauptung muss die Standards der normalen Wissenschaft für Glaubwürdigkeit, Bewertung und Replikation über den Produktlebenszyklus erfüllen, aber in aussagekräftigen Zeiträumen als Teil eines Wertbehauptungsprofilpakets gemeldet werden. Zusammengesetzte Angaben zur pauschalen Kosteneffektivität oder zum Kosten-Nutzen-Verhältnis sind inakzeptabel, ebenso wie andere Endpunktangaben (Bericht S. 91), denen es an auswertbarem empirischem Inhalt fehlt (z. B. Angaben zur Gesamtüberlebenszeit), oder zusammengesetzte Angaben zur gesundheitsbezogenen Lebensqualität (HRQoL), die nicht den Standards für Eindimensionalität entsprechen. Dies schafft die Voraussetzungen für detaillierte krankheits- oder zielgruppenspezifische Kernwertaussagen, zusammen mit möglichen ergänzenden Wertaussagen.

Alle Wertaussagen und die sie unterstützenden Protokolle müssen vereinbart werden, bevor die Verhandlungen über die Rezeptur fortgesetzt werden können. In vielen Fällen wird die akzeptable Evidenzbasis, die den erforderlichen Standards entspricht, dünn sein. Dies bedeutet nicht, dass Entscheidungen beiseitegeschoben werden, sondern dass die Wertaussagen darauf ausgerichtet sein können, Beweislücken zu schließen (z. B. vergleichende klinische Wirksamkeit), wobei die Vorlage einen vorläufigen Preis vorbehaltlich der Bewertung der Wertaussagen unterstützt. Dadurch wird der Schwerpunkt auf Angaben gelegt, die in einem relativ kurzen Zeitrahmen sinnvoll bewertet werden können.

DER FAIRE PREISKALKULATOR

Der vorgeschlagene Rechner für faire Preise ist das Gegenteil von dem, was man sich unter der Rolle von Beweisen bei der Unterstützung von Formulierungsentscheidungen vorstellt. In der Wirtschaftswissenschaft ist ein fairer Preis ein Preis, der zwischen Käufer und Verkäufer vereinbart wurde; er ist der ausgehandelte Preis. Natürlich kann es Abweichungen geben, Preisdiskriminierung, um Nachfrageunterschiede zwischen Marktsegmenten auszugleichen und den Gewinn zu maximieren, sowie Preissubventionen für ältere oder benachteiligte Bevölkerungsgruppen. Für keine dieser Optionen ist ein Kalkulator für faire Preise erforderlich; die Konzentration auf die Marktreaktion ist das Signal (z. B. bei der Preisgestaltung für Flugsitze).

Fantasy HTA verfolgt mit dem Vorschlag für einen Algorithmus für faire Preise einen anderen Ansatz. Der vorgeschlagene Ansatz besteht darin, sich auf die Rolle zu konzentrieren, Evidenzlücken nicht zu ignorieren, sondern einen Rahmen zu schaffen, um den Nutzen und die Auswirkungen eines Produkts auf der Grundlage ausgewählter Parameter und unterstützender Annahmen (keine Wahl) zu bewerten. Der Rechner für den fairen Preis ist ein absurder Ansatz. Der faire Preis wird so berechnet, dass er eine Rendite von 8 % der Forschungs- und Entwicklungskosten ergibt, wobei der Innovationsgrad angesichts des therapeutischen Werts des Produkts möglicherweise angepasst wird. In dem Maße, in dem die eingebauten Annahmen variiert werden können, besteht die Möglichkeit, dass die Hersteller versuchen, den größtmöglichen Wert aus der Berechnung herauszuholen, oder, was wahrscheinlicher ist, sie von vornherein als nichts weiter als grobes Rätselraten und irreführende Zeitverschwendung ablehnen.

Der Rechner ist ein absurdes Vehikel, das jedoch mit der relativistischen Position der aktuellen Standards im HTA-Bereich übereinstimmt, wo Preisempfehlungen auf der Grundlage von annahmeorientierten Modellsimulationen in vielen Gesundheitssystemen mit nur einem Kostenträger der Bezugspunkt sind. Auch wenn diese Modelle den Anschein erwecken, auf "realistischen" Annahmen zu beruhen, werden sie zunehmend als Fehlschlag angesehen, da keine der Ergebnisaussagen empirisch evaluierbar ist, ganz abgesehen von der Tatsache, dass die inkrementellen Kosten pro QALY nicht den Standards für grundlegende Messungen entsprechen.

Der Bericht kommt zu dem Schluss, dass der Rechner "als nützliche Ergänzung" zu bewerten ist, wenn auch möglicherweise überarbeitet und die Ergebnisse "als Lernsystem gewertet" werden. Was das Lernziel sein soll, abgesehen von der hier dargestellten negativen Sichtweise, ist nicht klar: Der Bericht kommt zu dem Schluss: "Dennoch kann der Preisansatz als Aufforderung an die pharmazeutischen Unternehmen verstanden werden, mehr Transparenz bei der Preisgestaltung zu bieten, denn nur so können Preisverhandlungen auf einer realistischeren Basis als bisher geführt werden". Dies ist, offen gesagt, Wunschdenken; die Hersteller und die Entscheidungsträger des Gesundheitssystems wären gut beraten, sich von einem solch simplen und naiven, von Annahmen geprägten Rahmen zu lösen. Jede Transparenz ist entweder erzwungen oder erfunden. Auch wenn die Evidenz bei der Produkteinführung in der Regel begrenzt ist, besteht die Antwort nicht darin, einen fiktiven, auf Annahmen beruhenden fairen Preis oder auf Annahmen beruhende Preisempfehlungen aus einem Checklisten-Algorithmus zu

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

erfinden, sondern ein Forschungsprogramm zu unterstützen, um vereinbarte Wertangaben als Grundlage für laufende vorläufige Verhandlungen zu bewerten.

Leider gibt es keine schnelle Lösung, um den mythischen Goldstandard eines "fairen" Preises zu erreichen. Der Rechner mit seinen festen Parametern und seiner begrenzten Auswahl an Annahmen ist eine unnötige, aber potenziell unterhaltsame Ablenkung. Er hat nicht nur die Voraussetzungen für die Anfechtung seiner Struktur und die Forderung nach mehr Flexibilität bei der Wahl der Parameter und Annahmen geschaffen, sondern auch die Schleusen für eine Vielzahl konkurrierender Fair-Price-Rechner geöffnet, die mit ihm konkurrieren. Es ist recht einfach, konkurrierende und "bessere" Fair-Price-Rechner zu entwickeln, wobei das von den Sponsoren des Rechners gesponserte Produkt immer einen "fairen Preis" erzielt. Dies ist bei zahlreichen Gelegenheiten mit veröffentlichten simulierten inkrementellen Kosten-pro-QALY-Behauptungen geschehen, die das Produkt des Sponsors zu einem behaupteten "kosteneffektiven" Preis begünstigen; es besteht kein Zweifel daran, dass wir den gleichen Ansturm auf die Veröffentlichung von Behauptungen über einen fairen Preis bei neuen und bestehenden Produkten erleben werden; die Tür für ein voreingenommenes Reverse Engineering ist weit offen [23].

Schließlich ist nicht klar, was der Algorithmus, der einen Vorschlag für einen "fairen Preis" unterstützt, erreichen soll. Es besteht sicherlich nicht die Absicht, die Standards der normalen Wissenschaft anzuerkennen, da der "faire Preis" nur eine Ordinalzahl ist. Als solcher kann er keine Angaben zum "Grad der Abweichung" vom aktuellen Marktpreis machen; wir können nur sagen, dass er höher oder niedriger ist, aber nicht um wie viel. Wenn es die Absicht war, einen Standard mit Intervall- (oder Verhältnis-) Eigenschaften zu schaffen, hätte dies von Anfang an deutlich gemacht werden müssen. Dies bedeutet auch, dass der Versuch, alternative Algorithmen zu entwickeln und die verschiedenen "fairen Preise" zu vergleichen, unmöglich ist, weil man keine Ordinalwerte vergleichen kann, um die Vorzüge eines Standards gegenüber einem anderen zu argumentieren. Anstatt als nützliche Ergänzung bei Preisverhandlungen angesehen zu werden, ist die Befürwortung eines "fairen Preises" am besten als vermeidbares Ablenkungsmanöver zu betrachten.

AUF DER SUCHE NACH ZUKÜNFTIGER EVIDENZ

Eines der Merkmale der deutschen, wenn nicht gar der europäischen Ansichten über HTA ist die Überzeugung, dass nationale Kommissionen oder andere bürokratische Einrichtungen, die jahrelange Ausschusssitzungen und politischen Kuhhandel (und Weinproben in Restaurants) erfordern, eine zentrale Rolle bei der Unterstützung von HTA spielen sollten. Aus amerikanischer Sicht scheint dies ein Overkill zu sein, eine unnötige Belastung. Die Position scheint zu sein, dass wir nicht darauf vertrauen können, dass die Hersteller in den Verhandlungen ehrlich sind, und dass die Regierung deshalb eingreifen muss. Im ersten Teil dieses Kommentars wurde eine Lösung vorgestellt, bei der sich die Parteien bei der Markteinführung eines Produkts und der vorläufigen Preisfestsetzung auf Wertansprüche einigen, um Evidenzlücken zu schließen und ein langfristiges Engagement einzugehen, das den finanziellen Zwängen und vor allem den Interessen von Patienten, Pflegepersonal und Anbietern am besten gerecht wird. Wertansprüche und die damit verbundenen Protokolle machen Transparenz erforderlich. Alle Behauptungen müssen nicht nur glaubwürdig, bewertbar und reproduzierbar sein und den Standards der normalen Wissenschaft entsprechen (die der Produktentwicklung zugrunde liegen), sondern auch sicherstellen, dass alle Behauptungen die Standards für grundlegende Beweise erfüllen. Dies ist keine Raketenwissenschaft und lässt sich leicht umsetzen. Die Ergebnisse und Protokolle können kaum als absichtlich undurchsichtig angesehen werden.

Der Bericht geht auf die Notwendigkeit ein, Evidenzlücken zu schließen, und schlägt vor, die Versorgungsforschung einzuführen, um die Verbesserung der Versorgung zu unterstützen, indem der Transfer von Therapieinnovationen aus klinischen Studien in die Wirksamkeitsforschung evaluiert und gefördert wird. Während das Konzept eines solchen Transfers von der Wirksamkeit in die Effektivität nicht neu ist, bietet der Bericht nur wenige praktische Vorschläge, wie dies geschehen kann, während gleichzeitig das Gebot der Einhaltung von Standards für die normale Wissenschaft bei der Effektivität beachtet wird - eine kritische Rolle von Standards in der normalen Wissenschaft und der grundlegenden Messung. Es reicht nicht aus, über Evidenzlücken zu sprechen; es muss anerkannt werden, dass Evidenzlücken in Wertaussagen übersetzt werden müssen, die den erforderlichen Standards entsprechen.

Der Schwerpunkt des Berichts liegt im klinischen Bereich. Er plädiert für eine Clearingstelle oder Kommission, die die Forschung und Zulassung "auf der Grundlage von Regeln und Kriterien" koordiniert, die im Voraus vereinbart werden, um Nutzen und Risiken innovativer Therapien zu bewerten. Bei Studien, die nicht im Interesse der Pharmahersteller liegen, müssen die Krankenkassen "in erster Linie die Studien initiieren", wobei sich die Hersteller an den Studienkosten beteiligen. Dies ist das hier vorgebrachte Argument, aber die Replikation klinischer Studien in einer realen Behandlungsumgebung ist nur eine der möglichen Kernforderungen für bestimmte Krankheitsbereiche.

Diese Empfehlung für eine Clearingstelle oder Kommission ist nur dann sinnvoll, wenn es Standards gibt, die einen Rahmen für die Erstellung und Bewertung von Wertangaben in einer realen Behandlungsumgebung bieten. Wir wissen, wie diese Standards in Bezug auf normale Wissenschaft und grundlegende Messungen aussehen sollten, und wir wissen auch, dass Protokolle zur Unterstützung der Bewertung erforderlich sind. Wertangaben, einschließlich vereinbarter Kernwertangaben für Krankheitsbereiche und Zielpatientengruppen, können eine Reihe möglicher Messgrößen für klinische Angaben, vom Patienten gemeldete Ergebnisse, den Arzneimittelverbrauch und andere

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Ressourcenauswirkungen umfassen. Mit anderen Worten: Es muss eine Einigung über den analytischen Rahmen oder die Leitlinien zur Unterstützung der Bewertung von Wertangaben über den Lebenszyklus eines Produkts erzielt werden. Natürlich sollte dies nicht das Vorrecht der Regierung sein (das ist ein einfacher Weg, um nichts zu erreichen), die über festgefahrene Interessen verfügt; die offensichtlichen Kandidaten sind die Krankenkassen.

SCHLUSSFOLGERUNGEN

Der vorliegende Innovationsbericht 2021 ist ein enttäuschendes Dokument. Erstens befürwortet er die Anwendung und Weiterentwicklung eines "fairen Preises" oder vieler konkurrierender Kalkulatoren, die keinerlei Chance haben, jemals vereinbart und angewandt zu werden, nicht nur innerhalb Deutschlands von den verschiedenen Akteuren des Gesundheitssystems, sondern auch im weiteren europäischen Umfeld. Es ist ein Nullsummenspiel, ein vermeidbares Ablenkungsmanöver. Der Begriff des fairen Preises mag ein Ideal sein, aber er hat keine messbare praktische Anwendung. Zweitens verkennt der Bericht bei seinem Vorschlag für einen Rahmen zur Behebung von Evidenzlücken die entscheidende Rolle eines übergreifenden Leitfadens für die Bewertung von krankheits- oder zielgruppenspezifischen Nutzenversprechen, der die erforderlichen Standards für die Bewertung einzelner Attribute festlegt, um ein Kernprofil von Nutzenversprechen zu erstellen; ein Profil, das während des Lebenszyklus des Produkts neu bewertet werden kann. Die Rolle grundlegender Messungen zur Untermauerung der Behauptungen wird nicht anerkannt, ganz zu schweigen von der Verpflichtung, Fortschritte beim Verständnis der langfristigen Auswirkungen eines Produkts auf Patienten und Pflegepersonal zu erzielen. Karl Poppers Konzept der Entwicklung objektiven Wissens ist nicht nur nicht vorhanden, sondern liegt auch außerhalb des Bezugsrahmens des Berichts [10].

Es ist dringend erforderlich, dass die Versicherer in Deutschland eine stärkere Kontrolle über Produkte und Geräte ausüben. Dies betrifft nicht nur die Evidenzstandards für neue Produkte, sondern auch das Management bestehender Produkte während ihres Lebenszyklus bzw. ihrer Patentlaufzeit. Der Fokus darf nicht nur auf dem Preis liegen, sondern auch auf den Ressourcen, die für eine effektive Produktbereitstellung erforderlich sind, wobei der Stückpreis im Vergleich zu anderen Ressourcen zur Unterstützung von Patienten und Leistungserbringern nur eine untergeordnete Rolle spielen kann. Rückmeldungen in einem aussagekräftigen Zeitrahmen sind von entscheidender Bedeutung; dies setzt jedoch voraus, dass eine Einigung über die Wertaussagen zur Krankheit oder zur Patienten-Zielgruppe besteht. Alle Wertaussagen müssen empirisch bewertbar sein und in ihrer Relevanz immer wieder neu bewertet werden. Sobald die wichtigsten krankheitsspezifischen Wertaussagen feststehen, obliegt es dem Hersteller, Entwürfe von Protokollen für deren Bewertung vorzulegen und die erforderlichen Ressourcen für die Berichterstattung über diese Aussagen bereitzustellen. Sicherlich kann dies vom Versicherer durch mögliche Informationen über die Ressourcennutzung und die damit verbundenen Stückkosten, einschließlich Produktaufnahme und -abbruch, und die Auswirkungen der Compliance-Muster auf die Wertaussagen zur Produktwirksamkeit erleichtert werden.

Zwei Schritte sind von entscheidender Bedeutung: Erstens muss der Versicherer klarstellen, dass nur glaubwürdige, bewertbare und reproduzierbare Wertansprüche akzeptiert werden, die den Standards für grundlegende Nachweise mit vereinbarten Bewertungsprotokollen entsprechen, und zweitens, dass Wertansprüche spezifisch für die Einführung eines Produkts oder Geräts bei einer bestimmten Patientengruppe sein müssen. Der erste Schritt erfordert einen übergreifenden Leitfaden, in dem die Rolle von Wertansprüchen und ihre Anwendung im Lebenszyklus des Produkts dargelegt wird. Der zweite Schritt erfordert eine Vereinbarung zwischen den Herstellern über die Wertansprüche, die zur Unterstützung der Produktbewertung und -verfolgung erforderlich sind, sowohl zur Unterstützung der

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

laufenden Überprüfungen von Krankheitsgebieten und therapeutischen Klassen als auch für mögliche ergebnisorientierte Verträge.

Die Frage eines "fairen" Preises muss zwischen den Parteien ausgehandelt werden; es ist sicherlich möglich, dies abstrakt zu betrachten, aber die Frage ist pragmatisch. Im Laufe des Produktlebenszyklus lässt er sich am besten als ein bewegliches Ziel beschreiben, da sich die Beweise für die vergleichende Wirksamkeit des Produkts, die Einhaltung der Vorschriften und die Einführung neuer Produkte häufen; hinzu kommen Wertangaben, die den erforderlichen Beweisstandards nicht genügen. Alle Wertangaben sind vorläufig und müssen fälschungssicher formuliert werden. Es gibt keine Abkürzungen zu einem "kosteneffizienten" fairen Preis; dies ist ein unmögliches Ziel. Kosten-pro-QALY-Schwellenwerte sind im Hinblick auf die Standards der grundlegenden Messung bedeutungslos und bieten denjenigen, die nicht bewertbare, auf Annahmen beruhende Modellaussagen rückgängig machen wollen, eine offene Tür. Alles, was getan werden kann, ist sicherzustellen, dass die Verhandlungen zwischen den Parteien zur Festlegung eines Einheitspreises und möglicher Zugangsbeschränkungen auf einer soliden Evidenzbasis beruhen müssen, die sich auf reale Fakten stützt. Wertbehauptungen bilden die erforderliche Evidenzbasis, wobei ein Profil der Ergebnisse der Wertbehauptungen zur Unterstützung der ersten und laufenden Verhandlungen vorgelegt wird.

Die hier vorgeschlagenen Standards stimmen mit dem im Bericht genannten Ziel überein, "Verhandlungsmöglichkeiten zwischen Krankenkassen und Herstellern zu fördern" (S. 27). Was uns trennt, sind die Evidenzstandards zur Unterstützung der Verhandlungen, insbesondere die Idee eines fairen Preiskalkulators. Die erforderlichen Evidenzstandards sind nun anerkannt und der Weg für ihre Einführung ist klar.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

- [1] Glaeske G. Innovations Report 2021. University of Bremen. SOCIUM, Mary-Somerville Str. 5, 28359 Bremen
- [2] Langley P. Nothing to Cheer About: Endorsing Imaginary Economic Evaluations and Value Claims with CHEERS 22 [version 1; peer review: 2 approved]. *F1000Research* 2022, **11**:248 (<https://doi.org/10.12688/f1000research.109389.1>)
- [3] Langley P. The Great I-QALY Disaster. *InovPharm*. 2020; 11(3): No 7
- [4] Langley PC and McKenna SP. Measurement, modeling and QALYs [version 1; peer review: 2 approved]. *F1000Research* 2020, **9**:1048 (<https://doi.org/10.12688/f1000research.25039.1>)
- [5] McKenna S, Heaney A. Composite outcome measurement in clinical research: the triumph of illusion over reality. *J Med Econ*. 2020;23(10):1196-1204
- [6] Brazier J, Peasgood T, Mukuria C et al. The EQ-HWB: Overview of the development of a measure of health and wellbeing and key results. *ValueHealth*. 2022;25(4):482-491
- [7] Husereau D, Drummond M, Augustovski F et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 22) Statement: Updated reporting guidance for health economic evaluations. *ValueHealth*. 2022;25(1):3-9
- [8] Husereau D, Drummond M, Augustovski F et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 2022 Explanation and Elaboration: A Report of the ISPOR CHEERS II Good Practices Task Force. *ValueHealth*. 2022;25(1):10-31
- [9] Drummond M, Sculpher M, Claxton K et al. *Methods for the Economic Evaluation of Health Care Programmes* (4th Ed.). New York: Oxford University Press, 2015
- [10] Popper K. *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press, 1972
- [11] Langley P. Formulary Submissions: Value Claims, Protocols and Outcomes Based Contracting in Rare Disease. *InovPharm*. 2023;13(3): No. 9
- [12] Bond T, Yan Z, Heene M. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 4th Ed. New York: Routledge, 2021
- [13] Andrich D, Marais I. *A Course in Rasch Measurement Theory: Measuring the Educational, social and Health Sciences*. Singapore, Springer: 2019
- [14] Wright B, Linacre J. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989; 70(12):857-60
- [15] Rasch G. *Probabilistic Models for some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research, 1960.
- [16] Andrich D. Understanding resistance to the data-model relationship in Rasch's paradigm: a reflection for the next generation. *J Appl Meas*. 2002;3(3):325-59

- [17] McKenna S, Heaney A, Langley P. Fundamental Outcome Measurement: Selecting Patient Reported Outcome Instruments and Interpreting the Data they Produce. *InovPharm*. 2021; 12(2): No. 17
- [18] Pigliucci M. Nonsense on Stilts: How to tell science from bunk. Chicago: University of Chicago Press, 2010
- [19] Mallinson T, Kozlowski A, Johnston M et al. Rasch Reporting Guidelines for Rehabilitation Research (RULER): the Ruler statement. *Arch Phys Med Rehab*. 2022;103:1477-86
- [20] Van de Winckel A, Kozlowski A, Johnston M et al. Reporting Guideline for RULER: Rasch Reporting Guideline for Rehabilitation Research: Explanation and Elaboration. *Arch Phys Med Rehab*. 2022;103:1487-98
- [21] Magee B. Popper. London: Fontana, 1974
- [22] Langley P. Evidentiary Standards for Patient-Centered Core Impact (PC-CIS) Value Claims. *InovPharm*. 2022;13(3): No. 15
- [23] Langley P. Facilitating bias in cost-effectiveness analysis: CHEERS 2022 and the creation of assumption-driven imaginary value claims in health technology assessment [version 1; peer review: 3 approved]. *F1000Research* 2022, **11**:993 (<https://doi.org/10.12688/f1000research.123709.1>)

A.1.29 Merck Healthcare Germany GmbH

Autorinnen und Autoren

- Gaupel, Ann-Christin
- Osowski, Ulrike

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Dr. Ann-Christin Gaupel
Dr. Ulrike Osowski
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Merck Healthcare Germany GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 2.1.3.2 (S. 20f.)	<u>Anmerkung:</u> In Abschnitt 2.1.3.2 stellt das IQWiG den Ablauf der Erstellung eines AbD-Konzepts dar. Während der Informationsbeschaffung werden externe Sachverständige, Betreiber bestehender Datenplattformen sowie Patientenorganisationen kontaktiert. Ein Einbezug der pharmazeutischen Unternehmen zu diesem Zeitpunkt ist wünschenswert und sinnvoll, da diese über Expertise im Indikationsgebiet sowie zu Registern besitzen, mit Registerbetreibern ggf. schon eine Zusammenarbeit angestoßen haben und somit die Konzepterstellung zielführend unterstützen können. <u>Vorgeschlagene Änderung:</u> Die betroffenen pharmazeutischen Unternehmer sollten als weitere Informationsquelle neben externen Sachverständigen, Betreibern geeigneter Datenplattformen sowie Patientenorganisationen aufgeführt werden: „Als Teil der Informationsbeschaffung werden zudem der pharmazeutische Unternehmer und die Betreiber bereits bestehender Datenplattformen kontaktiert, die für die Durchführung der geplanten AbD nach erster Einschätzung potenziell geeignet sind (insbesondere Indikationsregister).“ Entsprechend sollte auch die Abbildung erweitert werden.
Abschnitt 2.1.3.3 (S. 21f.)	<u>Anmerkung:</u> In Abschnitt 2.1.3.3 stellt das IQWiG den Ablauf der Recherche nach Indikationsregistern dar. Die Informationsbeschaffung erfolgt in verschiedenen, nicht näher spezifizierten Informationsquellen. Wie in der Anmerkung zu Abschnitt 2.1.3.2 erwähnt, besitzen die

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	pharmazeutischen Unternehmer tiefgehendes Wissen zu Indikationen und Indikationsregistern bzw. sind eventuell bereits an Planung und Aufbau eines Registers beteiligt. <u>Vorgeschlagene Änderung:</u> Bei der vom IQWiG definierten Recherche nach Indikationsregistern durch eine „fokussierte Informationsbeschaffung in verschiedenen Informationsquellen“ sollten die pharmazeutischen Unternehmen berücksichtigt werden.
Abschnitt 3.3.4 (S. 66-68)	<u>Anmerkung:</u> In Abschnitt 3.3.4 erläutert das IQWiG das Ziel und die Inhalte des Konzepts zur anwendungsbegleitenden Datenerhebung (AbD). Die Basis zu den Ausführungen zur anwendungsbegleitenden Datenerhebung in diesem und weiteren Abschnitten stellt der IQWiG Rapid Report A19-43 „Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung“ dar. Anhand des Konzepts sollen die methodischen Anforderungen dargestellt sowie geprüft werden, „ob die geplante AbD prinzipiell realisierbar ist“. Zu berücksichtigen sind bezüglich der Realisierbarkeit die oft geringen Patientenzahlen. Pharmazeutische Unternehmen entwickeln Medikamente z.B. auf Basis von genetischen Alterationen für immer spezifischere Indikationsgebiete und entsprechend kleine Patientenpopulationen. Dementsprechend soll die Umsetzbarkeit einer solchen AbD im Vorfeld mitberücksichtigt werden. Desweiteren sind die Anforderungen an die Studiendurchführung und Datenqualität sowie die Effektstärke „in einem Bereich von 2 bis 5 für das relative Risiko ... bezogen auf die untere bzw. obere Konfidenzintervallgrenze“ unter Adjustierung von Confoundern im Rahmen der AbD sehr hoch. Diese hohen Anforderungen unter Berücksichtigung geringer Fallzahlen stellen die Realisierbarkeit und die Eignung der AbD zum Nachweis eines Zusatznutzens infrage. <u>Vorgeschlagene Änderung:</u> Grundsätzlich sollte fortan die Möglichkeit der frühzeitigen Stellungnahme zu weitgehenden Ausarbeitungen wie der des IQWiG Rapid Reports A19-43 des IQWiG bestehen, um die wissenschaftliche Diskussion aller Beteiligten zu fördern. Die Realisierbarkeit einer AbD und ihre Verwertbarkeit für die frühe Nutzenbewertung unter Berücksichtigung der zu erwartenden Fallzahlen sowie der Effektstärke sollte stärker in den Fokus gesetzt werden. Mit Bezug auf die Effektstärke sollten sich die Schwellenwerte an den allgemein gültigen Schwellenwerten der Nutzenbewertung orientieren, da bereits eine Adjustierung für

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Confounder erfolgt ist. Die Anforderung eines Effekts im Bereich von 2 bis 5 bezogen auf die untere bzw. obere Konfidenzintervallgrenze zusätzlich zur Confounder-Adjustierung sollte fallen gelassen werden.
Abschnitt 9.1.1 (S. 159)	<u>Anmerkung:</u> Das IQWiG hat den Abschnitt zu den Kriterien für den Einschluss von Studien neu formuliert. Aus dem Text geht nicht eindeutig hervor, dass Studien für die Nutzenbewertung herangezogen werden können, wenn weniger als 80% der Patientinnen und Patienten das Einschlusskriterium bezüglich der Prüfintervention und Vergleichsintervention bzw. der Population erfüllen, wenn Analysen der relevanten Teilpopulation vorliegen oder hinreichend sicher plausibel oder nachgewiesen ist, dass die Ergebnisse auf die Zielpopulation anwendbar sind. <u>Vorgeschlagene Änderung:</u> Es sollte eine Präzisierung ähnlich der des Methodenpapier 6.1 erfolgen, dass Studien für die Nutzenbewertung herangezogen werden können, bei denen die Einschlusskriterien von mindestens 80% der Studienpopulation erfüllt sind. Studien, bei denen dies nicht der Fall ist, können herangezogen werden, wenn Analysen der relevanten Teilpopulation vorliegen oder hinreichend sicher plausibel oder nachgewiesen ist, dass die Ergebnisse auf die Zielpopulation anwendbar sind. Es ist nicht nachzuvollziehen, anhand welcher Kriterien das IQWiG über die Berücksichtigung einer Studie bei Informationen zu potentiellen Effektmodifikationen entscheiden wird. Daher sollte dieser Satz wieder entfernt werden.
Abschnitt 9.3.2 (S. 170)	<u>Anmerkung:</u> In Abschnitt 9.3.2 Statistische Signifikanz ergänzt das IQWiG den Satz „Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an“ um den Teilsatz „z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist“. Verschiedene Operationalisierungen eines Endpunkts sind gegebenenfalls aus wissenschaftlichen oder medizinischen Gründen sinnvoll und bieten als unterstützende, ergänzende Analysen oder Sensitivitätsanalysen einen Mehrwert. Die automatische Adjustierung für multiples Testen für die Bewertung von einem Endpunkt mit verschiedenen Operationalisierungen ist demnach nicht sinnvoll. <u>Vorgeschlagene Änderung:</u> Im Methodenpapier des IQWiG sollte erläutert werden, dass eine Darstellung verschiedener Operationalisierungen als ergänzende Analysen und Sensitivitätsanalysen nicht zur Adjustierung für

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	multiples Testen führt. Der Nebensatz „z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist“ sollte dementsprechend gestrichen werden.
Abschnitt 9.3.12 (S. 192)	<u>Anmerkung:</u> Das IQWiG ergänzt den Abschnitt 9.3.12 „Umgang mit variablen Beobachtungszeiten“ um Ausführungen zu Analysen zur „dauerhaften“ Veränderungen mit der Schlussfolgerung „solche Auswertungen werden im Falle einer verkürzten oder sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauer daher in der Regel nicht für die Nutzenbewertung herangezogen“. Der G-BA hat ebenfalls in seinem FAQ „Antworten auf häufig gestellte Fragen zum Verfahren der Nutzenbewertung“ diesen Aspekt thematisiert. Bei relevant unterschiedlichen Beobachtungszeiten zwischen den Studienarmen, ist die Auswertung der Zeit bis zur bestätigten oder dauerhaften Veränderung nicht sinnvoll interpretierbar und es muss die Zeit bis zur erstmaligen Veränderung vorgelegt werden. Liegen keine unterschiedlichen Beobachtungszeiten zwischen den Armen vor, ist die Auswertung der Zeit bis zur bestätigten oder dauerhaften Veränderung grundsätzlich inhaltlich sinnvoll. Es ist zu begründen, wieso die gewählte Auswertung zur erstmaligen, bestätigten oder dauerhaften Veränderung geeignet ist¹. Das IQWiG erläutert weitergehend, dass Datenanalysen mit unterschiedlich langer Beobachtungsdauer, die dies nicht adäquat berücksichtigen, nicht als verwertbar eingestuft werden. Im Vergleich zum Methodenpapier 6.1, wo ausgeführt wird, dass unterschiedlich lange Beobachtungsdauern ohne adäquate Berücksichtigung in der Datenanalyse dazu führen können, dass die entsprechenden Ergebnisse für die Nutzenbewertung nicht verwertbar sind, wird im Methodenpapier 7.0 wie folgt formuliert: „Insgesamt <i>sind</i> bei unterschiedlich langer Beobachtungsdauer zwischen den zu vergleichenden Gruppen Datenanalysen, die dies nicht adäquat berücksichtigen, <i>nicht für die Nutzenbewertung verwertbar</i>.“ Die Einstufung als nicht verwertbar sollte keine Regel sein, sondern einer Einzelfallprüfung unterliegen, in der das Ausmaß der Verzerrung und Ergebnissicherheit geprüft wird. Der jeweilige Kontext (Population, Intervention, Comparator, Endpunkt, und intercurrent event Strategie) und der gewählte Schätzer spielen eine entscheidende Rolle, ob und wie sich unterschiedliche Beobachtungszeiten auswirken. Das Ausmaß der möglichen Verzerrung und die Ergebnissicherheit der Analysen sollte entscheidend für deren Verwertbarkeit im Rahmen der Nutzenbewertung sein. Rufibach et al. beschreiben beispielsweise in ihrem Artikel „Comparison of Adverse Event Risks in

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Randomized Controlled Trials with Varying Follow-Up Times and Competing Events: Results from an Empirical Study", dass die Auswahl des Schätzers wichtiger ist als die Datenmerkmale wie z.B. „amount of follow-up“². Zitat: „We find that for both, avoiding bias and categorization of evidence with respect to treatment effect on AE risk into categories, the choice of the estimator is key and more important than features of the underlying data such as percentage of censoring, CE (<i>competing events</i>), amount of follow-up, or value of the gold-standard RR (<i>risk ratio</i>).“ Aus unserer Sicht trifft dies auch für patienten-berichtete symptomatische Ereignisse zu. Die Einstufung „nicht verwertbar“ kann deshalb nicht die Regel sein. Das IQWiG führt abschließend aus, dass „im Falle systematisch verkürzter Beobachtungsdauern die Ergebnisse davon betroffener Endpunkte (z.B. unerwünschter Ereignisse oder patient reported outcomes (PROs)) die Gesamtaussage zum Nutzen oder Zusatznutzen einer medizinischen Intervention in der Regel nicht dominieren“ können. Hierbei ist zu berücksichtigen, dass sich die Datenerhebung an der Praktikabilität und Interpretierbarkeit der PRO Daten orientieren sollte, wie von der EMA in der 'guideline on the evaluation of anticancer medicinal products' beschrieben: "PRO measures should be administered to study subjects at time points when there is a clear and hypothesis driven rationale for their use and when it is feasible to expect high levels of completion by the individual patient"³. Längere Beobachtungsdauern können aufgrund sinkender Rücklaufquoten mit einem hohen Verzerrungspotential einhergehen, welches die Aussagesicherheit einschränken würde. Die Interpretation der Ergebnisse ist, wie z.B. in der Onkologie, nach Progression und folgendem Behandlungsabbruch oder Therapiewechsel erschwert. Entsprechend sollte vom Begriff „systematisch verkürzte Beobachtungsdauer“ abgesehen werden. <u>Vorgeschlagene Änderung:</u> Entsprechend des G-BA FAQ sollte ergänzt werden, dass eine Analyse der Zeit bis zur dauerhaften Veränderung sinnvoll sein kann, solange die Beobachtungszeiten zwischen den Armen nicht relevant unterschiedlich sind. Zudem kann ergänzt werden, dass bei relevant unterschiedlichen Beobachtungszeiten zwischen den Studienarmen eine Analyse der Zeit bis zur erstmaligen Veränderung eine adäquate Operationalisierung sein kann. Datenanalysen mit unterschiedlichen Beobachtungszeiten, die diese nicht adäquat berücksichtigen, sollten nicht automatisch als nicht verwertbar eingestuft werden. Die ursprüngliche Formulierung sollte belassen werden.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Von einer geringeren Gewichtung auf Basis der Beobachtungsdauer bei der Bildung der Gesamtaussage zum Zusatznutzen ist abzusehen. Konkreter Vorschlag zur Ergänzung des Abschnitts (Ergänzung in fett hervorgehoben): S. 190: „Um solche Probleme zu vermeiden, sollten die Daten zu allen bewertungsrelevanten Endpunkten, also auch zu PROs und unerwünschten Ereignissen, vollständig erfasst werden, und zwar auch nach einem Behandlungsabbruch oder -wechsel [53], sofern eine sinnvolle Interpretation der Ergebnisse möglich ist. S. 191: Solche Auswertungen werden im Falle einer verkürzten oder sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauer daher in der Regel nicht für die Nutzenbewertung herangezogen. Eine Analyse der Zeit bis zur erstmaligen Veränderung kann jedoch eine adäquate Operationalisierung sein. Eine Analyse der Zeit bis zur dauerhaften Veränderung kann auch sinnvoll sein, solange die Beobachtungszeiten zwischen den Armen nicht relevant unterschiedlich sind. Konkreter Vorschlag zur Ersetzung folgenden Satzes aus Version 7.0 mit dem Satz aus Version 6.1: Version 7.0 S. 191: „Insgesamt sind bei unterschiedlich langer Beobachtungsdauer zwischen den zu vergleichenden Gruppen Datenanalysen, die dies nicht adäquat berücksichtigen, nicht für die Nutzenbewertung verwertbar.“ Ersetzen mit Version 6.1 S. 199: „Liegen bei den zu vergleichenden Gruppen unterschiedlich lange Beobachtungsdauern vor, ohne dass dies in der Datenanalyse adäquat berücksichtigt worden ist, so kann dies dazu führen, dass die entsprechenden Ergebnisse für die Nutzenbewertung nicht verwertbar sind.“

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. G-BA. Antworten auf häufig gestellte Fragen zum Verfahren der Nutzenbewertung - Gemeinsamer Bundesausschuss. <https://www.g->

ba.de/themen/arzneimittel/arzneimittel-richtlinie-anlagen/nutzenbewertung-35a/faqs/.

2. Rufibach, K. *et al.* Comparison of Adverse Event Risks in Randomized Controlled Trials with Varying Follow-Up Times and Competing Events: Results from an Empirical Study. *Stat. Biopharm. Res.* 1–14 (2022) doi:10.1080/19466315.2022.2144944.
3. European Medicines Agency. Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. www.ema.europa.eu/contact (2016).

A.1.30 MSD Sharp & Dohme GmbH

Autorinnen und Autoren

- Tränkl, Sandra
- Pointner, Julia

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Tränkl, Sandra; Dr.
Pointner, Julia; Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: MSD Sharp & Dohme GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 9.1.1 Kriterien für den Einschluss von Studien, S. 159	<u>Anmerkung:</u> Das IQWiG fasst die Kriterien für den Einschluss von Studien neu, indem es das Kriterium der Prüfung einer ggf. vorliegenden Effektmodifikation hinzufügt. „Falls geeignete Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium (Population bzw. Interventionen) vorliegen, muss über die Berücksichtigung der Studie anhand der Stärke der Effektmodifikation und dem Anteil der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen bzw. dem Grad der Abweichung der Interventionen entschieden werden.“ Der oben genannte Textabschnitt in Kapitel 9.1.1. lässt viel Spielraum für Interpretation. Folgende Punkte erscheinen aus Sicht von MSD unklar: • Es bleibt offen, wann es sich um „geeignete Informationen“ bezüglich einer Effektmodifikation handelt und wann eine solche Information als ungeeignet bewertet wird. • Zudem bleibt offen, wann eine Effektmodifikation als hinreichend stark erachtet wird, um zum Ausschluss einer Studie zu führen. • Des Weiteren ist unklar, ob eine hinreichend starke Effektmodifikation bei einem einzelnen oder wenigen Endpunkten ausreicht, um auch bei allen anderen Endpunkten die Übertragbarkeit in Frage zu stellen.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	• Des Weiteren erscheint das Vorgehen mit einer allgemein gültigen Grenze von weniger als 80% die Analysen der Gesamtpopulation einer Studie abzulehnen, wie es auch in früheren Versionen des Methodenpapiers angeführt wird, nicht immer als zielführend. So sollten auch in Situationen, in denen die relevante Teilpopulation weniger als 80% der Studienpopulation beträgt, medizinisch und biologisch plausible Faktoren mit einbezogen werden und situativ über eine Eignung der Studie entschieden werden. • Es sollte ferner konkretisiert werden, ob das Institut weiterhin die Anhebungsregel empfiehlt bzw. an dieser festhält, wenn Ergebnisse für die relevante Teilpopulation vorliegen, jedoch aufgrund des Powerverlusts kein Zusatznutzen vorliegt. <u>Vorgeschlagene Änderung:</u> Aufgrund der genannten Aspekte und Unsicherheiten ist der oben dargestellte Textabschnitt in dieser Form nicht aussagekräftig. Aus Sicht von MSD sollte der Abschnitt deshalb angepasst und konkretisiert werden.
Abschnitt 9.3.2 Statistische Signifikanztests und Konfidenzintervalle, S. 170	<u>Anmerkung:</u> Das IQWiG ergänzt, dass Methoden zur Adjustierung für multiples Testen z.B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist, angewendet werden können. Aus Sicht von MSD ist ein konsistentes Vorgehen bei der Kontrolle des Fehlers 1. Art im Rahmen der Nutzenbewertung nötig. So ist z.B. bei den patientenberichteten Endpunkten die Auswertung nach mehreren Analysearten - Responderanalysen und Analysen kontinuierlicher Daten - gewünscht. Liefert der pU diese, so besteht die Gefahr, dass aufgrund einer p-Wert Korrektur durch das IQWiG, signifikante Effekte nicht berücksichtigt werden. Ein derartiges Vorgehen kann für Auswertungen mit hohen Analysevolumina, wie z.B. Nutzenbewertungen nicht zielführend sein. <u>Vorgeschlagene Änderung:</u> Aus Sicht von MSD sollte konkretisiert werden, unter welchen Umständen eine Adjustierung für multiples Testen im Rahmen der Nutzenbewertung sinnvoll sein kann.
Abschnitt 9.3.12 Umgang mit variablen	<u>Anmerkung:</u> Das IQWiG ergänzt die Ausführungen zum Umgang mit variablen Beobachtungszeiten. Es führt aus, dass bei sich deutlich zwischen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Beobachtungszeiten, S. 190 f.	den Behandlungsarmen unterscheidenden Beobachtungsdauern eine Analyse der PRO zur dauerhaften Veränderung nicht sinnvoll interpretierbar ist und in der Regel nicht für die Nutzenbewertung herangezogen wird. Zudem fügt das IQWiG hinzu, dass im Falle systematisch verkürzter Beobachtungsdauer die Ergebnisse davon betroffener Endpunkte die Gesamtaussage zum Nutzen oder Zusatznutzen einer medizinischen Intervention in der Regel nicht dominieren können. Aus Sicht von MSD ist die adäquateste Auswertungsmethode patientenberichteter Endpunkte stets von der jeweils vorliegenden Indikations- und Studiensituation abhängig und daher spezifisch für jede Studie zu wählen. Ebenso ist eine Herabstufung der Relevanz patientenberichteter Endpunkte und Endpunkte der Kategorie Nebenwirkungen aufgrund der Beobachtungsdauer aus Sicht von MSD nicht sachgerecht. Eine Erhebung patientenberichteter Endpunkte sowie der Endpunkte zu Nebenwirkungen bis zum Behandlungsende (zzgl. Nachbeobachtungszeit) dient dazu sicherzustellen, dass die jeweiligen Ereignisse eindeutig auf die unterschiedlichen Behandlungen zurückzuführen sind. Nach Behandlungsabbruch kommt es häufig zu Folgetherapien, die die Ergebnisse verzerren können. Sowohl patientenberichtete Endpunkte als auch Endpunkte zu Nebenwirkungen sind ohne a priori Herabstufung der Relevanz in das Gesamtbild der Nutzenbewertung einzubeziehen. <u>Vorgeschlagene Änderung:</u> Eine Beobachtungsdauer bis zum Behandlungsabbruch oder -wechsel sollte nicht zu einer a priori Herabstufung der Relevanz der betroffenen Endpunkte führen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

A.1.31 Novartis Pharma GmbH

Autorinnen und Autoren

- Denzer-Lippmann, Melanie
- Florschütz, Kristina
- Kolb, Benjamin
- Lambertz, Lisa
- Marx, Almuth
- Nguyen, Hien
- Templin, Corinna

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Denzer-Lippmann, Melanie; Dr.
Florschütz, Kristina
Kolb, Benjamin; Dr.
Lambertz, Lisa
Marx, Almuth; Dr.
Nguyen, Hien
Templin, Corinna; Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Novartis Pharma GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 3.3.4 (S.66-68)	<u>Anmerkung:</u> Abschnitt 3.3.4 ist ein neues Kapitel im vorliegenden Entwurf zum Methodenpapier 7.0 zur Erstellung von Konzepten für anwendungsbegleitende Datenerhebungen (AbD). Das Kapitel basiert auf dem Rapid Report A19-43 des IQWiG (1). Anhand des AbD-Konzepts sollen die methodischen Anforderungen an die Durchführung der AbD beschrieben werden. <u>Vorgeschlagene Änderung:</u> Die wissenschaftliche Fragestellung, die im AbD-Konzept im PICO-Format beschrieben wird, ergibt sich laut IQWiG-Methodenpapier aus dem Informationsbedarf für die Nutzenbewertung. Da der medizinische Zusatznutzen eines Arzneimittels über alle Nutzen-dimensionen bewertet werden sollte, sollte aus Sicht von Novartis die Definition der Outcomes im AbD-Konzept nicht durch die Verfügbarkeit dieser Daten in einem möglicherweise geeigneten Register beeinflusst werden. Die Erfahrung im AbD-Konzept mit Fedratinib (2) zeigt, dass sich, entgegen der Anforderungen der Nutzenbewertung, die Wahl der im Konzept geforderten Outcomes nicht an der Fragestellung orientiert, sondern an der Verfügbarkeit der Daten im Register. Es sollte daher bereits bei der Konzept-erstellung geprüft werden, ob eine ganzheitliche Bewertung des Zusatznutzens unter Berücksichtigung der vorab definierten patientenrelevanten Endpunkte in der jeweiligen Indikation mit Hilfe der AbD möglich ist. Eine Fokussierung auf einzelne verfügbare Endpunkte kann nicht die Basis einer Nutzenbewertung mit Hilfe einer AbD sein, da ein ggf. vorhandener Zusatznutzen/ Schaden in

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	weiteren Nutzendimensionen hierdurch nicht in der Bewertung berücksichtigt wird. Im Rahmen der Konzepterstellung soll laut Methodenpapier geprüft werden, ob geeignete Datenquellen zur Beantwortung der AbD-Fragestellungen existieren. So sollen Register identifiziert werden, die die notwendigen Daten in ausreichender Qualität zur Verfügung stellen können. Aus dieser Eignungsprüfung wird zudem der Änderungs- und Erweiterungsbedarf abgeleitet, der ggf. erforderlich ist, damit das jeweilige Register als primäre Datenquelle für die Registerstudie zum Zwecke der AbD dienen kann. Sollte sich abzeichnen, dass der Erweiterungsbedarf eines potenziell geeigneten Registers mit einem nicht verhältnismäßigen Aufwand an Zeit und Ressourcen verbunden ist, sollte aus Sicht von Novartis von einer Verwendung des Registers als Datenquelle für eine AbD abgesehen werden. Bei fraglicher Eignung eines Registers als Datenquelle für eine AbD sollte vor Beginn der Protokollerstellung durch den betroffenen pharmazeutischen Unternehmer (pU) abschließend geklärt sein, ob erforderliche Anpassungen in der Registerstruktur möglich und umsetzbar sind. Novartis merkt an, dass bei der Planung der Studiendauer und des Studienumfangs im Rahmen einer Konzepterstellung für eine AbD nicht die in Modul 3 des Herstellerdossiers berechneten Zahlen zur Größe der Zielpopulation als Grundlage herangezogen werden sollten, da diese als theoretisch maximale Größe nicht der Anzahl der Patientinnen und Patienten entspricht, die im festgelegten Zeitraum realistisch für ein Register rekrutiert werden können. Es müssen Versorgungsaspekte, Patientenpräferenzen sowie regionale Unterschiede in der Versorgung berücksichtigt werden, um eine realistische Studienplanung zu ermöglichen.
Abschnitt 4 (S.83-111)	<u>Anmerkung:</u> Abschnitt 4 wurde im Vergleich zur vorhergehenden Version 6.1 des Methodenpapiers grundlegend überarbeitet. Ein neuer Aspekt ist das Heranziehen der zVT bzw. der vom pU als Therapieoption gewählten zVT in der Kosten-Nutzen-Bewertung im Gegensatz zu allen versorgungsrelevanten Arzneimitteln, wie es bislang der Fall war. <u>Vorgeschlagene Änderung:</u> Abschnitt 4 gibt keine klare Trennung zwischen Nutzenbewertung und Kosten-Nutzen-Bewertung (KNB) hinsichtlich ihrer

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Interdependenzen. In Abschnitt 4.1 wird festgehalten, es sei „Ziel der KNB [...], ökonomische Informationen als Zusatz zur Nutzenbewertung insbesondere für die Preisverhandlungen im Sinne einer Informationssynthese zusammenzufassen“ (S. 83). Es sollte darüber hinaus noch explizit klargestellt werden, dass ausschließlich die Nutzenbewertung nach §35a SGB V über den medizinischen Mehrwert entscheidet und diese nach Abschluss nicht erneut durch eine KNB angestoßen werden kann. Diese Option wird jedoch in Abschnitt 4.3.2 suggeriert („Im Fall, dass in der KNB ein Vergleich mit mehreren Vergleichstherapien und nicht nur mit der gewählten zweckmäßigen Vergleichstherapie stattfinden soll, könnte diese Betrachtung nicht auf Grundlage der bereits vorliegenden Nutzenbewertung erfolgen, sondern ggf. ist eine erneute Nutzenbewertung unter Einbeziehung der weiteren Komparatoren erforderlich.“ S. 88). Beide Verfahren haben jedoch unterschiedliche Perspektiven bzw. Intentionen und sollten deshalb klar voneinander abgegrenzt werden.
Abschnitt 4.3.3 (S.88-89)	<u>Anmerkung:</u> Im Zusammenhang mit der Gewichtung der Ergebnisse im Rahmen der Beschlussfassung einer KNB nennt das IQWiG Präferenz-erhebungen von Patientinnen und Patienten als eine Option: „Eine andere Möglichkeit ist eine Gewichtung der Ergebnisse unter Berücksichtigung von Verfahren der Präferenz-erhebung von Patientinnen und Patienten.“ <u>Vorgeschlagene Änderung:</u> An dieser Stelle mangelt es an einem Bewertungsrahmen bzw. Erwartungshorizont für eine vom pU durchgeführte Präferenz-erhebung. Aus unserer Sicht würde eine Konkretisierung der Methodik für eine Präferenz-erhebung die Akzeptanz der Ergebnisse im Rahmen einer KNB seitens des IQWiG deutlich erhöhen.
Abschnitt 4.7.2 (S.95-98)	<u>Anmerkung:</u> In diesem Abschnitt wird der 4-stufige Prozess zur Bestimmung der in das Modell eingehenden Kosten erläutert. Zu Stufe 3 „Bewertung der Ressourcen“ heißt es: „Wird der Zeitaufwand von Betroffenen oder Angehörigen in der Kostenbestimmung berücksichtigt, wird dieser mit dem Nettolohn bewertet.“ <u>Vorgeschlagene Änderung:</u> Aus unserer Sicht sollte hier eine exakte Definition von „Nettolohn“ sowie eine Begründung, weswegen in die Bewertung der Netto- und nicht der Bruttolohn einfließt, ergänzt werden.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 4.8 (S.98)	<u>Anmerkung:</u> Dieser Abschnitt behandelt knapp das Vorgehen zu Inflationsbereinigung und Diskontierung bei einer KNB. <u>Vorgeschlagene Änderung:</u> Auch wenn die gängige Praxis bei KNB gleiche Diskontierungsraten für Kosten als auch Nutzen ansetzt (3), so ist die Diskontierung des gesundheitlichen Nutzens umstritten bzw. diskutiert (4). Daher sollte die vom IQWiG vorgeschlagene differenzielle Diskontierung mindestens als regelhafte Sensitivitätsanalyse herangezogen werden.
Abschnitt 4.9 (S.98-103)	<u>Anmerkung:</u> In diesem Abschnitt werden die Grundlagen für das Konzept und die Struktur eines gesundheitsökonomischen Modells für die KNB beschrieben. Für die Extrapolation von Ergebnissen aus klinischen Studien nennt das IQWiG zwei Szenarien, die dabei in Betrachtung kommen. Vor dem Hintergrund, dass das IQWiG in Szenarioanalysen der KNB regelhaft einen Zeithorizont von 5 Jahren betrachten möchte (siehe Abschnitt 4.3.5, S. 90), erscheinen beide vorgeschlagenen Szenarien zur Extrapolation der Studienergebnisse zu konservativ und somit zu einseitig. So soll im „konservativen Szenario“ der beobachtete Effekt bzw. Unterschied zwischen den betrachteten Therapien abrupt nach Ende der klinischen Studie abreißen. Das ist weder bei einmalig zu verabreichenden Therapien (z. B. Gentherapien) noch bei dauerhaft anzuwendenden Arzneimitteln zu erwarten. Im „optimistischeren Szenario“ wird zwar angenommen, dass der Effekt nach Studienende noch anhält. Nichtsdestotrotz soll auch in diesem Szenario der Effekt anschließend abfallen oder enden. Dahingegen wird kein Szenario vorgeschlagen, welches annimmt, dass der beobachtete Effekt bzw. Unterschied zumindest anhält oder gar ansteigt, solange die Intervention gegeben wird. Es gibt kaum klinische Studien, die eine direkt vergleichende Beobachtungsdauer von fünf Jahren aufweisen, sodass im Regelfall nach Ansicht des IQWiG der extrapolierte Effekt zum Ende des betrachteten Zeithorizonts der KNB hin immer niedriger sein wird als in den Studien gezeigt – sowohl im „konservativen“ als auch im sogenannten „optimistischeren“ Szenario. Damit wird bei den

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	vorgeschlagenen Extrapolationsszenarien des IQWiG für die KNB der Nutzen des Arzneimittels regelhaft reduziert. <u>Vorgeschlagene Änderung:</u> Stattdessen wird vorgeschlagen, dass der in den klinischen Studien beobachtete Kurvenverlauf für die Extrapolation der Studienergebnisse herangezogen wird. Bei der Extrapolation sollte zudem nicht nur auf direkt vergleichende Studien zurückgegriffen werden. So können auch einarmige Extensionsstudien valide Daten zum Verlauf des beobachteten Effekts über einen längeren Zeitraum liefern.
Abschnitt 4.14 (S.108-110)	<u>Anmerkung:</u> In Abschnitt 4.14 werden die Informationsbeschaffung im Rahmen der KNB und die zu Verfügung stehenden Datenquellen behandelt. Im Zusammenhang mit „fokussierter Informationsbeschaffung“ schreibt das IQWiG: „Die Bestimmung der Nutzwerte kann im Rahmen klinischer Studien erfolgen. Alternativ oder ergänzend kann eine fokussierte Informationsbeschaffung nach Nutzwerten für die im Modell betrachteten Gesundheitszustände erforderlich sein. Falls vorhanden, werden validierte Studienfilter verwendet. Dies trifft z. B. auf die Recherche nach gesundheitsökonomischen Evaluationen [293] oder nach Nutzwerten [27] zu.“ <u>Vorgeschlagene Änderung:</u> In diesem Abschnitt sollte klar dargestellt werden, wie zu verfahren ist, wenn keine Bewertungen von Gesundheitszuständen auf Basis der in Abschnitt 4.6 erwähnten indirekten/direkten Instrumente verfügbar sind. Tabelle 8 (S. 110) spricht von „fokussierter Informationsbeschaffung“ zu Nutzwerten als „optional.“ Bedeutet das, dass eine Bewertung der Gesundheitszustände und somit die Durchführung einer Kosten-Nutzwert-Analyse als optional gelten?
Abschnitt 9.3.2 (S.170)	<u>Anmerkung:</u> Im vorliegenden Entwurf zum Methodenpapier 7.0 hat das IQWiG den Abschnitt 9.3.2 neu strukturiert und überarbeitet. Im Gegensatz zur Version 6.1 des Methodenpapiers wird nun im Zusammenhang mit der frühen Nutzenbewertung auch auf Adjustierung für multiples Testen eingegangen: „Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an [59], z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.“

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Es ist in der Nutzenbewertung gängige Praxis, dass Endpunkte auf verschiedene Arten analysiert werden. So werden stetige Endpunkte häufig sowohl als Veränderung zur Baseline als auch im Rahmen von Responderanalysen ausgewertet. Solche unterschiedlichen Operationalisierungen sind in der Regel hoch korreliert, was zu einer nur geringen Inflation des α-Fehlers führt. Die gängigen Adjustierungsverfahren gehen hingegen in der Regel von unkorrelierten Endpunkten aus. <u>Vorgeschlagene Änderung:</u> Es sollte eine Konkretisierung vorgenommen werden, in welchen Fällen/unter welchen Umständen das IQWiG Methoden zur Adjustierung für multiples Testen heranziehen würde, und welche Verfahren hierbei zum Einsatz kommen sollen.
Abschnitt 9.3.12 (S.190-191)	<u>Anmerkung:</u> Im vorliegenden Entwurf zum Methodenpapier 7.0 ergänzt das IQWiG in Abschnitt 9.3.12 zusätzliche Details zum Umgang mit PROs (patient reported outcomes) sowie mit unerwünschten Ereignissen bei unterschiedlichen Beobachtungszeiten in den Behandlungsarmen bzw. bei systematisch verkürzten Beobachtungsdauern. Im Fazit seien Datenanalysen, welche die unterschiedlich lange Beobachtungsdauer zwischen den Behandlungsgruppen „nicht adäquat berücksichtigen, nicht für die Nutzenbewertung verwertbar“. Im Falle der systematisch verkürzten Beobachtungsdauer könnten die Ergebnisse davon betroffener Endpunkte „die Gesamtaussage zum Nutzen oder Zusatznutzen einer medizinischen Intervention in der Regel nicht dominieren“. Weiterhin ergänzt das IQWiG, dass Auswertungen der „dauerhaften Veränderung“ im Falle einer verkürzten oder sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauer in der Regel nicht für die Nutzenbewertung herangezogen werden sollen. <u>Vorgeschlagene Änderung:</u> Patientenberichtete Endpunkte (PROs) sind ein wesentlicher Teil der Nutzenbewertung, da sie die Sicht der Patienten unmittelbar widerspiegeln. Indikationsbedingt kann es zu unterschiedlichen Beobachtungsdauern kommen. Dennoch sollten diese Daten unter Berücksichtigung der möglichen Verzerrung in der Nutzenbewertung vollumfänglich herangezogen werden. Darüber hinaus sollte die bislang anerkannte Operationalisierung „Zeit bis zur ersten Verschlechterung“ bei patientenberichteten Endpunkten weiterhin vollumfänglich akzeptiert werden. Eine Einschränkung hinsichtlich der PROs und unerwünschten

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Ereignisse zur Gesamtaussage zum Nutzen oder Zusatznutzen einer medizinischen Intervention ist nicht sachgerecht. Im Detail sollte aus unserer Sicht Folgendes ergänzt werden: • Konkretisierung – was ist ein „deutlich sich unterscheiden-der Beobachtungszeitraum zwischen den Behandlungs-armen“? Dies sollte unter Berücksichtigung des jeweiligen Therapiegebiets und der konkreten Indikation betrachtet werden. • Konkretisierung – was ist mit “systematisch verkürzter Beobachtungsdauer” gemeint? - o V.a. bei onkologischen Studien ist der Gesundheits-zustand der Patienten zu berücksichtigen. Eine Erfassung der PROs bis zum Tod kann in der Regel nicht umgesetzt werden (Gründe dafür sind z. B.: Patient lebt nicht für Studie; Patienten geht es ab einem gewissen Zeitpunkt so schlecht, dass PROs-Erfassung physisch nicht umsetzbar ist; etc.). - o Eine Erfassung bis Studienende ist nicht in jeder Therapiesituation adäquat umsetzbar. Es sollte festgehalten werden, dass eine angemessene Anzahl an PRO-Erfassungen nach Behandlungsende als adäquat anerkannt wird, auch für die „dauerhafte Veränderung“. Somit sollte sich keine Einschränkung hinsichtlich der PROs zur Gesamtaussage zum Nutzen oder Zusatznutzen einer medizinischen Intervention ergeben.
Abschnitt 9.3.8 (S.184)	<u>Anmerkung:</u> Im Falle einer nicht erfüllten Konsistenzannahme (ausreichende Ähnlichkeit und Homogenität der paarweisen Metaanalysen) auf Grund nichtexistierender Studiendaten, kann kein Beleg für einen Nutzen bzw. Schaden der zu bewertenden Intervention abgeleitet werden. <u>Vorgeschlagene Änderung:</u> In Fällen, in denen die aktuelle Evidenzlage ausschlaggebend für das Fehlen relevanter Informationen ist, muss dieses auch entsprechend gekennzeichnet werden. Ansonsten kann der Eindruck entstehen, dass es keinen Unterschied in Bezug auf Nutzen oder Schaden zwischen zwei Behandlungsalternativen gibt. Beispiel: „Es gibt bislang keine geeigneten Studiendaten, um einen Beleg für einen höheren/ niedrigeren Nutzen bzw. Schaden abzuleiten.“

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach §35a SGB V. Rapid Report A19-43. 2020.
2. Gemeinsamer Bundesausschuss. Zusammenfassende Dokumentation über einer Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Fedratinib (Myelofibrose) Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen. 2022.
3. Attema AE, Brouwer WBF, Claxton K. Discounting in Economic Evaluations. Pharmacoeconomics. 2018;36(7):745-58.
4. Severens JL, Milne RJ. Discounting health outcomes in economic evaluation: the ongoing debate. Value Health. 2004;7(4):397-401.

A.1.32 Novo Nordisk Pharma GmbH

Autorinnen und Autoren

- Kiencke, Peter
- Knerr-Rupp, Katrin

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Kiencke, Peter; Dr.
Knerr-Rupp, Katrin; Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Novo Nordisk Pharma GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.
Mit dem vorgelegten Entwurf für Version 7.0 insbesondere Kapitel 4 folgt das IQWiG seinem gesetzlichen Auftrag gemäß §35b (1) SGB V, transparente Methoden zur Kosten-Nutzen-Bewertung (KNB) von Arzneimitteln festzulegen. Der Gesetzgeber hat die KNB im Rahmen des AMNOG 2010 für den Fall vorgesehen, dass im regulären Prozess von Nutzenbewertung nach § 35a und Preisfindung nach § 130b keine angemessene Lösung für ein neues Arzneimittel im Sinne der Wirtschaftlichkeit gefunden werden kann. Demnach kann nach § 130b (8) SGB V eine der beiden Vertragsparteien (Pharmazeutischer Unternehmer oder GKV-SV) nach Abschluss eines Schiedsverfahrens eine KNB beim G-BA beantragen, der dann das IQWiG mit der Durchführung derselbigen beauftragt. Das Ergebnis der KNB soll dann als neue Grundlage zur Vereinbarung eines angemessenen Erstattungsbetrags dienen. Die KNB stellt somit ein letztes Mittel bei der Erstattungspreisfindung dar, das im Rahmen des AMNOG- Verfahrens seit 2010 jedoch noch nie zum Einsatz kam. Die KNB soll auch in Zukunft nicht zu einer vierten Hürde werden. Die Änderung des Abschnitts 4 des Methodenpapiers ist damit eher theoretischer als praktischer Natur. Im Rahmen des gesetzlich vorgesehenen Zwecks der KNB orientiert das IQWiG die Methodik an den internationalen wissenschaftlichen Standards der Gesundheitsökonomie. Jedoch wurden im Methodenentwurf 7.0 des IQWiG die methodischen Standards an einigen Stellen unklar beschrieben. Eine Kommentierung ist daher geboten.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung
	<i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 1.4 Gesundheitsökonomie, S. 10 und 4.1 Einführung, S. 83	<u>Anmerkung:</u> Das IQWiG formuliert die Zielsetzung der KNB wie folgt: „Ziel der KNB ist es, ökonomische Informationen als Zusatz zur Nutzenbewertung insbesondere für die Preisverhandlungen im Sinne einer Informationssynthese zusammenzufassen.“ Die KNB ist jedoch kein rechtlich vorgesehener Teil der Erstattungsbetragsverhandlungen und der entsprechenden Rahmenvereinbarung. Diese kann basierend auf § 35b SGB

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	V als Orientierung bei der Preisfindung dienen, wenn die entsprechenden Voraussetzungen erfüllt sind (https://www.g-ba.de/themen/arzneimittel/arzneimittel-richtlinie-anlagen/kosten-nutzen/). <u>Vorgeschlagene Änderung:</u> Der Satz zum Ziel der KNB sollte wie folgt angepasst werden: „Ziel der KNB ist es, nach abgeschlossener früher Nutzenbewertung bei Arzneimitteln mit neuen Wirkstoffen als ökonomische Orientierung (entsprechend § 35b SGB V) als Zusatz zur Nutzenbewertung zu dienen.“
Abschnitt 4.3.3 Gesundheitsökonomische Zielgröße und 4.6 Nutzwerte (im Fall einer Kosten-Nutzwert-Analyse), S. 85, 88-89 und S. 91	<u>Anmerkung:</u> Das IQWiG sieht die Möglichkeit einer Kosten-Nutzwert-Analysen als eine Form der Kosten-Effektivitäts-Analyse vor. Die Kosten sollen in Relation zu qualitätsadjustierten Lebensjahren (QALYs) gesetzt werden, wobei die Bewertung von Gesundheitszuständen in Form von Nutzwerten mit der Restlebenserwartung oder Verweildauer pro Gesundheitszustand kombiniert werden. Die Nutzwerte können dabei auch durch indirekte Verfahren z. B. basierend auf den Fragebögen EQ-5D, SF-6D sowie HUI 2 oder 3 ermittelt werden. Auswertungen basierend auf dem EQ-5D Index in den wurden zu Beginn des AMNOG in den Bewertungen des Zusatznutzens durch den G-BA unter der Endpunktkategorie „gesundheitsbezogene Lebensqualität“ bewertet. Dossierbewertung A11-23 Fingolimod – Nutzenbewertung gemäß § 35a SGB V 11.01.2012 Seite 35: https://www.g-ga.de/bewertungsverfahren/nutzenbewertung/15/#nutzenbewertung „Der pU zieht als Messinstrument für die gesundheitsbezogene Lebensqualität den Fragebogen EQ-5D heran und verweist auf eine Validierungspublikation zu Patienten mit Multipler Sklerose [12]. In der genannten Validierungspublikation wird der EQ-5D jedoch nicht als Instrument für die gesundheitsbezogene Lebensqualität validiert, sondern als Health-Utility Instrument. Dennoch wird der EQ-5D für die vorliegende Nutzenbewertung akzeptiert, da er zumindest Teilbereiche der gesundheitsbezogenen Lebensqualität abdeckt“ Dossierbewertung für Orphan Drugs Riociguat 1. August 2014 Seite 37-38: https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/110/#nutzenbewertung Einschätzung zur Qualität u. Patientenrelevanz Definition und Erhebung des Endpunkts: Der EQ-5D ist ein häufig genutzter, valider, generischer Gesundheitsfragebogen, der die selbstberichtete Lebensqualität eines Patienten in einer eindimensionalen Maßzahl von 5 (sehr gut) bis 15 (extrem niedrig) ausdrückt. Der Gesundheitszustand wird in fünf Dimensionen beschrieben: Beweglichkeit und Mobilität, die Fähigkeit, für sich selbst zu sorgen, alltägliche Tätigkeiten (z. B. Arbeit, Studium, Hausarbeit, Familie, Freizeit), Schmerzen und körperliche Beschwerden, Angst und Niedergeschlagenheit. Der EQ-5D-Fragebogen gehört zu den Single-Index-Instrumenten, wobei nicht für jede Lebensqualitäts-Dimension ein Einzelwert bestimmt wird, sondern die 5 Fragen mittels standardisierter Berechnungsvorgaben in einen Indexwert umgerechnet werden. In den Studien wurde der EQ-5D 3L Lebensqualität genutzt, bei dem die Antworten auf einer 3-Punkte-Likert-Skala erfasst werden, im Gegensatz zum EQ-5D mit 5-Punkte-Likert-Skala. Die MID für den EQ-5D wird allgemein mit 0,074 angegeben (Walter et al., 2005).

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Später wurden EQ-5D Index Werte stets aus den Nutzenbewertungen ausgeschlossen, da die Nutzwerte „nicht als eigentlicher Nutzenparameter interpretiert werden“ konnten. Dossierbewertung A15-32 Nivolumab (neues Anwendungsgebiet) – 12.11.2015 (Seite 49-50) https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/186/#nutzenbewertung „Gesundheitsbezogene Lebensqualität nach EQ-5D Index Wert: nicht eingeschlossen Neben der unter Morbidität aufgeführten VAS enthält der EQ-5D auch 5 Fragen zu unterschiedlichen Domänen mit jeweils 3 Antwortkategorien. Diese ordnete der pU der Endpunktkategorie gesundheitsbezogene Lebensqualität zu. Für den Indexwert und die Werte der einzelnen Domänen legte der pU in Modul 4 B MMRM Auswertungen zur Änderung der Mittelwerte zu Studienbeginn vor, sowie eine Sensitivitätsanalyse zurzeit bis zur Verschlechterung für den Indexwert. Es lagen keine Angaben dazu vor, wie der Indexwert ermittelt wurde und ob er anhand von Patienten im relevanten Anwendungsgebiet ermittelt wurde. Die Auswertungen auf Basis der der einzelnen Dimensionen können nicht interpretiert werden, da den Messwerten keine arithmetischen Eigenschaften zugeordnet werden können [16].“ Insgesamt erscheint damit der Umgang mit den Ergebnissen solcher Instrumente wie EQ-5D Index als inkonsistent, da sie als Option der Messung des Nutzens bei KNB darstellen sollen, zugleich jedoch nicht als Nutzenparameter im Rahmen einer Bewertung nach §35a SGB V darstellen. Das IQWiG stellt zudem selbst fest, dass in der Literatur verschiedene Limitationen des QALY-Konzepts diskutiert werden, insbesondere ethische, methodische sowie konzeptionelle Einschränkungen. Dabei wird auf verschiedene Literaturquellen verwiesen. Zugleich merkt das IQWiG an, dass das QALY-Konzept dennoch die derzeit am umfassendsten angewandte Methodik darstelle, „die explizit und mit einem theoretisch begründeten Konstrukt die Bewertung der Effektivität von Therapien in gesundheits-ökonomischen Evaluationen unter Berücksichtigung von Präferenzen ermöglicht“. Die erwähnten ethischen (z.B. Akzeptanz und Verwendung bei seltenen Erkrankungen) und methodischen Limitationen (z.B. Änderungssensitivität bei der Nutzwernerhebung) werden jedoch nicht erwähnt, obwohl sie von hoher Relevanz für die theoretische Anwendung in Deutschland wären und zu erheblichen Verzerrungen bei der Informationsbereitstellung für Preisverhandlungen bzw. damit zu Fehlallokationen bei daraus folgenden Erstattungen von Arzneimitteln führen können. Auf S. 85 gibt das IQWiG vor: „Nutzwerte basierend auf Bewertungen durch die Allgemeinbevölkerung sind als dann sinnvoll, wenn sich die Bewertungen nicht von denen durch Patientinnen und Patienten unterscheiden“. Dem wäre zu

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	entnehmen, dass Nutzwerte basierend auf Bewertungen durch die Allgemeinbevölkerung nicht akzeptiert werden, wenn Unterschiede vorliegen. Dabei werden die Unterschiede nicht weiter präzisiert. Zugleich wird außer Acht gelassen, dass solche Fallkonstellationen in der Praxis kaum realisierbar wären. Zugleich wird auf S. 92 angegeben: „Wenn davon auszugehen ist, dass Unterschiede existieren, besteht für die Ergebnisse von Analysen auf Grundlage von Bewertungen durch die Allgemeinbevölkerung die Unsicherheit, dass die Ergebnisse auf Basis von Bewertungen durch Patientinnen und Patienten anders ausfallen könnten. Dies ist bei der Einordnung der Ergebnisse der KNB zu berücksichtigen.“ Hieraus ist abzuleiten, dass auch Analysen auf Grundlage von Bewertungen durch die Allgemeinbevölkerung grundsätzlich akzeptiert werden könnten. <u>Vorgeschlagene Änderung:</u> Klarstellung zur Relevanz der Instrumente EQ-5D, SF-6D sowie HUI 2 oder 3 auch für die Nutzenbewertung nach §35a SGB V. Die ethischen, methodischen sowie konzeptionellen Limitationen des QALY-Konzepts und auch eine mögliche praktische Umsetzbarkeit in Deutschland sollten im Methodenpapier detaillierter diskutiert und dargestellt werden. Die Verwendung der Nutzwerte basierend auf Bewertungen durch die Allgemeinbevölkerung sollte präziser beschrieben werden.
Abschnitt 4.9 Gesundheitsökonomisches Studiendesign, S. 102	<u>Anmerkung:</u> Für die Extrapolation soll in Szenarien u.a. angenommen werden, „dass der in den klinischen Studien beobachtete Effekt/Unterschied zwischen den betrachteten Therapien über das Ende der klinischen Studie hinaus anhält und danach abfällt oder endet.“ Unberücksichtigt bleibt dabei die Möglichkeit einer Zunahme des Effekts. <u>Vorgeschlagene Änderung:</u> Ergänzung des Satzes „dass der in den klinischen Studien beobachtete Effekt/Unterschied zwischen den betrachteten Therapien über das Ende der klinischen Studie hinaus sich verstärkt, anhält, danach abfällt oder endet.“

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

A.1.33 Pfizer Pharma GmbH

Autorinnen und Autoren

- Leverkus, Friedhelm
- Böhme, Sarah
- Berger, Leonie
- Genet, Astrid
- Kauffmann, Stephanie
- Kisser, Agnes
- Kühne, Felicitas
- Kullack, Max
- Kürschner, Niclas
- Rauchensteiner, Stephan
- Trescher, Saskia

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Leverkus, Friedhelm
Böhme, Sarah
Berger, Leonie
Genet, Astrid; Dr.
Kauffmann, Stephanie
Kisser, Agnes; Dr.
Kühne, Felicitas
Kullack, Max
Kürschner, Niclas
Rauchensteiner, Stephan; Dr.
Trescher, Saskia; Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Pfizer Pharma GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Die Pfizer Pharma GmbH (nachfolgend: Pfizer) möchte positiv hervorheben, dass sich das IQWiG dazu entschieden hat, den Entwurf für Version 7.0 seiner Allgemeinen Methoden und die darin enthaltenen Änderungen gegenüber der Version 6.1 zur wissenschaftlichen Diskussion im Rahmen eines Stellungnahmeprozesses zu stellen. Pfizer beteiligt sich an der Diskussion und reicht die folgende Stellungnahme ein.

Im Rahmen der Stellungnahme werden die folgenden Punkte kommentiert:

- Kosten-Nutzen-Bewertung
 - o Verantwortlichkeit für Kosten-Nutzen-Analyse (KNA) vs. Kosten-Nutzen-Analysen-Bewertung (KNB)
 - o QALY basierend auf Nutzwerten
 - o Zielgrößen in der KNB
 - o Einbindung von Surrogatparametern
 - o KNB im internationalen Kontext
- Statistische Signifikanztests: Adjustierung für multiples Testen
- Umgang mit variablen Beobachtungszeiten
 - o Erhebung für PRO und UE nach Behandlungsabbruch oder -wechsel
 - o Unterschiedliche Beobachtungsdauer
- Kriterien für den Einschluss von Studien: Herangehensweise via Effektmodifikation
- Anwendungsbegleitende Datenerhebung (AbD)
 - o Prozess
 - Frühzeitige Abstimmungen und Beratungen für Planungssicherheit bei den langfristigen Projekten einer AbD
 - Fehlen bestimmter Teilaktivitäten des IQWiG im AbD-Prozess
 - Beteiligung von Sachverständigen
 - Veröffentlichung von Zwischenschritten im AbD-Prozess
 - Präzisierung, welche pharmazeutischen Unternehmen und anderen Stakeholder beteiligt werden
 - o Methodik
 - Projektskizze und IQWiG AbD-Konzept
 - Auswahl von Registern
 - Verständnis des Begriffs „zVT“ im Rahmen einer AbD
 - PICO-Schema und Fragen zu nichtinterventionellen versus interventionellen Studiendesigns
 - Bezugnahme auf andere laufende Datenerhebung(en) – Mehrfacheingaben in mehrere Studien vermeiden
 - Studieninterne Datenerhebung als Möglichkeit belassen, da es meist keine Idealtypischen Register in der Realität gibt
 - AbD mit vergleichender Evidenzgenerierung bei seltenen Erkrankungen nicht in jedem Fall umsetzbar
 - Realistische Fallzahlschätzungen und Zeitpläne aufstellen

- Verschieben von Nullhypothesen mit weiterhin „dramatischen“ Effekten bei seltenen Erkrankungen nicht umsetzbar
- Unklarheit über verbindliche und optionale methodische Vorgehensweisen im Rapid Report vs. Methodenpapier, z.B. bei Confounderidentifizierung etc.
- Vorgehen bei kleinen Fallzahlen
- Umgang mit Zwischenanalysen und Vergeblichkeitskriterien unklar
- Instrumentenvariable
- Indirekte Vergleiche mit versorgungsnahen Daten nicht kategorisch ablehnen, wenn RCT nicht möglich sind

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschnagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Kosten-Nutzen-Bewertung (KNB)	
2.1.4 Kosten-Nutzenbewertung gemäß §35b SGB V und 4.1 Einführung (S. 83 ff)	Verantwortlichkeit für Kosten-Nutzen-Analyse (KNA) vs. Kosten-Nutzen-Analysen-Bewertung (KNB) <u>Anmerkung:</u> Das IQWiG kommt seinem gesetzlichen Auftrag gemäß §35b (5) SGB V nach und legt transparente Methoden zur Kosten-Nutzen-Bewertung (KNB) von Arzneimitteln fest. Gesetzlich verankert ist die KNB schon seit 2007, eine praktische Relevanz hat sie bislang nicht erlangt. Nach wie vor stellt die KNB ein letztes Mittel bei der Erstattungspreisfindung dar. Seit 2010 ist sie jedoch noch nie zum Einsatz gekommen. Auch in Zukunft sollte sie nicht zur weiteren Hürde werden. Die Verantwortlichkeit, wer die Kosten-Nutzen-Analyse (KNA) durchführt, bleibt unklar. Im zeitlich begrenzten AMNOG-Verfahren erscheint es unrealistisch, dass KNA regelhaft in ausreichender Qualität durch das IQWiG durchgeführt werden können. Der pharmazeutische Unternehmer (pU) sollte in diesem Fall die Möglichkeit haben, KNA einzureichen. <u>Vorgeschnagene Änderung:</u>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Die Verantwortlichkeiten für KNA und KNB sollten eindeutiger beschrieben werden.
4.6 Nutzwerte (S. 91)	QALY basierend auf Nutzwerten <u>Anmerkung:</u> Das QALY-Konzept wird international kritisch gesehen, so zeigen bspw. Duru et al. 2002 (12), dass das Verhaltensmodell zur Bestimmung der Nutzwerte sich nicht in der realen Population widerspiegelt. Das Konzept beruht auf dem Utilitarismus. Dabei soll der Nutzen über die Gesellschaft maximiert werden. Das geht nicht immer mit den Wertvorstellungen in Deutschland überein. So ergeben sich bspw. für Viagra höhere QALYs als für Bypass-Operationen. Die psychometrischen Eigenschaften werden zum Teil als fraglich angesehen. Die Reliabilität kann gering sein. Verschiedene Methoden zur Messung der Utilities/Nutzwerte liefern verschiedene Ergebnisse. Von daher wird das QALY-Modell zum Teil für gesundheitsökonomische Entscheidungen in Frage gestellt. Das IQWiG beschreibt selbst die Schwierigkeiten des QALY-Konzepts. <u>Vorgeschlagene Änderung:</u> Auf Grund der kritischen Aspekte des QALY-Konzepts sollte der Fokus nicht auf diesem Konzept liegen. Es sollte klarer beschrieben werden, dass QALYs genutzt werden können, um eine Lebenserwartung und Lebensqualität miteinander zu verbinden. Positive Effekte in der Lebenszeit und bei klinischen Endpunkten sowie Surrogatparametern sollten auch aufgezeigt werden können.
4.12 Einordnung des ICERs (S. 107)	Zielgrößen in der KNB <u>Anmerkung:</u> Wir begrüßen den Verzicht auf strenge Effizienzgrenzen. Allerdings wird zwar das methodische Vorgehen des ICER beschrieben, aber über die Beurteilung des ICERs und zu welchen Schlussfolgerungen das konkrete Ergebnis eines ICERs und diesbezüglicher Sensitivitätsanalysen führt, bleibt offen. Einerseits wird dargestellt, dass mit der KNB weder ein Grad eines Zusatznutzens festgestellt noch ein konkreter Preis errechnet werden soll. Es bleibt unklar, wozu der ICER-

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Wert dann tatsächlich in der deutschen Preisfindung dienen sollte. <u>Vorgeschlagene Änderung:</u> Neben dem ICER sollten auch andere Zielgrößen und Methoden beschrieben werden, bei denen man evtl. die klinischen Nutzen-Schaden-Abwägungen - in die Zukunft extrapoliert - genauer in den Fokus nimmt (21).
4.3.3 Gesundheitsökonomische Zielgrößen (S. 141 f.)	Einbindung von Surrogatparametern <u>Anmerkung:</u> Das IQWiG beschreibt, dass die gesundheits- ökonomischen Modelle auf entscheidungsrelevanten Endpunkten der Nutzenbewertung beruhen sollen. Die Information aus Surrogatparametern bleibt dadurch unberücksichtigt. Das NICE kam zu der Erkenntnis, dass durch die Nicht-Berücksichtigung der Surrogatparameter die Effekte der verbleibenden patientenrelevanten Endpunkte – wie Gesamtüberleben – zu einem außergewöhnlich hohen ICER und daher einer Fehlbeurteilung der Erstattungsfähigkeit führte (25). <u>Vorgeschlagene Änderung:</u> Es müssen mehr als nur die entscheidungsrelevanten Endpunkte aus der Nutzenbewertung in die KNA- Modelle eingehen, z.B. ist die Information aus Surrogatparametern und deren Übergangs- wahrscheinlichkeit zu anderen Outcomes zu berück- sichtigen.
4. Kosten-Nutzen-Bewertung medizinischer Interventionen (S. 83 ff.)	KNB im internationalen Kontext <u>Anmerkung:</u> Pfizer begrüßt, dass im Rahmen der Änderungen die gegenwärtigen aktuellen wissenschaftlichen Erkenntnisse und Methoden zu gesundheits- ökonomischen Analysen im internationalen Kontext mit einfließen. Pfizer gibt allerdings zu bedenken, dass andere europäische Gesundheitssysteme die KNA zwar vergleichbar durchführen, die KNB aber sehr unterschiedlich gehandhabt wird. Als Beispiele wären hier Großbritannien und Frankreich zu nennen. Aus

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Großbritannien ist bekannt, dass es unter Anwendung der KNB-Methodik für neue, höherpreisige Arzneimittel zunehmend herausfordernd ist, den realen Wert der Innovationen abzubilden (30,32). Ferner stellt im aktuellen Entwurf die KNB „im Kontext des AMNOG-Verfahrens ein Instrument dar, das nach abgeschlossener früher Nutzenbewertung von Arzneimitteln zur Orientierung bei der Preisfindung für ein neues Arzneimittel dienen kann.“ Pfizer ist diesbezüglich skeptisch, was die Erfolgsaussichten einer möglichen KNB angeht: da eine bereits abgeschlossene Nutzenbewertung die Ausgangsbasis darstellen soll, scheint eine systematische Benachteiligung des zu prüfenden Arzneimittels wahrscheinlich, wenn das Ergebnis der Nutzenbewertung „kein Zusatznutzen“ lautet, insbesondere wenn der primäre Studienendpunkt im Rahmen des Verfahrens nicht als patientenrelevant akzeptiert wird. Als Beispiele hierfür wären die in klinischen Studien weit verbreiteten und von den Zulassungsbehörden anerkannten primären Endpunkte progressionsfreies Überleben (progression free survival, PFS) bei soliden Tumoren oder die (molekulare) Ansprechrate bei hämatologischen Tumoren zu nennen. <u>Vorgeschlagene Änderung:</u> Die Änderungen sollten hinsichtlich konkreter Anwendbarkeit in der Praxis überprüft werden.
Statistische Signifikanztests	
9.3.2 Statistische Signifikanztests und Konfidenzintervalle zur Entscheidungsunterstützung (S.170)	Adjustierung für multiples Testen <u>Anmerkung:</u> Das IQWiG beschreibt in Abschnitt 9.3.2 „Statistische Signifikanztests und Konfidenzintervalle zur Entscheidungsunterstützung“, dass für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist, Methoden zur Adjustierung für multiples Testen angewendet werden sollten. Die konkreten Anwendungsbedingungen für die beschriebenen Adjustierung für multiples Testen sind

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	aus Sicht von Pfizer auf Basis der Ausführungen im aktuellen Entwurf 7.0 unklar und in dieser Form im Nutzenbewertungsdossiers nicht anwendbar. Daher ist eine Konkretisierung dessen, was das IQWiG unter „einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist“ versteht, erforderlich. Im Rahmen der Nutzenbewertung wird vom pU routinemäßig für jeden Endpunkt eine primäre Analyse dargestellt, die zur Ableitung des Zusatznutzens dient. Die Operationalisierung des Endpunktes der primären Analyse ist normalerweise diejenige, die der im Protokoll präspezifizierten Version des Endpunktes am nächsten kommt. Hinzu kommen Sensitivitätsanalysen und supportive Analysen auf Basis weiterer Operationalisierungen dieses Endpunktes zur Feststellung der Empfindlichkeit des Effektes auf Veränderungen der Methodik, die jedoch nicht entscheidungsrelevant sind und nicht zur Ableitung des Zusatznutzens herangezogen werden. Typische Beispiele sind Responderanalysen im Vergleich zu einer stetigen Auswertung, der Vergleich von unterschiedlichen Imputationsstrategien für fehlende Werte, die Betrachtung von unterschiedlichen Estimands in der Anwesenheit von Intercurrent Events oder die Darstellung von weiteren a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten. Nach Ansicht Pfizers erfordert die Darstellung von Sensitivitätsanalysen keine Änderungen der Entscheidungsgrenzen der primären Analyse, d.h. eine Adjustierung für multiples Testen ist in diesem Fall nicht angebracht. Ein Beispiel, bei dem eine Korrektur des Signifikanzniveaus (α-Niveaus) aus Sicht von Pfizer erforderlich sein könnte, ist z. B., wenn für denselben PRO-Endpunkt zwei Responderanalysen als primäre Analysen vorgelegt werden, eine mit der im Protokoll präspezifizierten Responseschwelle und eine mit einer Responseschwelle, die mind. 15 % der Skalen-spannweite des verwendeten Erhebungsinstruments entspricht. Würden beide Operationalisierungen in die Entscheidung über einen Zusatznutzen einfließen, dann könnte eine Korrektur des α-Niveaus relevant sein.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Im Gegensatz dazu ist keine Adjustierung für multiples Testen erforderlich, wenn sowohl Progressionsfreies Überleben als auch Tod/Gesamtüberleben dargestellt werden. Es handelt sich hierbei um zwei unterschiedliche Endpunkte, die zur Dimension Morbidität bzw. Mortalität gehören. <u>Vorgeschlagene Änderung:</u> Generell kann die Empfehlung, eine Korrektur des Signifikanzniveaus bei Endpunkten, die mithilfe verschiedener Operationalisierungen untersucht worden sind, anzuwenden, beibehalten werden. Es sollte jedoch näher spezifiziert werden, dass diese nur für mehrere primäre Analysen für einen Endpunkt dient, sofern die Ableitung des Zusatznutzens auf mehr als einer Operationalisierung basiert. Die Darstellung von Sensitivitätsanalysen jeder Art erfordert keine Korrektur des Signifikanzniveaus.
Umgang mit variablen Beobachtungszeiten	
9.3.12 Umgang mit variablen Beobachtungszeiten (S. 190 f.)	Erhebung für PRO und UE nach Behandlungs- abbruch oder -wechsel <u>Anmerkung:</u> Das IQWiG fordert in Abschnitt 9.3.12 „Umgang mit variablen Beobachtungszeiten“ eine vollständige Erfassung der Daten zu patientenberichteten Endpunkten (PRO) und unerwünschten Ereignissen (UE) bis zum Studienende oder Tod, auch nach einem Behandlungsabbruch oder -wechsel. Diese sei im Rahmen der Nutzenbewertung erforderlich, um die Daten auf Basis des Treatment Policy Estimand und des Intention-to-treat (ITT) Prinzips auswerten zu können, auch bei unterschiedlichen Beobachtungsdauern in den Studienarmen aufgrund von Behandlungsabbruch oder -wechsel. Pfizer stimmt dieser Forderung nicht zu. Eine weitere Erfassung der PRO und UE nach einem Behandlungsabbruch oder -wechsel ist im Kontext der Fragestellung der Nutzenbewertung nicht zielführend, was im Folgenden begründet wird. Die Erhebung der PRO und UE bis zum Studienabbruch erfolgt in beiden Studienarmen gemäß ITT-Prinzip, das

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	kausale Rückschlüsse erlaubt. Das ITT-Prinzip wird durch die Zuteilung aller randomisierten Patient:innen auf die jeweilige Behandlungsgruppe (und nicht auf die tatsächlich erhaltene Behandlung) definiert, nicht über den Follow-up (bei dem alle Patient:innen unabhängig von Protokollabweichungen weiterverfolgt würden, um vollständig beobachtete Daten zu erhalten). Das ITT-Prinzip kann als eine Möglichkeit angesehen werden, mit Selektionsverzerrungen umzugehen, da die Randomisierung beibehalten wird. Es ist jedoch nicht dazu geeignet, fehlende Daten oder das Verhalten nach der Randomisierung spezifisch zu betrachten (26). Ein vollständiges Follow-up der Patient:innen, die die Studie protokollgemäß bis zum geplanten Studienende beendet haben, ist somit nicht notwendig, um die Daten nach ITT-Prinzip auswerten zu können. Bei onkologischen Indikationen enthält eine leitlinien-konforme Behandlung eine Sequenztherapie. In diesem Fall gehören das Ende der Therapie und der Behandlungswechsel zum Therapieplan. Eine weitere Erhebung wäre in diesem Fall nicht zielführend – entscheidungsrelevant für die Fragestellung der Nutzenbewertung ist die Beurteilung der Patient:innen in dem Teil der Sequenz, in dem die neuen Wirkstoffe eingesetzt werden. Die ursprüngliche Fragestellung könnte durch die Effekte von Folgebehandlungen nicht mehr adäquat beantwortet werden. Es bliebe unklar, wie Daten, die beispielsweise nach einem Behandlungswechsel erhoben wurden, überhaupt in Bezug auf den Vergleich zwischen Intervention und Kontrolle ausgewertet werden sollten, wenn sich neue Therapien in beiden Armen anschließen, deren Effekt komplett verschieden auf die betrachteten Endpunkte ausfallen kann. Auch wenn die weitere Erhebung der PRO und UE nach Therapieabbruch in manchen Indikationen und Fragestellungen theoretisch relevante Informationen liefern könnte, halten wir diesen Ansatz in der Praxis aus mehreren Gründen für nicht umsetzbar. Erstens sind Patient:innen nach einem Abbruch nicht mehr unter der Behandlung von Interesse und auch nicht mehr verblindet (bei verblindeten Studien). Dies könnte die Antworten auf PRO-Fragebögen sowie die UE-Meldung

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	nach dem Abbruch beeinflussen und dazu führen, dass PRO und UE, die vor und nach Therapieabbruch erhoben wären, systematisch verschieden und von daher nicht mehr vergleichbar wären. Zweitens könnten die Patient:innen auch auf eine andere Studie wechseln. Dies führt zu einer weiteren Limitation, denn um an einer anderen Studie teilnehmen zu können, müssen die Patient:innen ihre Einwilligung für die erste Studie i.R. zurückziehen. In diesem Fall ist eine weitere Erhebung der Studienendpunkte noch schwieriger. Somit sinken auch die Rücklaufquoten für PRO Fragebögen, wodurch die Ergebnisse nicht mehr bewertbar wären. Wir denken, dass die Fragestellung der Nutzenbewertung durch die Erfassung der PRO und UE nach Behandlungsabbruch oder -wechsel nicht immer richtig beantwortet wäre. Daher wird die im Entwurf der Allgemeinen Methoden 7.0 beschriebene Aussage, dass Beobachtungen „systematisch“, also gezielt und geplanter Weise verkürzt wären, von Pfizer sehr kritisch gesehen. <u>Vorgeschlagene Änderung:</u> Die Betrachtung von PRO und UE bis zum Behandlungsende ist angemessen, um die Daten im Rahmen der Nutzenbewertung unverzerrt gemäß ITT-Prinzip auswerten zu können. Die weitere Erhebung der PRO und UE nach Therapieabbruch ist in der Praxis meist nicht umsetzbar und wird mit großer Wahrscheinlichkeit verzerrt sein, da die Patient:innen nicht mehr unter der Behandlung von Interesse und auch nicht mehr verblindet sind.
9.3.12 Umgang mit variablen Beobachtungszeiten (S. 190 f.)	Unterschiedliche Beobachtungsdauern <u>Anmerkung:</u> Das IQWiG kündigt in Abschnitt 9.3.12 „Umgang mit variablen Beobachtungszeiten“ an, bei verkürzter oder deutlich unterschiedlicher Beobachtungsdauer zwischen den Armen die dauerhafte Verschlechterung für PRO i.d.R. nicht für die Nutzenbewertung heranzuziehen. Weiterhin seien bei unterschiedlich langer Beobachtungsdauer zwischen den zu vergleichenden Gruppen Datenanalysen zu PRO und UE insgesamt nicht für die Nutzenbewertung verwertbar.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Pfizer begrüßt prinzipiell, dass die Klarstellung bezüglich des Umgangs mit variablen Beobachtungszeiten Eingang in den Entwurf der Allgemeinen Methoden 7.0 gefunden hat, da die beschriebene Methodik bereits seit einiger Zeit gelebte Praxis in der Nutzenbewertung war (4). Pfizer ist hingegen nicht der Meinung, dass bei unterschiedlich langer Beobachtungsdauer zwischen den zu vergleichenden Gruppen die Datenanalysen zu PRO und UE insgesamt nicht für die Nutzenbewertung verwertbar seien. Bei Endpunkten, bei denen relevante unterschiedliche Beobachtungszeiten zwischen den Studienarmen vorliegen, ist die Zeit bis zur erstmaligen oder bestätigten Veränderung ein grundsätzlich inhaltlich sinnvoller Endpunkt, welcher Meinung auch der G-BA ist (4). Nicht sinnvoll in einem solchen Fall wäre die Anwendung einfacher Methoden basierend auf relativen Häufigkeiten oder Inzidenzdichten. Wenn die Erhebung der PRO und UE bis zum Studienabbruch ähnlich in beiden Studienarmen gemäß ITT-Prinzip erfolgte, sind kausale Rückschlüsse erlaubt, sofern adäquate Methoden der Überlebenszeitanalyse angewendet werden. Bei Hormesis-Prozessen beispielsweise wäre die Zeit bis zur erstmaligen Veränderung nicht repräsentativ für den gesamten Effekt unter Behandlung. Die Verläufe der beobachteten Veränderungen werden standardmäßig im Dossier dargestellt und sollten die Grundlage für die Wahl der adäquaten Auswertung sein. Die Zeit bis zur erstmaligen oder (wenn die Rücklaufquoten es erlauben) bis zur bestätigten Veränderung soll weiterhin vom pU im Dossier vorgelegt werden und vom IQWiG und G-BA für die Nutzenbewertung herangezogen werden. Auch bei unterschiedlichen Beobachtungszeiten zwischen den Studienarmen, die z.B. infolge eines vorzeitigen Beobachtungsendes aufgrund von Krankheitsprogression erfolgen, sollen die PRO und UE in der Gesamtaussage zum Zusatznutzen aufgrund ihrer hohen Patientenrelevanz eine wichtige Rolle spielen. <u>Vorgeschlagene Änderung:</u> Es sollte u.a. auch der Verlauf der beobachteten Veränderungen als Grundlage für die Wahl der geeignetsten Analysemethodik gewählt werden. Die Zeit

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	bis zur erstmaligen oder (wenn die Rücklaufquoten es erlauben) bis zur bestätigten Veränderung ist in diesem Fall für die Nutzenbewertung verwertbar.
Kriterien für den Einschluss von Studien	
9.1.1 Kriterien für den Einschluss von Studien (S. 159)	Herangehensweise via Effektmodifikation <u>Anmerkung:</u> Das IQWiG beschreibt in Abschnitt 9.1.1 „Kriterien für den Einschluss von Studien“, dass in bestimmten Fällen eine Effektmodifikation zur Beurteilung der Berücksichtigung der Studie heranzuziehen sei. Pfizer interpretiert den neugefassten Absatz so, dass die bestehende Regel für den Einschluss von Studien (es muss auf mindestens 80 % der eingeschlossenen Patient:innen das Einschlusskriterium (Population) zutreffen bzw. das Einschlusskriterium bezüglich der Prüfintervention/ Vergleichsintervention jeweils zu mindestens 80 % erfüllt sein) unverändert bleibt. Für Pfizer ist jedoch unklar, ob sich die Herangehensweise über die Effektmodifikation nur auf Studien bezieht, bei denen die vorab beschriebene 80%-Regel eintrifft, oder ob alle potenziell einzuschließenden Studien auf dieses Merkmal hin untersucht werden sollen. Weiterhin bleibt unklar, wie stark die Effektmodifikation sein müsste bzw. wie groß der Grad der Abweichung der Interventionen sein müsste, um einen Ein- bzw. Ausschluss einer Studie zu rechtfertigen. <u>Vorgeschlagene Änderung:</u> Pfizer bittet um eine Klärung bezüglich des Einsatzes der Herangehensweise über die Effektmodifikation, sowie um Spezifikation der akzeptablen Schwellenwerte für die „Stärke der Effektmodifikation“ sowie den „Grad der Abweichung der Intervention“.
Anwendungsbegleitende Datenerhebung (AbD)	
2.1 Produktspezifische Verfahrensabläufe (S. 11 ff.)	Prozess: Frühzeitige Abstimmungen und Beratungen für Planungssicherheit bei den langfristigen Projekten einer Abd <i>„Zu Bewertungen gemäß § 35a SGB V, in dem die Bewertung des Nutzens neuer Wirkstoffe auf Basis eines</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>Dossiers des pharmazeutischen Unternehmers geregelt ist, gehören Dossierbewertungen (siehe auch Abschnitt 3.3.3), AbD-Konzepte (siehe auch Abschnitt 3.3.4) sowie Recherchen nach Indikationsregistern zur Vorbereitung der Entscheidungsfindung des G-BA hinsichtlich der Beurteilung der Erforderlichkeit der Forderung einer AbD. Auftraggeber ist jeweils der G-BA. Für Bewertungen gemäß § 35a sind keine Anhörungen durch das Institut vorgesehen. Für Dossier- bewertungen und AbD-Konzepte erfolgt ein Stellungnahmeverfahren im weiteren Verfahren durch den G-BA.“</i> <u>Anmerkung:</u> Im derzeitigen Prozess erfolgt die Abstimmung zu den Anforderungen des IQWiG und G-BA nicht frühzeitig und verbindlich genug. Das IQWiG ist in der G-BA-Beratung nicht involviert, bestimmt aber die meisten methodischen Anforderungen im AbD-Prozess. Eine frühzeitige Einbindung aller Stakeholder (insbesondere auch Hersteller, Register-Betreiber, Zulassungsbehörden, Versorgungsforscher:innen, Methodiker:innen/ Statistiker:innen, Expert:innen für klinische nicht- interventionelle und interventionelle Studien) erfolgt nicht, was zu unrealistischen Anforderungen für eine Datenerhebungen führen kann. Eine Einschätzung zur Realisierbarkeit der konkret vorgeschlagenen AbD erfolgt nicht. <u>Vorgeschlagene Änderung:</u> Durch frühzeitige und verbindliche Beratungen des G- BA kann der aktuell noch hohen Planungsunsicherheit entgegengewirkt werden. Ein frühzeitiger Austausch hinsichtlich methodischer Anforderungen ist wichtig, und zwar bereits bei der ersten G-BA-Beratung, bei der Erarbeitung des IQWiG-Konzepts und insbesondere bei der Erarbeitung des Studienprotokolls. Alle Stakeholder sollten bei der Konzepterstellung eingebunden werden, also auch Register-Betreiber, Zulassungsbehörden, Industrieverbände und betroffene Unternehmen. Noch vor Finalisierung der AbD-Konzepte sollte es auch frühzeitig Anhörungen und Beratungen durch das IQWiG (bzw. gemeinsam mit G-BA und IQWiG) mit dem pU geben. Während der Erarbeitung des IQWiG-Konzepts und auch des Studienprotokolls und statistischen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Analyseplans (SAP) des pU sollte es Beratungsmöglichkeiten durch G-BA und IQWiG für pU und Register-Betreiber geben. Zudem kann es während einer bereits laufenden AbD zu Herausforderungen und notwendigen Änderungen kommen, die auch mit IQWiG und G-BA abgesprochen werden sollten. Dies ist aufgrund der Tragweite und Auswirkungen auf viele Beteiligte im deutschen Gesundheitssystem besonders bei solch langwierigen AbD wichtig. Es sollte eine Einschätzung zum Grad der Realisierbarkeit der AbD von IQWiG und G-BA zusammen mit externen Sachverständigen erfolgen und realistischerweise mit Expert:innen auch eine studienindividuelle, u.a. minimal-interventionelle Datenerhebung diskutiert werden, falls der Einbau in einer Register zu aufwändig ist oder die Zulassungs- und HTA-Anforderungen an valide Daten mit einem non-interventionellen Register nicht erfüllt werden können.
2.1.3.2 AbD-Konzept (S. 20)	Prozess: Fehlen bestimmter Teilaktivitäten des IQWiG im AbD-Prozess <u>Anmerkung:</u> In der Prozessbeschreibung fehlen neben dem reinen AbD-Konzept die anderen Aufgaben und Aktivitäten des IQWiG, z.B. Review des Studienprotokolls und später sicher auch die Bewertung der AbD-Auswertungen. <u>Vorgeschlagene Änderung:</u> Die weiteren Schritte, die durch das IQWiG erfolgen können oder müssen, sollten explizit auch alle im Prozessdiagramm genannt werden.
2.1.3.2 AbD-Konzept (S. 20 f.) und 2.1.1 Bericht (S. 32) „Auftragserteilung“	Prozess: Beteiligung von Sachverständigen <i>„Dabei werden regelhaft externe Sachverständige beteiligt (s. Abschnitt 2.2.2). Bei Bedarf wird auch der Wissenschaftliche Beirat des Instituts einbezogen. [...]“</i> <i>„Bei der Erarbeitung des AbD-Konzepts werden externe Sachverständige für die medizinische Expertise sowie Betroffene bzw. Patientenorganisationen für die Patientenperspektive eingebunden. Dies erfolgt primär über Fragebogen analog der Einbindung bei Dossierbewertungen (siehe Abschnitt 2.1.3.1). Als Teil der Informationsbeschaffung werden zudem die</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>Betreiber bereits bestehender Datenplattformen kontaktiert, die für die Durchführung der geplanten AbD nach erster Einschätzung potenziell geeignet sind (insbesondere Indikationsregister). Nach der Fertigstellung wird das AbD-Konzept an den G-BA versendet. Zeitnah im Anschluss daran erfolgen die Information des Vorstands der Stiftung, des Stiftungsrats und des Kuratoriums der Stiftung über den Versand. Das AbD- Konzept wird nach Beschluss des G-BA zur AbD auf der Website des Instituts veröffentlicht.“</i> <u>Anmerkung:</u> Der Auswahlprozess und Einbezug von externen Sachverständigen und Registerbetreibern wird zu allgemein dargestellt. Es wird nicht klar, welche externen Sachverständigen bei der Erstellung eines AbD- Konzepts obligatorisch und welche optional beteiligt werden. Wie genau bei der Auswahl von Registern und der Befragung von Registerbetreibenden und damit in Beziehung stehenden Expert:innen vorgegangen wird, bleibt intransparent. Dass externe Sachverständige und Betroffene nach Erhalt des IQWiG-Konzepts nur 4 Wochen Zeit haben, um eine Stellungnahme dazu einzureichen, ist nicht erwähnt. Die 4 Wochen sind außerdem viel zu kurz und es ist nicht nachvollziehbar, nach welchen Kriterien bei der Auswahl der Firmen, die bei der AbD beteiligt werden, vorgegangen wird. So wurden bisher manche Firmen mit zugelassenen Produkten in der Indikation nicht beteiligt, während andere Firmen ohne Phase III- Studien oder Produkten in der Vermarktung jedoch beteiligt wurden. Industrieverbände werden ebenfalls nicht beteiligt. <u>Vorgeschlagene Änderung:</u> Der Einbezug von Sachverständigen und Registerbetreibern sollte transparenter dargestellt werden. Es sollte explizit definiert werden, welche externen Sachverständigen genau beteiligt werden. Für AbD sind insbesondere Expert:innen der Versorgungsforschung, der Methodik bei der Analyse von Daten aus nicht-interventionellen Studien, medizinisch-naturwissenschaftliche Fachexpert:innen, Patientenorganisationen, Kliniker:innen und Klinik- personal, welches die Daten in Praxen oder

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Krankenhäusern letztlich erfassen muss, definitiv zu beteiligen. Die Kontaktaufnahme mit Betreibenden potenziell für die AbD geeigneter Datenplattformen muss einem transparenten Prozess folgen, der regulatorische und gesetzliche Bestimmungen berücksichtigt, den pharmazeutische Unternehmen und andere Beteiligte unterliegen, so dass nicht eine Sonderstellung eines einzelnen Register- oder Softwarebetreibers oder einer Agentur erzeugt wird. Industrieverbände wie vfa und bpi sollten beteiligt werden, da jede AbD mehrere Firmen aufgrund von Produkten in der zVT oder sich in Entwicklung befindenden Produkten betreffen kann. Die Frist für Beteiligungsverfahren und Fachausschüsse sollte hier auch genannt werden, und die bisherige Frist von 4 Wochen sollte verlängert werden, damit alle Beteiligten ausreichend viel Zeit haben, Stellung zu beziehen. Der betroffene pU sollte auch die Möglichkeit haben, Beteiligte und Expert:innen vorzuschlagen.
2.1.3.3. Recherche nach Indikationsregistern (S. 41)	Prozess: Veröffentlichung von Zwischenschritten im AbD-Prozess <i>„Der Auftrag des G-BA beschreibt die Indikation(en), für die eine Recherche nach Indikationsregistern stattfinden soll. In der Projektskizze werden die wesentlichen Schritte zur Informationsbeschaffung zusammengefasst. Bei der Recherche nach Indikationsregistern erfolgt innerhalb von 4 Wochen eine fokussierte Informations- beschaffung in verschiedenen Informationsquellen. Nach der Fertigstellung wird das Dokument an den G- BA versendet. Eine Veröffentlichung auf der Website des Instituts findet statt, wenn der G-BA ein Verfahren zur Forderung einer AbD im Indikationsgebiet einleitet.“</i> <u>Anmerkung:</u> Die Ergebnisse des IQWiG zu Registerrecherchen sowie das IQWiG-Konzept werden erst nach G-BA Beschluss veröffentlicht. So haben andere Unternehmen keine Möglichkeit zur frühzeitigen Planung, falls sie ein Produkt in der Indikation in der Entwicklung hätten. <u>Vorgeschlagene Änderung:</u> Eine Veröffentlichung von Registerrecherchen und AbD- Konzepten sollte immer direkt nach Abschluss des

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Berichts stattfinden, und nicht erst nach finalem G-BA Beschluss der AbD. Nur so können sich andere Registerbetreiber und betroffene pharmazeutische Unternehmer über die klaren Kriterien oder ggfs. notwendigen Anpassungen von Registern transparent und frühzeitig informieren.
2.2.5 Stellungnahmeverfahren (S. 40)	Prozess: Präzisierung, welche pharmazeutischen Unternehmen und anderen Stakeholder beteiligt werden <u>Anmerkung:</u> Für das AbD-Konzept scheint es andere Vorgehens- weisen bei der Stellungnahme/ Beteiligung zu geben als bei der üblichen Stellungnahmephase und Anhörung in der regulären Nutzenbewertung. Beispielsweise ist das Beteiligungsverfahren und die Fachaustausch-Anhörung dazu nicht öffentlich und nicht klar, warum welche Stakeholder beteiligt bzw. nicht beteiligt wurden. <u>Vorgeschlagene Änderung:</u> Es sollte eindeutig klargestellt werden, welche Regeln bei der Auswahl der beteiligten pU und anderer Stakeholder angewandt wird, z.B. Produkt in Vermarktung oder Produkt in Phase III Studie. Da eine AbD meist mehrere Beteiligte im Gesundheitssystem über mehrere Jahre betrifft, sollten auch externe Sachverständige, insbesondere Expert:innen der Versorgungsforschung, Kliniker und Registerbetreiber sowie Industrieverbände eine Stellungnahme abgeben können, genauso wie bei Stellungnahme-Verfahren zu Nutzenbewertungen üblich. Das Stellungnahme- verfahren sollte daher öffentlich sein.
2.1.3.2 Abd-Konzept (S. 21)	Methodik: Projektskizze und IQWiG Abd-Konzept <i>„In der Projektskizze werden die wesentlichen Schritte zur Informationsbeschaffung und der wissenschaftlichen Bewertung zusammengefasst. Die Projektskizze ist Grundlage für die Erstellung des Abd-Konzepts.“</i> <u>Anmerkung:</u> Einige der bisherigen IQWiG-AbD-Konzepte (5) oder Projektskizzen enthielten sehr allgemeine Aussagen zu Endpunkten, ohne konkrete Nennung von einzelnen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Symptomen oder Erhebungsinstrumenten, und ohne klare Aussagen zu den relevanten Confoundern, deren Identifizierung oder Adjustierung bei dieser Erkrankung notwendig ist. Ohne Informationen zu relevanten Confoundern ist selbst die Einschätzung, welches Register oder welche Datenbank für die AbD geeignet wäre, nicht abschließend möglich. <u>Vorgeschlagene Änderung:</u> Welche Aspekte in der Projektskizze und dem IQWiG-Konzept erfasst werden, sollte explizit spezifiziert werden. Bereits in der Projektskizze bzw. IQWiG-AbD-Konzept sollten erste konkrete Confounder bei der jeweiligen Erkrankung bzw. Therapie von IQWiG spezifiziert werden und grundlegende Überlegungen zur Confounder-Identifizierung, Relevanz der Confounder und Auswertung inkl. Endpunkte zumindest in Ansätzen klar sein.
2.1.3.3. Recherche nach Indikationsregistern (S. 41)	Methodik: Auswahl von Registern <u>Anmerkung:</u> Es bleibt unklar, warum bestimmte Register, die manchmal sogar noch Mängel aufweisen, bevorzugt ausgewählt werden, während andere Register, die ebenfalls Mängel aufweisen, Änderungen vornehmen müssten. Ob internationale Register in Frage kommen, bleibt unklar. Bei manchen seltenen Erkrankungen funktionieren Datenerhebungen und -auswertungen nicht einwandfrei, wenn nicht auch Patient:innen aus internationalen Registern einbezogen werden. <u>Vorgeschlagene Änderung:</u> Es sollte genauer und transparent spezifiziert werden, welche Kriterien bei der Registerrecherche angewandt werden und welche Abweichung von theoretischen Idealregistern vom IQWiG generell akzeptiert werden und welche nie akzeptiert werden würden. Eine Einschätzung zur Realisierbarkeit sollte von unabhängigen Expert:innen erfolgen, die weder bei pU, IQWiG/G-BA, noch Registerbetreibern arbeiten oder mit denen in Beziehung stehen. Es sollte klar definiert werden, ob deutsche Register bei einer Erkrankung ausreichen oder aufgrund der Seltenheit der Indikation oder der Therapielinie nur

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	internationale Registerlösungen möglich wären. Welche Länder für IQWiG dann im Sinne einer Ähnlichkeit mit dem deutschen Gesundheitssystem akzeptabel wären, sollte genauer spezifiziert werden. Eine verbindliche Absprache des IQWiG mit Registerbetreibern und pU sollte vorzeitig erfolgen und nicht erst nach Beauftragung der AbD.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß §35a Abs. 3b SGB V (AbD Konzepte) (S. 66)	Methodik: Verständnis des Begriffs „zVT“ im Rahmen einer AbD <i>„Das AbD-Konzept soll zum einen die methodischen Anforderungen an die Durchführung der AbD einschließlich der Auswertung zum Zwecke der Nutzenbewertung gemäß § 35a SGB V gegenüber einer zweckmäßigen Vergleichstherapie (siehe Abschnitt 3.3.3) beschreiben.“</i> <u>Anmerkung:</u> Der Begriff der zweckmäßigen Vergleichstherapie (zVT) ist nicht eins zu eins aus Zulassungsstudien übertragbar auf versorgungsnahe Daten und die AbD. Da sich bei einer nicht-randomisierten, nicht-interventionellen Studie die Verfügbarkeit und Auswahlmöglichkeiten für Therapien unkontrollierter ändern können als bei einer interventionellen Zulassungsstudie, kann man hier nicht von einer zVT im klassischen Sinne sprechen. Zudem werden anscheinend keine Fachgesellschaften bei der zVT-Bestimmung für das AbD-Konzept beteiligt, obwohl das mittlerweile gesetzlich verbindlich ist. Eine Herausforderung ist die Ähnlichkeit der Vergleichs- kohorte (zVT-Gruppe einer AbD): manchmal sind die Patient:innen in Verum-Arm und Vergleichsarm nicht ähnlich genug, um überhaupt eine Confounder- adjustierung mit statistischen Methoden durchführen zu können. In manchen Fällen verändern sich die Versorgungsanteile von Produkten in der zVT dramatisch oder es steht gar keine Alternativtherapie zur Verfügung, so dass eine AbD gar nicht dauerhaft realistisch als Vergleichsstudie durchgeführt werden kann. Manchmal sehen selbst Fachinformation, Leitlinien und Ärzt:innen die Produkte in der Versorgung gar nicht als ähnlich an und damit ist kein einwandfreier Vergleich möglich. Zudem sind für die zVT Versorgungseinschränkungen anscheinend gesetzlich nicht möglich, daher ist die

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Rekrutierung von zVT-Patient:innen nicht in der gleichen Weise wie im Verum-Arm regelhaft möglich, für den die G-BA VerFO die Versorgungseinschränkung gesetzlich zulässt. Das macht den Vergleich dann unausgewogen. <u>Vorgeschlagene Änderung:</u> Der Begriff einer zVT muss für das AbD Konzept klarer definiert werden, denn im Gegensatz zu Zulassungsstudien ändert sich bei der langfristigen AbD-Studie das Therapieumfeld oder Therapieentscheidungen, Risikofaktoren und Leitlinien über den Zeitverlauf einer AbD, so dass im ungünstigsten Falle evtl. gar keine Patient:innen mehr für die zVT zur Verfügung stehen könnten oder nicht ähnlich lange beobachtet werden können wie im Verum-Arm oder sogar im Laufe der Zeit anders adjustiert werden müssten, weil es andere Confounder gibt als zu Beginn der AbD Studie.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) (S. 67)	Methodik: PICO-Schema und Fragen zu nicht- interventionellen versus interventionellen Studien- designs <i>„Definition der wissenschaftlichen Fragestellungen, die sich aus dem Informationsbedarf für die Nutzen- bewertung gemäß § 35a SGB V ergeben: Diese werden im PICO-Format (Population, Intervention, Comparison, Outcome) beschrieben.“</i> <u>Anmerkung:</u> Beim PICO-Schema sind manche Angaben im IQWiG-Konzept zu allgemein gehalten und relevante Confounder werden nicht definiert. Ob eine Messung von jedem einzelnen Endpunkt zu festgelegten Zeitpunkten im nichtinterventionellen Design eines Registers/AbD überhaupt möglich ist, wird nicht geklärt. Eventuell könnten minimal-interventionelle klinische Prüfungen realistischer sein, welche jedoch dann unter die Gesetzgebung und Richtlinien für klinische Prüfungen fallen (10,19,27). <u>Vorgeschlagene Änderung:</u> Neben dem PICO-Schema sollten auch das Studiendesign, realistischer Zeitpunkt der non-interventionellen Messungen der verschiedenen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Endpunkte, und auch speziell die relevanten Confounder im AbD-Konzept bereits erarbeitet werden. Denn für eine nicht-interventionelle AbD lässt sich die Einschätzung, ob ein bereits bestehendes Register potenziell geeignet ist, nur mit dem Wissen zu den relevanten Confoundern und ob diese in dem Register bereits erhoben worden sind, treffen. Wenn ein Register nicht von vornherein alle relevanten Confounder in ihrer Registerstruktur erhebt, ist es sehr unrealistisch, dass zum Launch eines Produkts eine Registerstruktur verfügbar ist, in der alle Informationen zu Confoundern bereits erhoben sind. Bei manchen Registern müssen alle Patient:innen eventuell nochmals einwilligen, sobald neue Case Report Form (CRF)-items zu Patientencharakteristika, Confoundern oder Outcomes geschaffen werden müssen. Die Frage, ob ein nichtinterventionelles Design überhaupt zum Ziel führt und möglich ist, oder ob (minimal-)interventionelle Studiendesigns realistischer sind, sollte bereits frühzeitig im IQWiG Konzept erörtert werden.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) (S. 67)	Methodik: Bezugnahme auf andere laufende Daten- erhebung(en) – Mehrfacheingaben in mehrere Studien vermeiden <i>„Prüfung, ob die wissenschaftlichen Fragestellungen teilweise oder vollständig in bereits geplanten oder laufenden Datenerhebungen zum jeweiligen Arzneimittel adressiert werden: Darauf basierend werden die wissenschaftlichen Fragestellungen abgeleitet, für die dies nicht der Fall ist und für die daher eine AbD durchgeführt werden sollte (AbD- Fragestellungen).“</i> <u>Anmerkung:</u> In den ersten AbD-Konzepten und Registerrecherchen wurden laufende Datenerhebungen aus interventionellen, einarmigen Extensionsstudien oder EMA-Nachzulassungsaufgaben bisher immer abgelehnt. Allerdings ist die Vernachlässigung solch wichtiger Informationen unverhältnismäßig. Es wurde bereits mehrfach die Problematik diskutiert, dass gerade bei seltenen Erkrankungen Patient:innen mehrfach in verschiedene nationale und internationale Studien und Register eingegeben werden, was zu einem erheblichen publication bias führen kann, wenn immer wieder zu den

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	gleichen Patient:innen Auswertungen durchgeführt werden (9,17,33). <u>Vorgeschlagene Änderung:</u> Auch für eine AbD könnte die Information zu zVT / Vergleichstherapien parallel in anderen Studien abgefragt werden und über indirekte Vergleiche mit externen Kontrollarmen ausgewertet werden, anstatt eine komplett neue prospektive Datenerhebung zu fordern. Manchmal sind sogar historisch erhobene Informationen zielführender, da sich gerade im AbD-Design mit einer prospektiven Registerstudie über Jahre hinweg die Therapielandschaft auch stark ändern kann. Eine viel stärkere Absprache mit der EMA und deren Guidelines für Zulassungsstudien und Nachzulassungsaufgaben sollte im europäischen Raum zwischen Zulassungsbehörden und HTA-Behörden und EMA und EU-HTA erfolgen, um realistisch durchführbare AbD und Registerstudien umzusetzen. Es ist bei manchen Erkrankungen unrealistisch zu fordern, dass ein pU nur für Deutschland separat nochmals eine AbD Studie aufsetzt, wenn schon von EMA, FDA oder anderen HTA-Behörden Datenerhebungen gefordert wurden oder laufen, in der manchmal deutsche Zentren sogar beteiligt sind. Ohne dass auf die Realität und die Problematik von tatsächlichen Registern und laufenden interventionellen und non-interventionellen Studien eingegangen wird, ist nicht abzusehen, dass eine rein autarke AbD nur für Deutschland durchzuführen ist. Die Langzeiterhebung von Daten zu bestimmten ATMP-Produkten, z.B. Gentherapien, ist so wichtig, dass sie nicht an den idealtypischen Vorstellungen scheitern sollte, sondern pragmatisch auf die Realität eingegangen werden sollte. Zudem sollte der G-BA stärker und kritischer als bisher die Realisierbarkeit und Angemessenheit von Register- und Studienanforderungen in den Verfahren berücksichtigen. Die Erfolgsaussichten auf bessere Evidenz, die Umsetzbarkeit und die Verhältnismäßigkeit sollten im Rahmen eines Stellungnahmeverfahrens vor der Verfahrenseinleitung einer anwendungsbegleitenden Datenerhebung erörtert werden. Die methodischen Vorgaben müssen bezüglich ihrer Praxistauglichkeit einer kritischen Prüfung unterzogen werden, zum einen, um die speziellen Herausforderungen bei seltenen Erkrankungen zu würdigen, aber auch um zu gewährleisten, dass die Ableitung eines Zusatznutzens

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	ohne Randomisierung und unter Hinzunahme historischer Vergleiche möglich ist.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) (S. 67)	Methodik: Studieninterne Datenerhebung als Möglichkeit belassen, da es meist keine Ideal- typischen Register in der Realität gibt <i>„Prüfung, ob bereits Datenplattformen existieren, die prinzipiell als primäre Datenquelle für die Durchführung von Studien zur Beantwortung der AbD-Fragestellungen dienen können: Die im Rapid Report A19-43 des Instituts [393] dargestellte Analyse der Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V hat ergeben, dass neben der studienindividuellen Datenerhebung insbesondere (Indikations-)Register eine geeignete Datenquelle für Studien zur Beantwortung der AbD- Fragestellungen darstellen können (Registerstudien). Voraussetzung hierfür ist, dass das jeweilige Register die notwendigen Daten in ausreichender Qualität zur Verfügung stellen kann. Dies umfasst neben der Datenerhebung u. a. die Planung, Auswertung und Publikation der Ergebnisse der zugehörigen Register- studie [364]. Im AbD-Konzept werden daher potenziell geeignete Register identifiziert (siehe Abschnitt 8.2) und die jeweiligen Registerbetreiber um weiterführende Informationen zum Register anhand eines strukturierten Fragebogens gebeten. Anhand dieser und weiterer öffentlich verfügbarer Informationen werden die Register charakterisiert und ihre Eignung für die Durchführung einer Registerstudie zur Beantwortung der AbD- Fragestellungen beurteilt, auch auf Basis national und international verwendeter Qualitätskriterien [242,248,299,500,691]. Aus dieser Eignungsprüfung wird zudem der Änderungs- und Erweiterungsbedarf abgeleitet, der ggf. erforderlich ist, damit das jeweilige Register als primäre Datenquelle für die Registerstudie zum Zwecke der AbD dienen kann.“</i> <u>Anmerkung:</u> Es wird zwar von studieninternen Datenerhebungen gesprochen, aber meist doch auf reine (Ideal-)Register für eine AbD abgezielt. Register haben meist größere Mängel als studieninterne Datenerhebungen. Sie benötigen aufgrund von bisherigen Governance-

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Strukturen meist viel länger zur Umstellung von bestimmten CRF-items und anderen Vorkehrungen als eine studieninterne Datenerhebung. Die kategorische Unterscheidung von studieninterner Datenerhebung versus Registern ist nicht möglich in der Realität. Häufig ist es so, dass aufgrund der GCP (Good Clinical Practice), GMP (Good Manufacturing Practice) (2) und GEP (Good Epidemiological Practice) Anforderungen von Zulassungsbehörden nur eine studieninterne Datenerhebung durchgeführt werden kann, weil man mit einem existierenden Register niemals alle GCP (11,15) und GEP Anforderungen zu Auditierung (16), Query-Management, Audit-Trails, Qualitätssicherungsmaßnahmen, CFR Part 11 compliance (1) der elektronischen Systeme etc. sicherstellen könnte. Der Umbau eines Registers erfordert daher mehr Aufwand als eine Planung einer studieninternen Datenerhebung oder Verwendung von Daten aus laufenden EMA-beauftragten Extensionsstudien. Methodische Anforderungen an Auswertungen sind hoch und ggf. hohe Hürde für positive Ergebnisse <u>Vorgeschlagene Änderung:</u> Es sollte transparenter und klarer erklärt werden, wann studieninterne Datenerhebungen und wann Registerstudien zielführend sind. Die im Rapid Report erbrachte Darstellung zu obligaten und optionalen/ case-by-case zu entscheidenden Qualitätskriterien für Register sollten genauer festgelegt werden, damit nachvollziehbar ist, wann genau ein Register in seiner jetzigen Form akzeptiert wird und wann ein Umbau noch kurz- und mittelfristig realisierbar ist. Das IQWiG sollte hier pragmatischer auf realistische Anforderungen eingehen, anstatt in der Praxis nicht umsetzbare Idealregister und komplexe Auswertungen zu fordern. Denn so bleibt eigentlich derzeit fast nur die studienindividuelle Datenerhebung als die realistische Möglichkeit, um allen Zulassungs- und HTA-Richtlinien für valide Daten zu entsprechen.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) (S. 67)	Methodik: AbD mit vergleichender Evidenzgenerierung bei seltenen Erkrankungen nicht in jedem Fall umsetzbar <i>„Beschreibung der Art der AbD: Hierzu gehören das Studiendesign und die primäre Datenquelle für die</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>Datenerhebung. Aufgrund der prinzipiellen Frage- stellung der Nutzenbewertung (Vergleich gegenüber einer zweckmäßigen Vergleichstherapie) handelt es sich dabei in der Regel um eine vergleichende Studie, die unter der Einschränkung von [242] § 35a SGB V Abs. 3b ohne Randomisierung durchgeführt wird [6,393]. Die Auswahl der primären Datenquelle ergibt sich aus dem Ergebnis der vorhergehenden Prüfung der Eignung vorhandener Datenquellen.“</i> <u>Anmerkung:</u> Nicht bei allen Produkten (insbesondere bei Produkten mit Orphan Drug-Status, conditional approval oder exceptional circumstance) sind bereits Therapie- alternativen auf dem deutschen Markt verfügbar. Vergleichstherapien sind besonders bei so genannten Solisten nicht verfügbar; d.h. Therapien für Erkrankungen ohne derzeitige medikamentöse Therapiemöglichkeiten. Hohe Anforderungen an Studiendurchführung und Registerqualität erfordern eine teilweise unrealistische Weiterentwicklung des Registers – solche idealen Register existieren in der Realität nicht in Deutschland (s. IGES/vfa Analyse (3) und BMG-Gutachten für Register (8)). Teilweise können in dem Register keine „interventionellen Schritte“ im geforderten Maß eingebaut werden. Die geforderte Datenqualität ist schwer zu erreichen mit derzeit in Deutschland oder international existierenden Registern. Sie sind und werden auch in Zukunft für andere Zwecke aufgesetzt, und nicht nur für die Nutzenbewertung. Governance-Strukturen für Register müssten geändert werden, um in manchen Fällen überhaupt Daten- auswertungen für HTA durchführen zu können. Der pU wird vom G-BA für die AbD beauftragt, daher ist er demnach auch verantwortlich und das muss miteingeplant werden, damit register-basierte Studien für HTA-Zwecke vom pU einwandfrei möglich sind. Bei der Schätzung, ob innerhalb eines akzeptablen Zeitraums prinzipiell ausreichend Patient:innen rekrutiert und beobachtet werden können, sollte geprüft werden, ob Veränderungen (durch neue Zulassungen oder andere äußere Einflüsse wie z. B. Neugeborenen Screening) in den zu untersuchenden Populationen zu erwarten sind. Folgende Szenarien sind hier vorstellbar:

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	1. Ursprüngliche Population bleibt unverändert 2. Ursprüngliche Population nimmt ab und verschwindet nach und nach 3. Ursprüngliche Population verschwindet nach sehr kurzer Zeit Bei Solisten haben wir das Problem, wenn das Medikament zugelassen ist, eine vergleichbare Vergleichsgruppe in der Versorgung zu finden. Denn die Patient:innen, die das Medikament nicht bekommen, unterscheiden sich in der Regel von denen, die das neue Medikament bekommen. <u>Vorgeschlagene Änderung:</u> Es sollte genau geklärt werden, in welchem Fall keine vergleichende AbD-Studie möglich ist. Für die genannten Szenarien 2 und 3 sollten Kriterien zur Überprüfung der Realisierbarkeit einer AbD und daraus folgende Maßnahmen beschrieben werden. Gerade bei Erkrankungen, für die bisher keine Therapiealternativen existieren, ist es fraglich, ob genügend Patient:innen im Kontrollarm zur Verfügung stehen werden. Patient:innen mit seltenen Erkrankungen wird hier zum Teil das erste Mal eine Therapiealternative zur Verfügung gestellt.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) (S. 67 f.)	Methodik: Realistische Fallzahlschätzungen und Zeitpläne aufstellen <i>„Beschreibung der Dauer der AbD: Dazu gehört zum einen die Dauer der Beobachtung der einzelnen Patientinnen und Patienten, die sicherstellen soll, dass relevante Endpunkte in der vorliegenden Indikation und Anwendungssituation beurteilt werden können. Zum anderen ist die Dauer der AbD auch dadurch bestimmt, dass genügend Patientinnen und Patienten bzw. Ereignisse (die notwendige Fallzahl) eingeschlossen und beobachtet werden können, um aussagekräftige Daten für die AbD zu sammeln. Hierfür werden epidemiologische Daten wie Prävalenz und Inzidenz der Erkrankung sowie der Grad der Vollzähligkeit der Datenquelle betrachtet.“</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Anmerkung:</u> Dieser Punkt scheint bei den bisherigen Präzedenzfällen nicht abschließend gelöst worden zu sein. Einerseits wurden internationale Datenlösungen abgelehnt, andererseits eine Stichprobengröße für Deutschland geschätzt, die nicht realistisch für ein deutsches Register erscheint und die Prävalenz/Inzidenz (tatsächlich zur Verfügung stehende Gesamtpopulation in Deutschland.) gar nicht berücksichtigte, oder es wurde gar keine Einschätzung zur Stichprobengröße gegeben. Wenn gar keine Stichprobenschätzung durchgeführt wurde, bleibt unklar, wie der Umfang, Verhältnismäßigkeit und Realisierbarkeit der AbD geschätzt werden kann. Fallzahlschätzungen bezogen sich in manchen Fällen nur auf einen einzelnen Endpunkt; jedoch werden in der Nutzenbewertung ja alle patientenrelevanten Endpunkte aus den Kategorien Mortalität, Morbidität, Lebensqualität und Sicherheit miteinander abgewogen (6). Zudem hängt auch der Zeitraum von der Dynamik der Therapielandschaft ab. Der Rekrutierungszeitraum, d.h. die Zeit bis genügend Patient:innen überhaupt eine Therapie erhalten haben, muss hier auch einberechnet werden. Teilweise kann es sich damit um Zeiträume über mehr als 5 oder 10 Jahre handeln; bis dahin könnte ein Arzneimittel dann bereits den Unterlagenschutz verloren haben, sodass keine AMNOG-Nutzenbewertung mehr erfolgen kann. <u>Vorgeschlagen Änderung:</u> Realistische Fallzahlschätzungen sollten transparent im IQWiG-Konzept dargestellt werden und für mehr als einen patientenrelevanten Endpunkt und verschiedene Szenarien berechnet werden. Internationale Datenlösungen bei seltenen Erkrankungen mit geringen Fallzahlen sollten berücksichtigt werden. Rekrutierungs- und klinische Set-up-Phasen, Inzidenz und Prävalenz im konkreten Anwendungsgebiet, Patientenpräferenzen für Therapien etc. müssen bei der Zeitplanung und Stichprobenschätzung beachtet werden. Die Transparenz und Nachvollziehbarkeit mit Angaben und Referenzen hinsichtlich der Vorgehensweisen, z. B. bei der Fallzahlberechnung, sollte gegeben sein.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a	Methodik: Vershobenen Nullhypothesen mit weiterhin „dramatischen“ Effekten bei Seltene Erkrankungen nicht umsetzbar

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abs. 3b SGB V (AbD-Konzepte) (S. 68)	<i>„Beschreibung des Umfangs der AbD: Aus einer nicht randomisierten Studie kann aufgrund potenziell unbekannter Confounder aus den in der Studie beobachteten Effekten erst ab einer bestimmten Effektstärke eine Aussage zum Nutzen oder Schaden einer Intervention abgeleitet werden (siehe Abschnitte 3.1.4 und 3.2.2), z. B. in einem Bereich von 2 bis 5 für das relative Risiko (bzw. 0,2 bis 0,5 bei mortalitätssenkenden Interventionen) bezogen auf die untere bzw. obere Konfidenzintervallgrenze [393]. Dies wird bei der Frage, ob die geplante AbD prinzipiell sinnvoll durchführbar ist, berücksichtigt. Basierend auf den vorliegenden Erkenntnissen zum Arzneimittel und zur zweckmäßigen Vergleichstherapie wird abgeschätzt, ob innerhalb eines akzeptablen Zeitraums prinzipiell ausreichend Patientinnen und Patienten rekrutiert und beobachtet werden können, um mit einer AbD grundsätzlich aussagekräftige Ergebnisse für eine Nutzenbewertung generieren zu können.“</i> <u>Anmerkung:</u> Die Forderung, dass die untere bzw. obere Konfidenzintervallgrenzen über einem Bereich von 2 bis 5 für das relative Risiko (bzw. unter 0,2 bis 0,5 bei mortalitätssenkenden Interventionen) liegen müssen, ist zu extrem – insbesondere, da es sich bei einer AbD (gemäß GSAV-Gesetz) meist um eine seltene Erkrankung oder Last-Line-Therapien mit wenigen Patient:innen handelt. Dass für mortalitätssenkende Interventionen die gleichen Grenzwerte (nur reziprok) wie für alle anderen Endpunkte vorausgesetzt werden, steht im Kontrast zur Vorgehensweise aus der Zielgrößenkategorie-Tabelle 4 und 5 bei RCT im Methodenpapier für die reguläre Nutzenbewertung. Eine Begründung oder Referenz für die Wahl dieser Grenzwerte wird nicht genannt. Zudem fehlt die Angabe eines Bereiches für stetige Endpunkte. Schließlich bleibt die Frage, ob die AbD das richtige Modell ist, um darauf basierend einen Zusatznutzen festzustellen, der wiederum Grundlage für die Preisfindung ist. Gerade bei den kleinen Fallzahlen mit Vollerhebung in einer AbD könnten Pay-for-Performance Modelle sinnvoller und angemessener als eine AbD sein.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Vorgeschlagene Änderung:</u> Analog zu den Zielgrößenbestimmungen für RCT (Tabelle 4 und 5 des IQWiG Methodenpapiers) sollten hier auch für unterschiedliche Endpunktkategorien und Ausprägungen andere Zielgrößen, bspw. in Form von RR, bestimmt werden. Die Wahl der konkreten Grenzwerte zur Bestimmung des Ausmaßes des Zusatznutzens muss außerdem begründet und mit Referenzen belegt werden. Außerdem sollten auch Grenzwerte für stetige Zielgrößen definiert werden. Es sollte zudem genau spezifiziert werden, ab welchen Kriterien bei welchen Endpunkten das IQWiG und G-BA zur Erkenntnis kommt, dass eine AbD nicht in einem akzeptablen Zeitraum durchgeführt werden kann. Der größeren Ergebnisunsicherheit bei seltenen Erkrankungen, die bei der Zulassung bereits berücksichtigt wird, sollte auch bei der HTA-Bewertung Rechnung getragen werden. Die Forderung von dramatischen Effekten, die die Akzeptanzschwelle nochmals erhöht, steht dem entgegen. Die verschobene Nullhypothese sollte als Forderung aufgegeben werden. Es ist eine zu extreme Forderung, besonders bei seltenen Erkrankungen (z. B. ist es kaum möglich, dass bei mortalitätssenkenden Interventionen das 95%-Konfidenzintervall vollständig unter 0,5 bzw. 0,2 liegt), insbesondere da für die Berücksichtigung der Ergebnisse bereits vorausgesetzt wird, dass für relevante Confounder angemessen adjustiert wurde. Bei manchen seltenen Erkrankungen sollte alternativ eher ein Pay-for-Performance Modell statt einer AbD in Betracht gezogen werden; allerdings sollte vermieden werden, dass AbD als auch P4P Modelle beide gleichzeitig durchgeführt werden müssen. Eine fallspezifische Entscheidung für die eine ODER die andere Alternative ist erforderlich, um Doppelstrukturen zu vermeiden.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) (S. 68)	Methodik: Unklarheit über verbindliche und optionale methodische Vorgehensweisen im Rapid Report vs. Methodenpapier, z.B. bei Confounder- identifizierung etc. <i>„Beschreibung grundsätzlicher Anforderungen an die Auswertung der Daten: Die Basis hierfür bilden die im Rapid Report A19-43 beschriebenen Konzepte zur</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>Auswertung versorgungsnaher Daten zum Zwecke der Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V [393]. Aufgrund des notwendigen nicht randomisierten Designs der Studie für die AbD stellt dabei die adäquate Adjustierung für Störgrößen (Confounder) einen wesentlichen Bestandteil dieser Anforderungen dar (siehe hierzu Abschnitt 9.3.6).</i> <i>Das AbD-Konzept enthält abschließend ein Fazit zu folgenden Aspekten:</i> ▪ <i>die Fragestellung(en), die Gegenstand der Datenerhebung und Auswertungen sein soll(en),</i> ▪ <i>die Anforderungen an Art, Dauer und Umfang der Datenerhebung,</i> ▪ <i>die Methodik der Datenerhebung,</i> ▪ <i>die Methodik der Auswertung durch den pharmazeutischen Unternehmer sowie</i> ▪ <i>eine Einschätzung dazu, ob die geplante AbD prinzipiell realisierbar ist.</i> <i>Das Fazit stellt die Empfehlung an den G-BA für das weitere Verfahren dar, in dem der G-BA abschließend über die Forderung einer AbD entscheidet [284]. Sofern der G-BA eine AbD fordert, muss der pU ein Studienprotokoll und einen statistischen Analyseplan (SAP) erstellen. In diesen Dokumenten ist die Planung, Durchführung und Auswertung der Studie zum Zwecke der AbD zu beschreiben. Das IQWiG kann sowohl bei der Beratung im Vorfeld der Erstellung des Studienprotokolls und SAPs als auch bei der Eignungsprüfung dieser Dokumente im Auftrag des G- BA eingebunden werden.“</i> <u>Anmerkung:</u> Es wird zwar auf den Rapid Report (20) verwiesen, aber eine ausführliche Beschreibung welches statistisch- methodische Verfahren bei welcher Anzahl von Patient:innen oder Studienteilnehmern zu wählen ist, wird auch im Rapid Report nicht genau spezifiziert. Confounderadjustierung mit Propensity-Score- Methoden könnten für viele seltene Erkrankungen nicht adäquat sein, da eine Mindestanzahl von ca. 10 Patient:innen pro Confounder verfügbar sein sollte. Dies

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	wird auch vom IQWiG in mehrere AbD-Verfahren so dargestellt (5). Die Confounderidentifizierung wird nur mit wenigen Referenzen beschrieben (29), die vieles unklar lassen. Zudem existieren zwar Methoden, um die Überlappung zu untersuchen, allerdings gibt es noch kein etabliertes Kriterium, um zu bewerten, ob eine ausreichende Überlappung vorliegt. Es wird in diesem Abschnitt nicht auf den Fall einer AbD mit mehreren Produkten aus dem neuen Rapid Report von Mai 2022 (22) eingegangen. <u>Vorgeschlagene Änderung:</u> Das IQWiG sollte in den Allgemeinen Methoden auf einschlägige Literatur verweisen, welche Methodik zur Adjustierung in welchem Fall verwendet werden soll. Matching-adjusted oder naive indirekte Vergleiche sollten in bestimmten Fällen ebenfalls akzeptiert werden. Es sollte eindeutig sein, welche Passagen aus dem Rapid Report zu versorgungsnahen Daten tatsächlich verbindlich für die AbD oder Indirekte Vergleiche mit versorgungsnahen Daten sind und diese verbindlichen Vorgaben sollten im Methodenpapier konkretisiert werden. Da AbD vor allem bei seltenen Erkrankungen mit kleinen Fallzahlen relevant sind (im Median beträgt bei der Gruppe von Medikamenten, die in den Anwendungskreis einer AbD fallen würden, die GKV-Zielpopulation nur ca. 100 Patient:innen p.a. (28)), wäre es wünschenswert, auf die methodischen Probleme, die mit kleinen Fallzahlen einhergehen (wie auch durch EMA in dem Dokument „clinical trials with small populations“ (13) angeregt), einzugehen (z. B. sind einige genannte Methoden wie Propensity Score Matching nur mit ausreichend großen Fallzahlen möglich) und konkrete Methoden dafür anzubieten. Zudem sollte spezifiziert werden, dass das IQWiG immer beim Review des Studienprotokolls und SAPs einbezogen werden sollte. Denn gemäß BMG-Entscheid zur AbD (19) darf die AbD nicht im Nachhinein von G-BA oder IQWiG abgelehnt werden und daher sind eindeutige transparente Angaben, die realistisch umsetzbar sind, von IQWiG und G-BA im Vorfeld der AbD zu treffen. Ein Kriterium zur Bewertung der Überlappung sollte definiert werden.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Die Vorgehensweise bei einer AbD mit mehreren Produkten (22) innerhalb eines Registers oder einer studieninternen Datenerhebung sollte auch hier im Methodenpapier genau beschrieben werden.
9.3.6 Adjustierung und multifaktorielle Verfahren (S. 176 ff.)	Methodik: Vorgehen bei kleinen Fallzahlen <i>„Da eine AbD in der Regel eine vergleichende Studie ohne Randomisierung darstellt, spielt auch bei AbDs eine adäquate Adjustierung für Confounder eine zentrale Rolle (siehe Abschnitt 3.3.4). Hier ist es erforderlich, vorab alle relevanten Confounder (inklusive wichtiger Wechselwirkungen) mithilfe eines systematischen Ansatzes zu identifizieren [587], vollständig zu erheben und bei der Datenanalyse in angemessener Form im Modell zu berücksichtigen [393].</i> <i>Für eine Datenanalyse mit adäquater Adjustierung für Confounder gibt es verschiedene kausale Methoden (z. B. multifaktorielle Regressionsmodelle, Instrumentenvariablen, Propensity Scores), die in einem detaillierten SAP zu beschreiben sind [12,163,698,757]. Hierbei ist es insbesondere erforderlich</i> ▪ <i>das kausale Modell exakt zu beschreiben, z. B. mithilfe kausaler Grafiken,</i> ▪ <i>die Annahmen des kausalen Modells darzustellen,</i> ▪ <i>fundiert zu begründen, warum diese Annahmen im konkreten Anwendungsfall zu rechtfertigen sind,</i> ▪ <i>darzustellen, welche Verfahren und Kriterien bei der Modellanpassung verwendet werden sowie</i> ▪ <i>zu erläutern, mit welchen Sensitivitätsanalysen die Robustheit der Ergebnisse überprüft werden.</i> <i>Im Rahmen einer AbD spielt zur Adjustierung für Confounder insbesondere die Propensity-Score-Methode eine Rolle [250,393], u. a. weil es hier Ansätze gibt die zentralen Annahmen der kausalen Inferenz, nämlich ausreichende Positivität, Überlappung und Balanciertheit zu prüfen und darzustellen [34]. Diese Annahmen sind aber auch für andere kausale Methoden erforderlich. Falls eine dieser Annahmen mit einem Propensity-Score-Ansatz nicht erreichbar ist, so ist es keine sinnvolle Option z. B. stattdessen ein</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>multifaktorielles Regressionsmodell zu verwenden [34,773].</i> <i>Unter Positivität versteht man, dass die im Rahmen der AbD zu vergleichenden Interventionen für alle Patientinnen und Patienten, die in die Auswertung eingehen, eine Therapieoption darstellen. Dies ist im Rahmen der Einschlusskriterien für die AbD zu prüfen und zu beschreiben. In der zur Verfügung stehenden Datenbasis muss eine ausreichende Überlappung zwischen den beiden zu vergleichenden Gruppen bezüglich der Verteilung des Propensity Scores gegeben sein. Weiterhin ist es erforderlich, dass die beiden zu vergleichenden Gruppen bezüglich aller relevanten Confounder ausreichend balanciert sind.</i> <i>Um im Rahmen von Propensity-Score-Ansätzen eine ausreichende Überlappung und Balanciertheit zu erreichen, existieren zahlreiche Methoden [34,162,180,464,758]. Durch Ausschluss von Patientinnen und Patienten in sich nicht überlappenden Bereichen des Propensity Scores kann die Überlappung und Balanciertheit verbessert werden, aber es ändert sich auch unter Umständen die Zielpopulation, auf die sich letztlich die Effektschätzung bezieht. Daher ist diese Population genau zu beschreiben und es ist zu untersuchen, ob diese der ursprünglichen Fragestellung entspricht. Ist dies nicht der Fall lässt sich unter Umständen die ursprüngliche Fragestellung mit den zur Verfügung stehenden Daten nicht adäquat beantworten [180,393].“</i> <u>Anmerkung:</u> Es werden keine Lösungen beschrieben, wie mit sehr kleinen Fallzahlen oder sehr ungleichen Verteilungen in den Studienarmen, fehlenden Werten oder in der klinischen Praxis nicht erhebbaren Endpunkten bei allen Patient:innen umgegangen werden soll. Multifaktorielle Regressionsanalysen und Propensity-Score Methoden sind nur bei ausreichend großen Fallzahlen und nur unter bestimmten Voraussetzungen möglich (5). Für Confounderidentifizierung und -adjustierung fehlen pragmatische Lösungsansätze.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Vorgeschlagene Änderung:</u> Matching-adjusted indirect comparisons (MAIC) oder naive indirekte Vergleiche sollten bei AbD mit sehr kleinen Fallzahlen pro Studienarm akzeptiert werden. Das IQWiG könnte z. B. einen Ausprägungsgrad des Zusatznutzens vergeben und mit einer weiteren Bewertungskategorie die Zuverlässigkeit und Aussagewahrscheinlichkeit der Methodik graduell einstufen, anstatt den Zusatznutzen komplett abzuspochen (s. ICER Institut in USA (31)). Eine Klärung oder Grenzwerte wären wünschenswert, was mit „angemessener Form“ gemeint ist. Es sollte auch beschrieben werden, wie mit fehlenden Werten bei einer AbD umgegangen werden soll. Wenn der pU z. B. Confounder in der Auswertung nicht berücksichtigt, weil mehr als 30% der Daten fehlen, wird kritisiert, dass dies die Relevanz eines Confounders nicht beeinträchtigen würde. Interessant wäre es zu erfahren, wie man in diesem Fall damit umgehen sollte. Sind z.B. Imputationen gewünscht oder würde eine Sensitivitätsanalyse mit Größe und Richtung des Bias helfen? Man könnte in einer Sensitivitätsanalyse die Größe und Richtung des Bias beschreiben und nicht pauschal die komplette Analyse ablehnen. Bei der Balanciertheit-Überprüfung sollte auch eine Guidance in Anlehnung an NICE (23) und EUnetHTA (14) erfolgen. Diese beschreiben z. B., dass die standardized mean difference < 0.1 sein sollte.
3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte) (S. 68 ff.)	Methodik: Umgang mit Zwischenanalysen und Vergeblichkeitskriterien unklar <u>Anmerkung:</u> Es wird nicht beschrieben, wie die andauernde Evaluierung der Daten im Rahmen von Zwischenanalysen erfolgen soll (das GSAV-Gesetz spricht von mind. einer Zwischenauswertung nach 18 Monaten (7)). Eine kontinuierliche Evaluierung der Daten in Zwischenanalysen erhöht den Aufwand und es ist fraglich, wie diese dann in der finalen Nutzenbewertung nach AbD-Auswertung herangezogen werden und in der nachfolgenden Preisverhandlung berücksichtigt werden. Wie in der Nutzenbewertung und Preisverhandlung mit einer abgebrochenen AbD aufgrund von Vergeblichkeit

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	umgegangen wird, bleibt unklar und es gibt diesbezüglich keine Planungssicherheit. <u>Vorgeschlagene Änderung:</u> „Futility“- bzw. Vergeblichkeits-Kriterien müssen eindeutig vom IQWiG bereits im Konzept vorgeschlagen werden und eindeutig beschrieben werden, ab welchem Zeitpunkt (>18 Monate) überhaupt über den Abbruch gemäß der a-priori im Studienprotokoll festgelegten Vergeblichkeitskriterien entschieden werden darf.
9.3.6 Adjustierung und multifaktorielle Verfahren (S. 177)	Methodik: Instrumentenvariable <i>„Für eine Datenanalyse mit adäquater Adjustierung für Confounder gibt es verschiedene kausale Methoden (z. B. multifaktorielle Regressionsmodelle, Instrumenten- variablen, Propensity Scores), die in einem detaillierten SAP zu beschreiben sind [12,163,698,757].“</i> <u>Anmerkung:</u> Pfizer begrüßt, dass die Instrumentenvariablen als Möglichkeit zur Confounderkontrolle erwähnt werden. Sie ist eine innovative Methode gerade bei Beobachtungsdaten (24) trotzdem eine Kausalattribution ähnlich eines randomisierten Studiendesigns mit geringeren Fallzahlen zu ermöglichen. <u>Vorgeschlagene Änderung:</u> Jedoch sollte diese Vorgehensweise im Methodenpapier mehr ausgeführt werden und als realistische Option angeboten werden.
9.3.8 Indirekte Vergleiche (S. 183 ff.)	Methodik: Indirekte Vergleiche mit versorgungs- nahen Daten nicht kategorisch ablehnen, wenn RCT nicht möglich sind <i>„Allerdings müssen adäquate Methoden für indirekte Vergleiche verwendet werden. Die Anwendung nicht adjustierter indirekter Vergleiche (d. h. die Verwendung einzelner Arme aus verschiedenen Studien) wird abgelehnt [54].60]. Dies gilt auch für Methoden für indirekte Vergleiche, in denen über Modellierungen mit starken Annahmen über die unbekannten Effekte [122[133] oder mithilfe von Verfahren aus dem Bereich der Kausalmodelle für Beobachtungsstudien mit</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>untestbaren Annahmen [656669] versucht wird, Effektschätzungen trotz fehlender Brückenkomparatoren zu ermöglichen. Es werden ausschließlich adjustierte indirekte Vergleiche über adäquate Brückenkomparatoren akzeptiert. Hierzu zählen insbesondere das Verfahren von Bucher et al. [101][112] sowie die oben erwähnten Methoden der Netzwerk-Metaanalysen“</i> <u>Anmerkung:</u> Im Kapitel 9.3.8. Indirekte Vergleiche des IQWiG Methodenpapiers 7.0 genauso wie im IQWiG Rapid Report zu versorgungsnahen Daten (20) wird die Methodik sowohl von indirekten Vergleichen mit Brückenkomparator (z. B. nach Bucher oder Netzwerk-Metaanalysen) als auch von naiven und Confounder-adjustierten indirekten Vergleichen (ITC) mit externen Kontrollarmen (ECA) beschrieben. Allerdings schreibt das IQWiG auch im Entwurf für Methodenpapier 7.0, dass nicht adjustierte indirekte Vergleiche und adjustierte indirekte Vergleiche (ITC) ohne Brückenkomparator kategorisch abgelehnt werden. <u>Vorgeschlagene Änderung:</u> Wenn Counfounder-adjustierte indirekte Vergleiche mit externen Kontrollarmen (ECA) mit versorgungsnahen Daten sowohl im Rapid Report als auch an anderer Stelle im IQWiG Methodenpapier erklärt werden und in der Nutzenbewertung auch bereits häufiger bewertet worden sind, kann es nicht bei diesem kategorischen Ausschluss von ITC mit ECA in bleiben, sondern es müssten Fälle im Methodenpapier beschrieben werden, in denen diese Vergleiche eine Möglichkeit darstellen (z. B. in Situationen, in denen RCT in der Praxis - gerade bei seltenen Erkrankungen - nicht möglich sind, und in denen Zulassungsbehörden diese Methode sowohl in Guidelines als auch in scientific advice meetings empfehlen). Das IQWiG sollte Bewertungskategorien zu Ergebnissicherheit der Evidenz in Abstufungen einführen, anstatt kategorisch die verfügbare Information abzulehnen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. Administration USDoHaHSFaD. Guidance for Industry. Part 11, Electronic Records; Electronic Signatures — Scope and Application. 2003.
2. Agency EM. Good manufacturing practice 2022 [Available from: <https://www.ema.europa.eu/en/human-regulatory/research-development/compliance/good-manufacturing-practice>].
3. Berkemeier F, Gothe H, Kurepkat M. Können anwendungsbegleitende Datenerhebungen überhaupt realisiert werden? Monitor Versorgungsforschung. 2020;2.
4. Bundesausschuss G. Antworten auf häufig gestellte Fragen zum Verfahren der Nutzenbewertung [Available from: <https://www.g-ba.de/themen/arzneimittel/arzneimittel-richtlinie-anlagen/nutzenbewertung-35a/faqs/>].
5. Bundesausschuss G. Arzneimittel-Richtlinie/Anlage XII: Autologe Anti-CD19-transduzierte CD3+ Zellen (rezidiertes oder refraktäres Mantelzell-Lymphom) – Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen 2022 [updated 01.11.2022. Available from: <https://www.g-ba.de/beschluesse/5540/>].
6. Bundesausschuss G. Beschluss. Arzneimittel-Richtlinie/Anlage XII: Valoctocogen Roxaparvovec (schwere Hämophilie A) – Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen 2023 [Available from: <https://www.g-ba.de/beschluesse/5866/>].
7. Bundesgesundheitsministerium. Gesetz für mehr Sicherheit in der Arzneimittelversorgung 2019 17.11.2021. [Available from: [https://urldefense.com/v3/__https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/G/GSAV_bgb119_S.1202_150819.pdf__;!!H9nueQsQ!_yLdnfYfltOxBmFdNbzOdabrUKEvk8Bdx2hypLK1nULVQwatX0Te8jh-M5G3LXSrwp9sTQOQz5VWQI7fCPwoKrwSor1TJHgT9VfsWA\\$](https://urldefense.com/v3/__https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/G/GSAV_bgb119_S.1202_150819.pdf__;!!H9nueQsQ!_yLdnfYfltOxBmFdNbzOdabrUKEvk8Bdx2hypLK1nULVQwatX0Te8jh-M5G3LXSrwp9sTQOQz5VWQI7fCPwoKrwSor1TJHgT9VfsWA$)].
8. Bundesministerium für Gesundheit (BMG). Gutachten zur Weiterentwicklung medizinischer Register zur Verbesserung der Dateneinspeisung und -anschlussfähigkeit 2021 [Available from: <https://www.bundesgesundheitsministerium.de/service/publikationen/details/gutachten-zur-weiterentwicklung-medizinischer-register-zur-verbesserung-der-dateneinspeisung-und-anschlussfaehigkeit-2.html>].
9. Calvez T, Chambost H, Claeysens-Donadel S, d'Oiron R, Goulet V, Guillet B, et al. Recombinant factor VIII products and inhibitor development in previously untreated

- boys with severe hemophilia A. *Blood*. 2014;124(23):3398-408.
10. Commission E. Clinical trials - Regulation EU No 536/2014 2022 [Available from: https://health.ec.europa.eu/medicinal-products/clinical-trials/clinical-trials-regulation-eu-no-5362014_en].
 11. Commission E. Clinical trials - Directive 2001/20/EC 2022 [Available from: https://health.ec.europa.eu/medicinal-products/clinical-trials/clinical-trials-directive-200120ec_en].
 12. Duru G, Auray JP, Béresniak A, Lamure M, Paine A, Nicoloyannis N. Limitations of the methods used for calculating quality-adjusted life-year values. *Pharmacoeconomics*. 2002;20(7):463-73.
 13. EMA European Medicines Agency. Guideline on clinical trials in small populations. [www]. WWW: EMA European Medicince Agency; 2006 [updated 27.07.2006. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-trials-small-populations_en.pdf].
 14. Eunethhta. D4.3.2 Direct and indirect comparisons. 2022.
 15. European Medicines Agency. ICH E6 (R2) Good clinical practice (GCP). 2021 [Available from: <https://www.ema.europa.eu/en/ich-e6-r2-good-clinical-practice>].
 16. European Medicines Agency. Guideline on registry-based studies 2021 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies_en-0.pdf].
 17. Fischer K, Lewandowski D, Marijke van den Berg H, Janssen MP. Validity of assessing inhibitor development in haemophilia PUPs using registry data: the EUHASS project. *Haemophilia*. 2012;18(3):e241-6.
 18. Gemeinsamer Bundesausschuss (G-BA). Arzneimittel-Richtlinie/Anlage XII: Autologe Anti-CD19-transduzierte CD3+ Zellen (rezidiertes oder refraktäres Mantelzell-Lymphom) – Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen. [updated 01.11.2022. Available from: <https://www.g-ba.de/beschluesse/5540/>].
 19. Gesundheit Bf. Beschluss des Gemeinsamen Bundesausschusses gemäß § 91 SGB V vom 16. Juli 2020. Änderung der Verfahrensordnung (VerfO): Änderung im 5. Kapitel – Verfahren zur Forderung einer anwendungsbegleitenden Datenerhebung nach § 35a Absatz 3b SGB V. 2020.
 20. Gesundheitswesen IfQuWi. [A19-43] Wissenschaftliche Ausarbeitung von Konzepten zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V - Rapid Report 2020. 2020.
 21. Gesundheitswesen IfQuWi. Überprüfung der Altersgrenzen im Mammografie-Screening Programm. 2022.
 22. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). [A21-37] Wissenschaftliche Ausarbeitung eines Konzeptes zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung nach § 35a SGB V in der Situation des Marktzugangs mehrerer Arzneimittel einer Wirkstoffklasse. [updated 22.10.2022. Available from: https://www.iqwig.de/download/a21-37_abd-bei-marktzugang-mehrerer-arzneimittel-einer-wirkstoffklasse_rapid-report_v1-0.pdf].
 23. Kent S., Kincaid L., Sharma M., Rowark S., Duffield S., Ayyar Gupta V., et al. NICE real-world evidence framework. 2022.

24. KVTRSAo S. Answering causal questions using observational data. The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel. 2021.
25. Latimer NR, Abrams KR. NICE Decision Support Unit Technical Support Documents. NICE DSU Technical Support Document 16: Adjusting Survival Time Estimates in the Presence of Treatment Switching. London: National Institute for Health and Care Excellence (NICE) Copyright © 2014 National Institute for Health and Clinical Excellence, unless otherwise stated. All rights reserved.; 2014.
26. Leuchs AK, Brandt A, Zinserling J, Benda N. Disentangling estimands and the intention-to-treat principle. Pharm Stat. 2017;16(1):12-9.
27. Medizinprodukte BfAu. Klinische Prüfung. Wirksamkeit nachweisen. Verträglichkeit feststellen. 2022 [Available from: https://www.bfarm.de/DE/Arzneimittel/Klinische-Pruefung/_node.html].
28. Nutzenbewertung IPz. Weiterentwicklung des AMNOG mit Augenmaß und Evidenz. 2022;15.
29. Pufulete M, Mahadevan K, Johnson TW, Pithara C, Redwood S, Benedetto U, et al. Confounders and co-interventions identified in non-randomized studies of interventions. Journal of Clinical Epidemiology. 2022;148:115-23.
30. Reed B. Breast cancer drug rejected for NHS use on cost-benefit grounds The Guardian2016 [Available from: <https://www.theguardian.com/society/2016/dec/29/breast-cancer-drug-kadcyla-rejected-for-nhs-use-on-cost-benefit-grounds>].
31. Review IfCaE. Evidence Rating Matrix [Available from: <https://icer.org/evidence-rating-matrix/>].
32. Seefat MR, Cucchi DGJ, Dirven S, Groen K, Zweegman S, Blommestein HM. A Systematic Review of Cost-Effectiveness Analyses of Novel Agents in the Treatment of Multiple Myeloma. Cancers (Basel). 2021;13(22).
33. Volkens P, Hanschmann KM, Calvez T, Chambost H, Collins PW, Demiguel V, et al. Recombinant factor VIII products and inhibitor development in previously untreated patients with severe haemophilia A: Combined analysis of three studies. Haemophilia. 2019;25(3):398-407.

A.1.34 Roche Pharma AG

Autorinnen und Autoren

- Knoerzer, Dietrich
- Fecker, Julian

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Dr. Dietrich Knoerzer, Head of Biostatistics and Epidemiology
Julian Fecker, Biostatistician
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation:
Roche Pharma AG
Market Access


☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Mit dem vorgelegten Entwurf für Version 7.0 insbesondere Kapitel 4 folgt das IQWiG seinem gesetzlichen Auftrag gemäß §35b (1) SGB V, transparente Methoden zur Kosten-Nutzen-Bewertung (KNB) von Arzneimitteln festzulegen.

Gesetzlich verankert ist die KNB schon seit 2007, eine praktische Relevanz hat sie bislang nicht erlangt. Der Gesetzgeber hat die KNB im Rahmen des AMNOG 2010 in der Regel als letztes Mittel implementiert, für den Fall dass im regulären Prozess von Nutzenbewertung nach § 35a und Preisfindung nach § 130b keine angemessene Lösung für ein neues Arzneimittel im Sinne der Wirtschaftlichkeit gefunden werden kann.

Die KNB stellt somit ein letztes Mittel bei der Ersatzpreisfindung dar, das im Rahmen des AMNOG- Verfahrens seit 2010 jedoch noch nie zum Einsatz kam. Zwar ist gesetzlich eine KNB grundsätzlich denkbar. Die Erfahrungen in Deutschland zeigen aber, dass die gegenwärtig in Deutschland angewendete und hier entwickelte Methodik des AMNOG-Verfahrens der KNB hinsichtlich des Marktzugangs von neuen, innovativen Therapeutika im internationalen Vergleich überlegen ist.

Die Änderung des Abschnitts 4 des Methodenpapiers ist damit eher theoretischer als praktischer Natur. Im Rahmen des gesetzlich vorgesehenen Zwecks der KNB verändert das IQWiG die Methodik hin zu den internationalen wissenschaftlichen Standards der Gesundheitsökonomie. Jedoch wurden im Methodenentwurf des IQWiG diese methodischen Standards fraglich interpretiert und umgesetzt. Auf diese will Roche im folgenden mit den Anmerkungen zu Kapitel 4 Bezug nehmen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 4.3.2 Komparatoren (S. 87-8)	<u>Jahrestherapiekosten</u> “Die Jahrestherapiekosten sind im Dossier zur Nutzenbewertung gemäß § 35a SGB V dagegen für alle vom G-BA als zVT bestimmten Therapien anzugeben (5. Kapitel 1. Abschnitt § 6 der Verfahrensordnung des G-BA) [284].” <u>Anmerkung:</u> Es ist unklar, wie Jahrestherapiekosten zu berechnen sind. Darüber hinaus machen diese nicht bei allen Indikationen Sinn und ergeben bei einer unsachgemäßen Berechnung ein verzerrtes Bild. Beispiele: - Dosierung in Abhängigkeit von Patientencharakteristika wie Gewicht oder Körperoberfläche - wie ist dabei vorzugehen (z.B. Verwendung von Statistiken der Patientenpopulation (Mittelwert, Median, ..) oder Berücksichtigung der kompletten Verteilung) - Behandlung von drop-outs und unterschiedlichen Therapiedauern der verschiedenen Therapien - Behandlung mit einer fixen bestimmten Anzahl von Zyklen <u>Vorgeschlagene Änderung:</u> Die Darstellung der Therapiekosten im Rahmen der KNB (siehe Kap. 4.7) unterscheidet sich grundsätzlich von der geforderten Darstellung der Jahrestherapiekosten im Dossier gemäß § 35a SGB V.
Abschnitt 4.3.2 Komparatoren (S. 87-8)	<u>Anmerkung:</u> Es werden ergänzende Sensitivitätsanalysen beschrieben, bei denen nur die Therapiekosten variiert werden, nicht aber die therapeutischen Effekte. Solche Analysen sollten nur dann erfolgen, wenn die therapeutische Äquivalenz stringent nachgewiesen ist (was sowohl Effektivität als auch Sicherheit/Nebenwirkungen einschließt). Ansonsten sind solche ergänzenden Analysen irreführend. <u>Vorgeschlagene Änderung:</u> Folgender Text sollte dem ersten Paragraphen auf Seite 88 angefügt werden: <i>“Zudem muss die Äquivalenz der verschiedenen Therapien, insbesondere bezüglich Wirksamkeit und Sicherheit, in solchen Fällen stringent nachgewiesen sein.”</i>
Abschnitt 4.3.3 Zielgrößen	Der erste Paragraph auf S. 89 besagt:

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
(S. 89)	<i>“Sofern mehrere Endpunkte als gesundheitsökonomische Zielgröße der KNB beauftragt werden, kann der Entscheidungsträger selbst im Rahmen der Beschlussfassung eine Gewichtung der Ergebnisse für die verschiedenen Endpunkte vornehmen. Eine andere Möglichkeit ist eine Gewichtung der Ergebnisse unter Berücksichtigung von Verfahren der Präferenzhebung von Patientinnen und Patienten.”</i> <u>Anmerkung:</u> Eine solche Gewichtung verschiedener Endpunkte ist wissenschaftlich keineswegs etabliert. Damit ist die praktische Umsetzung dieses Paragraphen unklar. Auch ist nicht ersichtlich, wie die kohärente Bewertung verschiedener Produkte langfristig sichergestellt werden kann, da keine Kriterien und Prinzipien für eine solche Gewichtung gegeben werden. Das Gleiche gilt für die Methode der Präferenzhebungen. <u>Vorgeschlagene Änderung:</u> Im Falle mehrerer, relevanter Endpunkte soll nur eine sehr beschränkte Anzahl Endpunkte als gesundheitsökonomische Zielgröße der KNB beauftragt werden, die sich auch sinnvoll in einer modellbasierten Evaluation abbilden lassen. In Fällen mit mehreren Endpunkten soll keine Gewichtung der Endpunkte stattfinden.
Abschnitt 4.3.4 Perspektive (S. 89)	<u>Anmerkung:</u> Es wird erwähnt, dass die gesellschaftliche Perspektive ergänzend Berücksichtigung finden kann. Für manche Therapiegebiete kann erwartet werden, dass ein beträchtlicher Teil der Effekte von einer GKV-Perspektive nicht erfasst werden. <u>Vorgeschlagene Änderung:</u> Die gesellschaftliche Perspektive sollte immer berücksichtigt werden. Der Referenzfall sollte grundsätzlich immer beide Perspektiven beinhalten. Damit soll der Basisfall jeweils zweimal durchgeführt werden, einmal mit der Perspektive der GKV-Versichertengemeinschaft und einmal mit der gesellschaftlichen Perspektive. Die Tabelle 6 “Übersicht über die Elemente eines Referenzfalls” ist entsprechend anzupassen.
Abschnitt 4.3.4 Perspektive	<u>Anmerkung:</u> Es werden verschiedene Perspektiven angesprochen:

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
(S. 89)	- GKV-Versichertengemeinschaft - (reine) GKV-Perspektive - Sozialversicherungsperspektive - Perspektive einzelner Sozialversicherungsträger - gesellschaftliche Perspektive - erweiterten Perspektive <u>Vorgeschlagene Änderung:</u> Es wäre hilfreich, diese Perspektiven inklusive der entsprechenden Variationen in den Referenzfällen zu definieren.
Abschnitt 4.3.5 Zeithorizont (S. 90)	<u>Anmerkung:</u> Der letzte Satz zu diesem Abschnitt möchte folgende arbiträre Festlegung machen: "In Szenarioanalysen wird regelhaft ein Zeithorizont von 5 Jahren betrachtet." Die Festlegung eines fixen Zeithorizonts auf 5 Jahre scheint aber wenig sinnvoll, sondern sollte vielmehr einer fall- bzw. indikationsbezogenen sinnvollen Einschätzung folgen . <u>Vorgeschlagene Änderung:</u> Szenarioanalysen mit verschiedenen Zeithorizonten sollten gemacht werden. Die dabei zu berücksichtigenden Zeiträume der Szenarien sollten von Fall zu Fall festgelegt werden.
Abschnitt 4.5 Effektivität (S. 90)	<u>Anmerkung:</u> "Das Ergebnis der KNB stellt das Ergebnis der vorangegangenen Nutzenbewertung gemäß § 35 a SGB V nicht infrage. Das bedeutet z. B., dass in Situationen, in denen im Rahmen des Verfahrens der Nutzenbewertung kein Zusatznutzen nachgewiesen wurde, durch die KNB auch kein Zusatznutzen suggeriert wird." Es erscheint höchst unwahrscheinlich, dass diese Annahme der eigentlichen Intention des Gesetzgebers zur KNB entspricht. Vielmehr lässt etwa in §35b Abs.3 S.3 SGB V erkennen, dass ein Beschluss zur KNB auch eine neue Bewertung zum Zusatznutzen erlaubt, die vom ursprünglichen Beschluss abweichen kann: "Mit dem Beschluss werden insbesondere der Zusatznutzen sowie die Therapiekosten bei Anwendung des jeweiligen Arzneimittels festgestellt." <u>Vorgeschlagene Änderung:</u> Die beiden oben zitierten Sätze sollten ersatzlos gestrichen werden.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 4.6 Nutzwerte (S. 92)	<u>Anmerkungen:</u> a) Bestimmung von Nutzwerten sollte breiter gefasst sein (Patienten können nicht immer solche Gesundheitszustände bewerten, z.B Alzheimer, Pädiatrie). In diesen Fällen sollte eine Bewertung durch Familienangehörige, Betreuungspersonen oder medizinisches Fachpersonal in Betracht gezogen werden. b) Nutzwerte auf Basis von Bewertungen von Gesundheitszuständen durch die Allgemeinbevölkerung werden nur dann zugelassen, <i>“wenn sich die Bewertungen zwischen der Allgemeinbevölkerung und Patientinnen und Patienten nachgewiesen nicht unterscheiden”</i>. Um diesen Nachweis zu erbringen, müssten im Prinzip Bewertungen der Gesundheitszustände von Patientinnen und Patienten (direkte Nutzwerte) vorliegen - in welchem Falle die Bewertungen der Allgemeinbevölkerung nicht notwendig sind. Eine direkte Messung der Nutzwerte erfordert komplexe Methoden, wie z.B. Standard Gamble oder Time Trade off, die in klinischen Studien für Patient:innen und Personal eine große Belastung bedeuten würden, dementsprechend finden sie praktisch keine Anwendung in klinischen Studien (RCT). Soll die gesundheitsbezogene Lebensqualität in RCT erhoben werden, kommen gewöhnlicherweise indirekte Messmethoden mittels Fragebögen, z.B. EQ-5D, zum Einsatz, um Nutzwerte zu ermitteln. Die Nutzung von indirekten Nutzwerten in gesundheitsökonomischen Evaluationen ist in der Gesundheitsökonomie gängige Praxis und kann als internationaler Standard gesehen werden. c) Das Mapping krankheitsspezifischer auf generische Instrumente für die KNB wird nicht empfohlen, obwohl das in Ausnahmefällen eine sinnvolle und effiziente Option darstellt, die auch international üblich ist. <u>Vorgeschlagene Änderungen:</u> a) ..., dass für die KNB im Kontext des AMNOG-Verfahrens die in die Analyse einfließenden Nutzwerte vorrangig auf Bewertungen durch Patientinnen und Patienten basieren sollen. Ist es aufgrund des Zustands der Patienten (z.B. Demenz oder Pädiatrie) nicht möglich, deren Nutzwerte zu erheben, sollten diese durch Einbeziehung des nächsten Umfelds der Patienten erhoben werden können (z.B. Familienangehörige, Betreuungspersonen oder

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	medizinisches Fachpersonal). b) Der Nachweis, dass sich die Bewertungen zwischen Allgemeinbevölkerung und Patientinnen und Patienten nicht unterscheiden, sollte fallen gelassen werden. Die Nutzung von indirekten Nutzwerten, basierend auf Beurteilungen der Allgemeinbevölkerung, sollte zugelassen werden, da dieses Vorgehen der gängigen Praxis der Gesundheitsökonomie entspricht. c) Die Möglichkeit des Mapping sollte zumindest in begründeten Ausnahmefällen möglich sein.
Abschnitt 4.7.1 Kostenkategorien (S. 93 f.)	<u>Anmerkung und vorgeschlagene Änderung:</u> Siehe auch Kommentar zu Abschnitt 4.3.4 Perspektive Es wäre hilfreich, die Perspektiven klar zu definieren und eine Tabelle einzufügen, die erklärt, welche Kosten unter welcher Perspektive anwendbar sind.
Abschnitt 4.7.1 Kostenkategorien, A) Direkte Kosten (S. 93)	<u>Anmerkung:</u> Das ICER wird sich nicht verändern, wenn die Zeitkosten einer anderen Kostenkategorie zugeordnet werden. Es ist jedoch interessant zu sehen, wie sich der Wert durch das Berücksichtigen oder Weglassen verändert. Insofern macht es Sinn, die Zeitkosten separat auszuweisen oder ggfs. eine Sensitivitätsanalyse mit vs. ohne durchzuführen. Eine Sensitivitätsanalyse, bei der diese entweder den direkten oder indirekten Kosten zugerechnet werden, macht unseres Erachtens nach keinen Sinn. <u>Vorgeschlagene Änderung:</u> Da Zeitkosten – in Abhängigkeit von der betrachteten Perspektive – nicht immer trennscharf den direkten, nicht medizinischen Kosten oder den indirekten Kosten zugeordnet werden können, sollten diese gesondert ausgewiesen werden.
Abschnitt 4.7.1 B) Indirekte Kosten (S. 93 f.)	<u>Anmerkung und vorgeschlagene Änderung:</u> Zur Bestimmung der indirekten Kosten gibt es unterschiedliche Methoden (z.B. Humankapital- oder Friktionskostenansatz), welche zu sehr unterschiedlichen Ergebnissen führen können. Des Weiteren geht aus dem Abschnitt nicht hervor, ob indirekte Kosten aus unbezahlter Arbeit in die Analyse aufgenommen werden dürfen oder nicht. Es wäre hier wünschenswert zu definieren, welche Analysen in der Basisanalyse und welche in Szenario- oder Sensitivitätsanalysen gemacht werden sollen.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 4.7.1 Kostenkategorien, F) Investitionskosten (S. 94)	<u>Anmerkung:</u> Der Text besagt, dass <i>“Implementierungskosten angemessen zu berücksichtigen”</i> sind. Weitere Angaben werden nicht gemacht. <u>Vorgeschlagene Änderung:</u> Der Abschnitt sollte ausgebaut werden mit konkreten Angaben wann und wie eine solche <i>“angemessene Berücksichtigung”</i> erfolgen soll. Zum Beispiel müsste bei den Infrastrukturkosten, die für die neue Therapie notwendig sind, die aber dem ganzen Therapiegebiet zugute kommen (z.B. die Anschaffung von MRT Geräten) genau definiert werden, wie diese anteilig bei der spezifischen Therapie Berücksichtigung finden.
Abschnitt 4.9 Studiendesign, Extrapolation (S. 102)	<u>Anmerkung:</u> Für die Extrapolation werden ein konservatives und ein optimistischeres Szenario beschrieben. Das <i>“optimistischeres”</i> Szenario ist dabei keineswegs <i>“optimistisch”</i>. Damit wird ein eingeschränktes und somit verzerrtes Spektrum der möglichen Ergebnisse hergestellt. Um den Einfluss der Annahmen bezüglich Extrapolation umfassend und ausgewogen beurteilen zu können, sollte ein <i>“optimistisches”</i> Szenario eingeschlossen werden. Für bestimmte Therapien ist zu erwarten, dass eine anhaltende Umkehrung der pathologischen Prozesse erreicht werden kann. In diesen Fällen bleibt der therapeutische Effekt über den ganzen Zeithorizont der KNB bestehen. Ferner wird auf andere publizierte Modelle verwiesen. Es ist nicht klar, ob dafür eine systematische Literaturrecherche erforderlich ist und wenn ja, mit welchen Suchkriterien. <u>Vorgeschlagene Änderung:</u> Für die Extrapolation sollten mindestens folgende drei Fälle betrachtet werden: - konservative Annahme: die beobachteten Effekte halten nicht über die Studie hinaus an. - optimistische Annahme: die beobachteten Effekte bleiben nach Ende der Studie über den ganzen Zeithorizont bestehen. - <i>“mittlere”</i> Annahme: die Effekte halten über das Studienende hinaus an, fallen aber kontinuierlich ab. Die Szenarien sollten entsprechend angepasst werden und diese drei Annahmen abbilden.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 3.3.4 „Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD- Konzepte)“	<u>Anmerkung:</u> Zur Vorbereitung eines GBA-Beschlusses über eine durchzuführende AbD erstellt das IQWiG ein AbD-Konzept. Das IQWiG schlägt in Abschnitt 3.3.4 vor in einem AbD-Konzept neben weiteren Themen den Umfang einer AbD zu beschreiben. Dazu heißt es: <i>„Beschreibung des Umfangs der AbD: Aus einer nicht randomisierten Studie kann aufgrund potenziell unbekannter Confounder aus den in der Studie beobachteten Effekten erst ab einer bestimmten Effektstärke eine Aussage zum Nutzen oder Schaden einer Intervention abgeleitet werden (siehe Abschnitte 3.1.4 und 3.2.2), z. B. in einem Bereich von 2 bis 5 für das relative Risiko (bzw. 0,2 bis 0,5 bei mortalitätssenkenden Interventionen) bezogen auf die untere bzw. obere Konfidenzintervallgrenze [393]. Dies wird bei der Frage, ob die geplante AbD prinzipiell sinnvoll durchführbar ist, berücksichtigt. Basierend auf den vorliegenden Erkenntnissen zum Arzneimittel und zur zweckmäßigen Vergleichstherapie wird abgeschätzt, ob innerhalb eines akzeptablen Zeitraums prinzipiell ausreichend Patientinnen und Patienten rekrutiert und beobachtet werden können, um mit einer AbD grundsätzlich aussagekräftige Ergebnisse für eine Nutzenbewertung generieren zu können“</i> Im Folgenden möchte Roche auf weitere Aspekte genauer eingehen: 1. Roche begrüßt einen Überprüfungsschritt der Realisierbarkeit einer AbD bevor diese seitens des GBA beschlossen wird. Neben den vom IQWiG aufgelisteten Punkten sollte überprüft werden, ob die Realisierbarkeit der AbD durch die Verschreibungspraxis oder Veränderungen in der Therapielandschaft beeinträchtigt wird. Der Einfluss dieser Ereignisse auf die zu erwartenden Populationsgrößen sollte daher genau geprüft werden. 2. In den bisherigen Verfahren wurde vom IQWiG vorgeschlagen, im Falle vorliegender Unsicherheit in den Annahmen zur Fallzahlplanung (wie z.B. der Effektstärke) auf eine Fallzahlplanung und deren Realisierbarkeit

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	zunächst zu verzichten und diese erst in den Zwischenanalysen durchzuführen. Es stellt sich die Frage, wie bei fehlender Fallzahlplanung der Umfang einer AbD beschrieben werden kann. 3. Die Ableitung eines Zusatznutzens soll erst ab einer gewissen Effektstärke erfolgen. Hierfür wählt das IQWiG einen Bereich von 2 bis 5 für das relative Risiko (bzw. 0,2 bis 0,5) und verweist auf den Rapid Report [1], in dem es wie folgt heißt: <i>“Da die Erfüllung der genannten Qualitätsanforderungen Voraussetzung für die Betrachtung der Effekte ist, sollte dieser Schwellenwert deutlich unterhalb des Wertes für den „dramatischen Effekt“ (relatives Risiko von 5-10) liegen, z. B. in einem Bereich von 2-5 für das relative Risiko. Die konkrete Schwelle ergibt sich durch die Qualität der Daten im Einzelfall.”</i> Für Roche stellen sich hierzu folgende Fragen: • Der vorgeschlagene Bereich von 2-5 für das relative Risiko lässt sich nicht aus dem Methodenpapier oder dem Rapid Report herleiten. Die wissenschaftliche Rationale, auf der die Festlegung basiert, sollte genau erläutert werden. • Wie kann die Qualität der Daten operationalisiert werden? • Wie hängt der Bereich von 2 bis 5 für das relative Risiko für die Ableitung des Zusatznutzens von der Qualität der Daten ab? • Ab welcher Effektstärke kann für stetige Endpunkte ein Zusatznutzen abgeleitet werden? <u>Vorgeschlagene Änderung:</u> • Bei auftretenden Veränderungen, die die Therapielandschaft der jeweiligen Indikation betreffen, sollte ermittelt werden, welches der folgenden Szenarien vorliegt:

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	1. Ursprünglich erwartete Populationsgröße bleibt unverändert 2. Ursprünglich erwartete Populationsgröße nimmt ab und verschwindet nach und nach 3. Ursprünglich erwartete Populationsgröße verschwindet nach sehr kurzer Zeit Während im Fall des Szenario 1 kein Einfluss auf die Realisierbarkeit der AbD anzunehmen ist, sollte dagegen in Szenario 3 keine Realisierbarkeit und keine Relevanz der AbD angenommen werden und damit kein Beschluss zur Durchführung einer AbD in der entsprechenden Population erfolgen. Für Szenario 2 sollten Kriterien zur Überprüfung der Realisierbarkeit einer AbD und daraus folgende Maßnahmen beschrieben werden. • Roche schlägt für die Definition eines Bereiches bei stetigen Endpunkten vor, sich an den für Cohen's d [2] definierten Bereichen für kleine (SMD zwischen 0,2 und 0,5), mittlere (SMD zwischen 0,5 und 0,8) und große (SMD ab 0,8) Effekte zu orientieren.
Abschnitt 9.3.6: Adjustierung und multifaktorielle Verfahren	<u>Anmerkung:</u> In Abschnitt 9.3.6 beschreibt das IQWiG, dass bei Anwendung von Propensity-Score-Methoden die Annahmen wie z.B. Überlappung zu prüfen und darzustellen sind. Dazu schreibt das IQWiG: „Um im Rahmen von Propensity-Score-Ansätzen eine ausreichende Überlappung und Balanciertheit zu erreichen, existieren zahlreiche Methoden [34,162,180,464,758].“ Aus Sicht von Roche ist dies ein wichtiges Thema, welches einer gemeinsamen normativen Einigung bedarf. Zwar existieren Methoden, um die Überlappung der Propensity Scores zwischen den zu vergleichenden Gruppen zu untersuchen, dennoch gibt es

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	bisher kein etabliertes Kriterium, um zu bewerten, ob eine Überlappung als ausreichend anzusehen ist. <u>Vorgeschlagene Änderung:</u> Roche hält zur Lösung von dieser und ähnlich gelagerten Fragestellungen ein Gremium für unabdingbar, das aus IQWiG, G-BA, Akademie, und Industrievertretern besteht und schon verschiedentlich von Roche und der AG Biostatistik vorgeschlagen wurde. In diesem Gremium sollte gemeinsam eine Konvention erarbeitet werden, die für die Anwendung der Propensity Score Methodik eine ausreichende Überlappung zweier Gruppen darstellt.
9.3.2 Statistische Signifikanz und Konfidenzintervalle	Mit Blick auf statistische Signifikanztests wird im Entwurf des Methodenpapiers beschrieben, dass es in Abhängigkeit der Fragestellung und der geplanten Aussage „<i>erforderlich sein [kann], ein anderes Signifikanzniveau zu fordern</i>“. Aus Sicht von Roche kann dies sinnvoll sein (nicht umsonst sind Methoden des multiplen Testens entwickelt worden), muss aber nicht in allen Fällen zielführend sein. Ganz allgemein sei auf den Allgemeinplatz verwiesen, dass die Nutzenbewertung auf Fragen abhebt, die in erster Linie auf Schätzprobleme abheben (i.e. PE & CI) denn auf reines Testen und eine Adjustierung bei ersterem eher unüblich ist. Letzteres sind Fragestellungen im vornehmlich regulatorischen Umfeld. Eine Überschneidung (i.e. Testprobleme in der Nutzenbewertung) gibt es aus Sicht von Roche bei der Prüfung von Effektmodifikatoren, etwa bei den Subgruppenanalysen und den zugehörigen Interaktionstests. Die Fragestellung ist hierbei aber auch eine andere als in der sonstigen Nutzenbewertung. Hier stellt sich die Frage: Gibt es eine Effektmodifikation oder nicht? Für diesen Aspekt der Nutzenbewertung macht aus Sicht von Roche eine multiple Teststrategie durchaus Sinn. Für Roche passen demgegenüber Fragen zu Sensitivitätsanalysen und verschiedenen Szenarien nicht in das Schema für das multiple Testen. Hierbei sollen v.a. Aspekte der Robustheit der Analysen,

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	von veränderten Rahmenbedingungen bzw. der Definition der Endpunkte näher betrachtet werden. Insbesondere bei - unterschiedlichen Effektmaßen, - unterschiedlichen Skalenniveaus, - Slicing/ZVT-Zuweisung etc, - biologisch leicht variierten Operationalisierungen des Endpunktes würde eine multiple Adjustierung der Idee dieser Sensitivitätsanalysen zuwiderlaufen. <u>Vorgeschlagene Änderung:</u> Die sehr allgemein gewählte Formulierung im Methodenpapier sollte spezifiziert und stark auf die Effektmodifikation eingegrenzt werden. Es sollte darüber hinaus deutlich gemacht werden, dass die Multiplizitätsadjustierung nicht für Sensitivitätsanalysen und Szenarien gilt.

Literaturverzeichnis

1. IQWiG-Bericht Nr. 863. Rapid Report. Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimittel nach § 35a SGB V: Version 1.1. Stand 13.05.2020
2. Jacob Cohen, Hrsg. Statistical Power Analysis for the Behavioral Sciences (Revised Edition): The t Test for Means; 1977.

A.1.35 Sanofi-Aventis Deutschland GmbH

Autorinnen und Autoren

- Hahn, Steffen
- Zietze, Henny

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Hahn, Steffen
Zietze, Henny
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation:
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
9.3.12 Umgang mit variablen Beobachtungszeiten (S.190 – 191)	<u>Anmerkung:</u> Es wird kritisiert, dass „bei sich deutlich zwischen den Behandlungsarmen unterscheidenden Beobachtungsdauern sogenannte „dauerhafte“ Veränderungen (Verbesserungen bzw. Verschlechterungen) nicht sinnvoll interpretierbar sind, da eine „dauerhafte“ Veränderung in dem länger beobachteten Behandlungsarm aufgrund der dafür notwendigen ununterbrochenen Über- bzw. Unterschreitung des Schwellenwerts (Responsekriterium) potenziell schwerer zu erreichen ist“. Dass deutliche Unterschiede in den Beobachtungsdauern zwischen den Behandlungsarmen zu einer Verzerrung führen können, ist unstrittig. Allerdings ist nicht klar, ab wann es sich um einen deutlichen Unterschied handelt. Zudem ist es nicht nachvollziehbar, warum im Falle unterschiedlicher Beobachtungsdauern zwischen den Behandlungsarmen die Analysen der Zeit bis zur dauerhaften Veränderung generell als nicht interpretierbar von der Nutzenbewertung auszuschließen sind. Die dauerhafte Verbesserung, definiert als eine ab einem Zeitpunkt in allen Folgevisiten gemessene Verbesserung oberhalb der Responseschwelle, ist nach obiger Argumentation in dem länger beobachteten Behandlungsarm potenziell schwerer zu erreichen. Wurde beispielsweise das Eintreten einer Progression als Grund für die Beendigung der Beobachtung festgelegt und es liegt ein Vorteil für Verum bezüglich des progressionsfreien Überlebens vor, so wird der Verum-Arm länger beobachtet. Folglich

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	wäre eine dauerhafte Verbesserung bezüglich des PRO Endpunktes im Verum-Arm potenziell schwerer (oder zumindest erst zu einem späteren Zeitpunkt) zu beobachten. Daraus resultiert eine mögliche Verzerrung zugunsten von Kontrolle. Zeigt sich nun dennoch ein Vorteil zugunsten von Verum, so ist es nicht nachvollziehbar, warum dieser nicht berücksichtigt werden sollte. Möglich wäre zudem auch der Fall, dass eine Verschlechterung im Kontrollarm früh (vor einer potenziellen Zensurierung) eintritt und dann bestehen bleibt. In diesem Fall würde eine im Vergleich zum Verum-Arm kürzere Beobachtungsdauer im Kontrollarm die Zeit bis zur Verschlechterung nicht beeinflussen. Falls sich also in einer Auswertung ein konsistenter Vorteil für die Zeit bis zur ersten, bestätigten (in zwei aufeinanderfolgenden Visiten) und dauerhaften Verschlechterung zeigt und auch bei Patienten mit längerer Beobachtungsdauer im Kontrollarm nur in Einzelfällen eine Verbesserung eintritt, so ist nicht ersichtlich, weshalb die Ergebnisse der Analyse der Zeit bis zur dauerhaften Verschlechterung nicht interpretierbar sein sollten. <u>Vorgeschlagene Änderung:</u> Die dauerhafte Veränderung sollte auch bei unterschiedlichen Beobachtungsdauern nicht pauschal als nicht interpretierbar verworfen werden. Es gibt Situationen, in denen die Ergebnisse durchaus sinnvoll interpretierbar sind, da eine Verzerrung ausgeschlossen werden kann oder diese zugunsten der Kontrolle ausfällt. Daher ist es geboten im Einzelfall zu entscheiden, ob die dauerhafte Veränderung für die Nutzenbewertung herangezogen werden kann.
9.3.2 Statistische Signifikanztests und Konfidenzintervalle (S. 170)	<u>Anmerkung:</u> Bei der Erläuterung des multiplen Testproblems in systematischen Übersichten wurde ein Beispiel für die sinnvolle Anwendung einer Adjustierung hinzugefügt. „Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an [59], z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.“ Es bleibt allerdings weiterhin unklar, wann eine Adjustierung als sinnvoll und möglich erachtet wird und es ist auch nicht ersichtlich, was die Änderung im Methodenpapier bezwecken soll. Vielmehr ist das Konzept der Adjustierung für multiples Testen im Kontext der Nutzenbewertung nicht sinnvoll auf post-hoc Analysen anwendbar. Verschiedene Operationalisierungen für einen Endpunkt haben häufig supportiven Charakter und eine Adjustierung des Signifikanzniveaus ist somit nicht zielführend.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Vorgeschlagene Änderung:</u> Aus dem Methodenpapier sollte klar ersichtlich sein, dass verschiedene Operationalisierungen, wie z.B. bei supportiven oder Sensitivitätsanalysen, nicht regelhaft zu einer Anwendung einer Adjustierung für multiples Testen führen. Eine Erläuterung anhand von Beispielen, wann eine Adjustierung für multiples Testen in der Nutzenbewertung anzuwenden ist und wie diese dann umzusetzen ist, sollte eingefügt werden. Die Voraussetzungen hierfür sollten klar definiert werden.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

A.1.36 SmartStep Consulting GmbH

Autorinnen und Autoren

- Volmer, Timm

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Volmer, Timm; Dr.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: SmartStep Consulting GmbH
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>
Keine Anmerkungen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
Abschnitt 4.6 (S. 91 f.)	Nutzwerte (im Fall einer Kosten-Nutzwert-Analyse) <u>Anmerkung:</u> Das IQWiG formuliert im Abschnitt 4.6, dass die Nutzwerte vorrangig auf Bewertungen durch Patientinnen und Patienten basieren sollen und beruft sich hierbei auf die systematische Literaturübersicht von Helgesson et al.. <u>Vorgeschlagene Änderung:</u> Im Kontext internationaler Erhebungen sollte die Bewertung der Nutzwerte auf der Allgemeinbevölkerung beruhen. Helgesson et al. merken an, dass ein erheblicher Verzerrungseffekt entstehen kann, da Patienten, zum Selbstschutz, ihre neue Krankheitssituation zu optimistisch bewerten und folglich eine endgültige Empfehlung der Bewertungsweise nicht möglich ist. (Helgesson et al. 2020) Des Weiteren stellt das IQWiG selbst fest, dass die Datenlage eingeschränkt ist im Vergleich zu Nutzwerten basierend auf der Allgemeinbevölkerung und das letzteres auch der internationalen Praxis entspricht. Dieses Vorgehen entspricht zudem den Empfehlungen der Modeling Good Research Practices Task Force Working Group der ISPOR (Brazier et al. 2019)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 9.3.12 (S. 190 ff.)	Umgang mit variablen Beobachtungszeiten Nachbeobachtung von unerwünschten Ereignissen (UEs) bis 30 Tage nach Behandlungsabbruch bzw. -wechsel und Relevanz der Ergebnisse für die Nutzenbewertung <u>Anmerkung:</u> Das IQWiG formuliert im Abschnitt 9.3.12 zur Erhebung von UEs und der im Rahmen klinischer Studien gängigen Praxis der Nachbeobachtung der UEs von bis zu 30 Tagen nach Behandlungsabbruch- bzw. -wechsel, dass diese Art der Operationalisierung „für die Nutzenbewertung unzureichend“ sei. <u>Vorgeschlagene Änderung:</u> Es wird sich dafür ausgesprochen, dass die Formulierung „Die häufig im Rahmen klinischer Studien angewendete Praxis, unerwünschte Ereignisse nur bis maximal 30 Tage nach Behandlungsabbruch oder -wechsel zu erfassen, ist für die Nutzenbewertung unzureichend.“ auf „Die häufig im Rahmen klinischer Studien angewendete Praxis, unerwünschte Ereignisse nur bis maximal 30 Tage nach Behandlungsabbruch oder -wechsel zu erfassen, kann zu Einschränkungen in der Bewertbarkeit der Relevanz der Daten für die Nutzenbewertung führen, so dass insgesamt eine längere Beobachtungsdauer wünschenswert ist.“ angepasst wird. Ein genereller Ausschluss von der Bewertung ist nicht hinreichend begründet dargelegt. Darüber hinaus erfolgt die regelhafte Nachbeobachtung von Sicherheitssignalen im Rahmen der gesetzlich vorgesehenen Pharmakovigilanz-Aktivitäten wie beispielsweise der Erstellung von Periodic Safety Update Reports (PSUR) sowie ggf. in Form von Post Authorization Safety Studies (PASS). Somit ist die Erfassung von unerwünschten Ereignissen für den Lebenszyklus eines Arzneimittels sichergestellt; etwaige neue Sicherheitssignale können mit Hilfe der vorgenannten Maßnahmen erfasst und entsprechend in die relevanten Produktinformationstexte aufgenommen werden. Des Weiteren sieht das IQWiG selbst, dass bei „informativer Zensierung“ „ggf. ein hohes Verzerrungspotenzial“ entsteht. Dies impliziert somit, dass es nicht zur vollständigen Nichtberücksichtigung der Daten kommen sollte, sondern dies insbesondere bei der hier vorliegenden nicht informativen Zensierung lediglich die Aussagesicherheit beeinflussen kann.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
	Verkürzte Beobachtungszeiten bei Patient Reported Outcomes (PROs) <u>Anmerkung:</u> Das IQWiG formuliert auch für die PROs hinsichtlich verkürzter oder sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauern eine zu stark einschränkende und nicht den Einzelfall berücksichtigende Formulierung. Das IQWiG schreibt, dass Auswertungen mit unterschiedlicher Beobachtungsdauer „daher in der Regel nicht für die Nutzenbewertung herangezogen“ werden. <u>Vorgeschlagene Änderung:</u> Die Formulierung „daher in der Regel nicht für die Nutzenbewertung herangezogen“ sollte angepasst werden auf „kann innerhalb der Nutzenbewertung zu einer eingeschränkten Bewertbarkeit der vorgelegten Daten führen“.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

Helgesson G, Ernstsson O, Astrom M et al. Whom should we ask? A systematic literature review of the arguments regarding the most accurate source of information for valuation of health states. Qual Life Res 2020; 29(6): 1465-1482.

Brazier J, Ara R, Azzabi I et al. Identification, Review, and Use of Health State Utilities in Cost-Effectiveness Models: An ISPOR Good Practices for Outcomes Research Task Force Report. Value Health. 2019 Mar; 22(3): 267-275.

A.1.37 UCB Pharma GmbH

Autorinnen und Autoren

- Andreas, Jens-Otto
- Brunnert, Marcus
- Dunkel, Jochen
- Hoesel, Volker
- Kumke, Thomas
- Marchowez, Marina

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Jens-Otto Andreas
Marcus Brunnert
Jochen Dunkel
Volker Hoesel
Thomas Kumke
Marina Marchowez
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen) ☒ im Namen folgender Institution / Organisation: UCB Pharma GmbH ☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>
Eine Kennzeichnung der geänderten Stellen im Vergleich zur Vorgängerversion wäre für das Stimmungsverfahren eine Erleichterung.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Kapitel 9 „Informationsbewertung“ /Abschnitt 9.1.1 (Seite 159)	<u>Anmerkung:</u> Die Kriterien zum Einschluss relevanter Studien werden im Entwurf zu Version 7 des Methodenpapiers überarbeitet. Dabei ergeben sich aus zwei Änderungen Unklarheiten im Vergleich zur Version 6.1: - Wird in Version 6.1 noch ausdrücklich das Kriterium für den Ausschluss von Studien von der Nutzenbewertung beschrieben, so ist dies im Entwurf zu Version 7.0 nicht mehr der Fall. D.h. es ergibt sich nun die Frage, ob in jedem Fall Studien mit weniger als 80% Patienten, die das Einschlusskriterium bzgl. Vergleichs und -Interventionsarm für Studien erfüllen, weiterhin generell ausgeschlossen werden. - Der Satz „Liegen für eine Studie Analysen der relevanten Teilpopulation bzw. der relevanten Prüf- und Vergleichsinterventionen vor, wird auf diese Analysen zurückgegriffen.“ wird nun im Vergleich zur Version 6.1 vorgezogen. Werden nun auch Studien akzeptiert, die das 80% Kriterium nicht erfüllen? - Der Satz „Falls geeignete Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium (Population bzw. Interventionen) vorliegen, muss über die Berücksichtigung der Studie anhand der Stärke der Effektmodifikation und dem Anteil der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen bzw. dem Grad der Abweichung der Interventionen entschieden werden.“ wird im Entwurf zu Version 7.0 neu aufgenommen.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Dabei bleibt unklar was geeignete Informationen über eine Effektmodifikation sind. <u>Vorgeschlagene Änderung:</u> Aus Sicht von UCB, sollte das Kriterium zum Ausschluss von Studien präzisiert werden. Insbesondere sollten die geeigneten Informationen über eine Effektmodifikation näher erläutert werden. Dabei sollten objektivierbare Kriterien definiert werden.
Kapitel 9 „Informationsbewertung“ /Abschnitt 9.3.2 (Seite 170)	<u>Anmerkung:</u> Im Entwurf der Allgemeinen Methoden Version 7.0 wird ein Hinweis auf die mögliche Akzeptanz von bayesschen Methoden bei Metaanalysen ergänzt und auf Abschnitt 9.3.7 verwiesen. Allerdings erfolgt in Abschnitt 9.3.7 keine Überarbeitung zu den bayesschen Verfahren. Kriterien für die Anwendung von bayesschen Methoden werden in den beiden Abschnitten (9.3.2 und 9.3.7) nicht genannt. Es bleibt offen, wann bayessche Verfahren akzeptiert werden. Der Einfluss des Satzes: „Wenn sinnvoll und möglich, wendet das Institut Methoden zur Adjustierung für multiples Testen an [1], z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist.“ auf zukünftige Nutzenverfahren ist unklar. <u>Vorgeschlagene Änderung:</u> Sowohl Kriterien für die Anwendung von bayesschen Verfahren bei Metaanalysen als auch allgemeine Vorgaben für bayesschen Verfahren sollten aus Sicht von UCB im IQWiG Methodenpapier Version 7.0 ergänzt werden. Aus Sicht von UCB sollte klargestellt werden, wann das Institut ggf. Methoden zur Adjustierung für multiples Testen anwendet. Das hier gewählte Beispiel ist nicht notwendig und z.T. auch irreführend, da verschiedene Operationalisierungen eines Endpunktes gegebenenfalls aus wissenschaftlichen oder medizinischen Gründen sinnvoll sind, beispielsweise als unterstützende oder auch Sensitivitätsanalysen. Zu einem Endpunkt ist eine Hauptanalyse definiert. Dadurch entfällt die Notwendigkeit der Adjustierung für multiples Testen.
Kapitel 9 „Informationsbewertung“	<u>Anmerkung:</u> Es wird von einer ausreichenden Überlappung und Balanciertheit zwischen den zu vergleichenden Gruppen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigelegt werden.</i>
/Abschnitt 9.3.6 (Seiten 175 bis 178)	bzgl. der Propensity Scores gesprochen. Aber es bleibt unklar was „ausreichend“ meint. <u>Vorgeschlagene Änderung:</u> Im Entwurf der Allgemeinen Methoden Version 7.0 sollte aus Sicht von UCB die ausreichende Überlappung und Balanciertheit zwischen den zu vergleichenden Gruppen bzgl. Propensity Scores unter Verwendung in der Literatur vorgeschlagener Kriterien näher beschrieben werden.
Kapitel 9 „Informationsbewertung“ /Abschnitt 9.3.12 (Seiten 190 und 191)	<u>Anmerkung:</u> Zwei neue Absätze zu Verfahren bei ungleichen Beobachtungsdauern wurden in Version 7.0 ergänzt. Es bleibt aber unklar, wie schon in Version 6.1, ab wann unterschiedliche Beobachtungsdauern zu einem Ausschluss der Analysen von der Nutzenbewertung führen. <u>Vorgeschlagene Änderung:</u> Aus Sicht von UCB sollte ein objektives Kriterium ergänzt werden, das die unterschiedlichen Beobachtungsdauern bzgl. ihrer Nutzenbewertungsrelevanz beurteilen lässt.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

- (1) Bender R, Lange S. Adjusting for multiple testing – when and how? J Clin Epidemiol 2001; 54(4): 343-349.

A.1.38 Verband Forschender Arzneimittelhersteller e. V. (vfa)

Autorinnen und Autoren

- Rasch, Andrej
- Bussilliat, Paul

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Rasch, Andrej; Dr.
Bussilliat, Paul
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☒ im Namen folgender Institution / Organisation: Verband Forschender Arzneimittelhersteller e.V. - vfa
☐ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

<i>Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.</i>
In den Änderungen im Entwurf der Allgemeinen Methoden 7.0 ergeben sich an mehreren Stellen wichtige Änderungen, die für die forschenden Pharmaunternehmen von Bedeutung sind. Die Änderungen beziehen sich auf das nun im Methodenpapier erstmals aufgenommene Vorgehen im Rahmen der anwendungsbegleitenden Datenerhebung (AbD) hinsichtlich des Verfahrensablaufs (Abschnitt 2.1.3.2), der Konzepte für eine AbD (Abschnitt 3.3.4), der Recherche nach Indikationsregistern (Abschnitt 2.1.3.3) sowie auf Ergänzungen zur Adjustierung für Confounder (Abschnitt 9.3.6). Grundlage hierfür bilden in großen Teilen die Inhalte des Rapid Report A19-43 (Konzepte zur Auswertung versorgungsnaher Daten zum Zwecke der Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V) in der Version 1.1 vom 14.05.2020. Da bei Erstellungen der Rapid Reports des IQWiG ein Stimmnahmeverfahren nicht vorgesehen war, ergibt sich nun erstmalig die Gelegenheit hierzu Stellung zu nehmen. Wesentliche Änderungen ergeben sich zudem für das Vorgehen einer Kosten-Nutzen-Bewertung (Abschnitte 1.4 und 4). Für die Nutzenbewertung von Arzneimitteln wurden außerdem weitere Anpassungen zu Kriterien für den Einschluss von Studien (Abschnitts 9.1.1), zu statistischer Signifikanz und Konfidenzintervallen (Abschnitt 9.3.2) sowie zur Auswirkung systematisch verkürzter Beobachtungszeiten bei der Erhebung patientenberichteter Endpunkte (PROs) (Abschnitt 9.3.12) vorgenommen. Zu den o.g. Änderungen wird im Detail Stellung genommen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung
	<i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Abschnitt 1.3.3 Grundlegende Strategien der EbM, S. 7	<u>Anmerkung:</u> Bei der Frage nach einer Definition des Nutzens (oder des Schadens) einer medizinischen Maßnahme, ergänzt das IQWiG die Reduktion beeinträchtigender Nebenwirkungen. Die Beeinträchtigung von Patientinnen und Patienten kann jedoch sehr subjektiv und heterogen sein bzw. verschiedene Lebensbereiche betreffen. <u>Vorgeschlagene Änderung:</u>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Der Begriff „beeinträchtigende Nebenwirkung“ sollte genauer erläutert werden, um klarzustellen, inwiefern sich eine Beeinträchtigung auch auf solche Bereiche wie das tägliche Arbeiten oder den Einfluss auf das soziale Umfeld bezieht.
Abschnitt 2.1.3.2 AbD-Konzept, S. 20-21	<u>Anmerkung:</u> Das IQWiG stellt den Ablauf bei der Erstellung eines AbD-Konzepts dar. Die frühzeitige Möglichkeit zur Einbindung von pharmazeutischen Unternehmen (pU) ist hierbei nicht vorgesehen. Der Verzicht auf die Expertise der pU ist nicht nachvollziehbar. Die pU verfügen über spezifisches Wissen im jeweiligen Anwendungsgebiet sowie ggf. Erfahrungen zu Registerstudien und Registern. Unter Umständen können solche Register sogar bereits im Aufbau sein. Eine frühzeitige Einbindung des pU erscheint daher sinnvoll, um ein praxisnahes, tragfähiges und realisierbares Konzept zu erstellen. Die Einbindung der medizinischen Expertise von externen Sachverständigen sowie von Betroffenen bzw. Patientenorganisationen ist für die Konzepterstellung elementar. In der bisherigen Praxis ist unklar, ob das in einem hinreichenden Ausmaß stattfindet. <u>Vorgeschlagene Änderung:</u> Einbindung des pU im Rahmen der Informationsbeschaffung (entsprechende Ergänzung in der Darstellung auf S. 20): „Kontaktaufnahme mit Betreibern potenziell für die AbD geeigneter Datenplattformen sowie pharmazeutischer Unternehmen (z. B. Indikationsregister, Produktregister)“ Es sollte explizit aufgeführt werden, dass bei der Erarbeitung des AbD-Konzepts externe Sachverständige und Fachgesellschaften für die medizinische Expertise sowie Betroffene bzw. Patientenorganisationen für die Patientenperspektive eingebunden werden (entsprechende Ergänzung in der Darstellung auf S. 20).
Abschnitt 2.1.3.3 Recherche nach Indikationsregistern, S. 21-22	<u>Anmerkung:</u> Das IQWiG stellt den Ablauf der Recherche nach Indikationsregistern dar. Dabei heißt es, dass „eine fokussierte Informationsbeschaffung in verschiedenen Informationsquellen“ erfolge. Die Darstellung ist nicht hinreichend spezifisch. Eine Anfrage beim betroffenen pU wird dabei nicht explizit als Teil der Informationsbeschaffung beschrieben. Das ist unverständlich, da betroffene pUs ggf. Indikationsregister in Planung haben oder ein solches sogar zeitnah in Betrieb geht. Hierbei handelt es sich um hilfreiche

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Informationen, die im Rahmen der Recherche nach Indikationsregistern gesammelt und entsprechend an den G-BA übermittelt werden sollten. <u>Vorgeschlagene Änderung:</u> Im Rahmen der Recherche nach Indikationsregistern sollte der betroffene pU als eine Informationsquelle adressiert werden.
Abschnitt 3.3.4 Konzepte für anwendungsbegleitende Datenerhebungen gemäß § 35a Abs. 3b SGB V (AbD-Konzepte). S. 66-68	<u>Anmerkung:</u> Das Konzept der anwendungsbegleitenden Datenerhebung kann insgesamt für bestimmte Therapiesituationen eine sinnvolle Möglichkeit der Evidenzgenerierung bieten, wenn eine randomisiert-kontrollierte Studie nicht möglich ist. In Abschnitt 3.3.4 werden die Inhalte und methodischen Grundzüge des Konzepts zur AbD skizziert, die in großen Teilen die Inhalte des Rapid Report A19-43 widerspiegeln. Die Erstellung eines AbD-Konzepts dient schlussendlich dem G-BA zur Prüfung, ob die geplante AbD prinzipiell realisierbar ist (vgl. S. 66). Die bisher veröffentlichten Konzepte zeigen jedoch deutlich, dass bei den Anforderungen die Umsetzbarkeit unzureichend berücksichtigt wurde. Teilweise erweisen sich die Vorgaben bisher als praxisfern. Bspw. werden die geforderten Spitzenwerte in der Datenqualität derzeit in deutschen Patientenregistern fast nie erreicht und sind auch hinsichtlich der Weiterentwicklung kaum erreichbar. Auch geht das geforderte Maß der kontrollierten Studiendurchführung am Vorbild klinischer Studien weit über die übliche Praxis von Versorgungsstudien hinaus. Zusätzlich ist die Methodik der Auswertung zur Beurteilung der AbD-Ergebnisse nicht beschrieben. Die Besonderheiten der AbD bei der Datenerhebung werden nicht berücksichtigt. Die Anforderungen an die Confounder-Adjustierung erscheinen als kaum in der Praxis realisierbar. Zugleich ist die besondere Situationen der Confounder-Adjustierung für kleine Fallzahlen nicht beschrieben, obwohl bestimmte Methoden nur bei größeren Stichproben realistisch und sinnvoll möglich sind. Nicht sachgerecht ist außerdem die Anforderung eines dramatischen Effekts, da sich die Vorgaben zudem auf Konfidenzintervallgrenzen und nicht auf die Punktschätzer beziehen („Aus einer nicht randomisierten Studie kann aufgrund potenziell unbekannter Confounder aus den in der Studie beobachteten Effekten erst ab einer bestimmten Effektstärke eine Aussage zum Nutzen oder Schaden einer Intervention abgeleitet werden (siehe Abschnitte 3.1.4 und

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	3.2.2), z. B. in einem Bereich von 2 bis 5 für das relative Risiko (bzw. 0,2 bis 0,5 bei mortalitätssenkenden Interventionen) bezogen auf die untere bzw. obere Konfidenzintervallgrenze“). Diese Anforderungen unterscheiden sich somit de facto nicht von den üblichen Anforderungen der Nutzenbewertung gemäß dem Abschnitt 3.2.2 (Dramatischer Effekt), wonach ein dramatischer Effekt vorliegt, wenn das beobachtete relative Risiko von 5 bis 10 vorliegt. Die AbD-Vorgaben für die Effekte berücksichtigen somit weder die zusätzlich zu erfolgende Confounder-Adjustierung noch die Besonderheiten der Therapiesituation (insb. die kleinen Patientenzahlen), die gemäß Auftrag des G-BA zur AbD jeweils vorliegt. Unter der zusätzlichen Berücksichtigung der Vorgaben an die Confounder-Adjustierung sind die Anforderungen an die Effekte in der AbD damit sogar konservativer als in der Nutzenbewertung nach §35 SGB V. Zugleich soll abgeschätzt werden, „ob innerhalb eines akzeptablen Zeitraums prinzipiell ausreichend Patientinnen und Patienten rekrutiert und beobachtet werden können, um mit einer AbD grundsätzlich aussagekräftige Ergebnisse für eine Nutzenbewertung generieren zu können.“ Hier ist keine nachvollziehbare Prüfung vorgesehen, inwiefern Veränderungen (durch neue Zulassungen oder andere äußere Einflüsse wie z.B. Neugeborenen-Screening) zu erwarten sind und zu welchen Auswirkungen dies (auch kurzfristig) auf die ursprüngliche Population führt. Ebenso unklar ist, wie Situationen ohne sinnvolle Behandlungsalternativen beurteilt werden. Unter solchen methodischen Bedingungen erscheint der Nachweis eines Zusatznutzens, insbesondere bei seltenen Erkrankungen, nur in absoluten Ausnahmefällen machbar. Die Vorgaben gefährden so die Realisierbarkeit der AbD und damit auch die vom Gesetzgeber gesetzten Ziele nach einer vermehrten Einbeziehung und Bewertung von anwendungsbegleitenden Daten im AMNOG. Die hier und vor allem im Rapid Report A19-43 beschriebenen Vorgaben sind bezüglich ihrer Praxistauglichkeit einer kritischen Prüfung zu unterziehen, zum einen, um die speziellen Herausforderungen bei besonderen Therapiesituationen zu würdigen, aber auch um zu gewährleisten, dass die Ableitung eines Zusatznutzens ohne Randomisierung möglich ist. <u>Vorgeschlagene Änderung:</u>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Eine kritische Würdigung und daraus resultierende grundlegende Überarbeitung des Rapid Reports A19-43 ist notwendig. Die Methodik der Auswertung zur Beurteilung der AbD-Ergebnisse muss aufgenommen werden. Insb. muss von den der Anforderung eines dramatischen Effekts (im Bereich von 2 bis 5 bezogen auf die untere bzw. obere Konfidenzintervallgrenze) zusätzlich zur umfangreichen Confounder-Adjustierung Abstand genommen werden, da es sich hier um eine praxisferne, überhöhte Anforderung handelt. Die Fallzahlplanung sollte Situationen ohne sinnvolle Behandlungsalternativen nachvollziehbar berücksichtigen und eine Prüfung enthalten, ob Veränderungen zu erwarten sind und zu welchen Auswirkungen auf die Population dies führt. Das Vorgehen für Confounder-Identifizierung sollte deutlicher beschrieben werden.
Abschnitt 9.3.6 Adjustierung und multifaktorielle Verfahren, S. 175-177	<u>Anmerkung:</u> Das IQWiG erachtet es als „<i>erforderlich, vorab alle relevanten Confounder (inklusive wichtiger Wechselwirkungen) mithilfe eines systematischen Ansatzes zu identifizieren [587], vollständig zu erheben und bei der Datenanalyse in angemessener Form im Modell zu berücksichtigen [393]. [...] Im Rahmen einer AbD spielt zur Adjustierung für Confounder insbesondere die Propensity-Score-Methode eine Rolle</i>“. Bei der Forderung eines systematischen Ansatzes zur Identifikation relevanter Confounder verweist das IQWiG als methodische Grundlage ausschließlich auf die Publikation von Pufulete et al. Hier wäre eine breitere und nachvollziehbare Evidenzgrundlage erforderlich. Zugleich zeigen sich auch bei der Confounder-Adjustierung die idealtypischen Anforderungen an Studien im Rahmen der AbD. Ob die Vorgaben in der Praxis erfüllbar sind, ist zu hinterfragen. Dabei stellt sich für zahlreiche Fragestellungen die Herausforderung, ähnliche Vergleichskohorten in der Praxis zu generieren, da die Vergleichskohorten absehbar nicht ähnlich genug sein können, um eine akzeptable Confounder-Adjustierung mit statistischen Methoden durchführen zu können. So z.B. wenn die Versorgung sich in einem Ausmaß verändert, so dass die Patienteneigenschaften sich wesentlich unterscheiden. Zugleich stellt das IQWiG hinsichtlich der Annahmen fest: <i>„Falls eine dieser Annahmen mit einem Propensity-Score-Ansatz nicht erreichbar ist, so ist es keine sinnvolle Option z.</i>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>B. stattdessen ein multifaktorielles Regressionsmodell zu verwenden.“ Diese Einschränkung für die alternative Nutzung von multiplen Regressionsanalysen war im IQWiG Rapid Report A19-43 so noch nicht enthalten. Die Einschränkung der methodischen Vorgaben erfolgt dabei ohne nachvollziehbare Begründungen.</i> <i>Die Prüfkriterien für die Confounder-Adjustierung sind teils unklar definiert. So wird z.B. festgehalten: „Weiterhin ist es erforderlich, dass die beiden zu vergleichenden Gruppen bezüglich aller relevanten Confounder ausreichend balanciert sind.“</i> <u>Vorgeschlagene Änderung:</u> Die Confounder-Adjustierung sollte hinsichtlich der Prüfkriterien genauer beschrieben werden und auf die praktischen Aspekte der sich verändernden Versorgung eingehen. Es sollte keine Einschränkung der methodischen Optionen für die Confounder-Adjustierung im Vergleich zu dem vorliegenden Rapid Report A19-43 erfolgen.
Abschnitt 1.4 Gesundheitsökonomie, S. 10 und 4.1 Einführung, S. 83	<u>Anmerkung:</u> Das IQWiG formuliert die Zielsetzung der KNB wie folgt: „Ziel der KNB ist es, ökonomische Informationen als Zusatz zur Nutzenbewertung insbesondere für die Preisverhandlungen im Sinne einer Informationssynthese zusammenzufassen.“ Die KNB ist jedoch kein rechtlich vorgesehener Teil der Erstattungsbetragsverhandlungen und der entsprechenden Rahmenvereinbarung. Diese kann basierend auf § 35b SGB V als Orientierung bei der Preisfindung dienen, wenn die entsprechenden Voraussetzungen erfüllt sind (https://www.g-ba.de/themen/arzneimittel/arzneimittel-richtlinien-anlagen/kosten-nutzen/). <u>Vorgeschlagene Änderung:</u> Der Satz zum Ziel der KNB sollte wie folgt angepasst werden: „Ziel der KNB ist es, nach abgeschlossener früher Nutzenbewertung bei Arzneimitteln mit neuen Wirkstoffen und nach einem Schiedsspruch als ökonomische Orientierung (entsprechend § 35b SGB V)) bei der erneuten Erstattungsbetragsvereinbarung zu dienen.“
Abschnitte 1.4 Gesundheitsökonomie und 4 Kosten-Nutzen-Bewertung medizinischer	<u>Anmerkung:</u> Laut IQWiG ist das Ziel einer KNB nicht die Ableitung von Aussagen über den Zusatznutzen. Die Modellierungen und Extrapolationen können demnach nicht für die Ableitung eines Belegs, Hinweises oder Anhaltspunkts

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
Interventionen, S. 10, S. 83-85, S. 98-99, S. 102	herangezogenen werden (S. 10). Die Modellierungen eignen sich demnach nicht als Basis für einen Nachweis des Zusatznutzens und haben insbesondere prognostische Bedeutung (S. 99). In der KNB geht es laut IQWiG also darum, mögliche Szenarien im Hinblick auf die Kosten abzubilden (S. 10). Zugleich führt das IQWiG aus, dass im Kontext der KNB in der Regel Modellierungen durchgeführt werden, für deren Erstellung Informationen erforderlich sind, die vom Inhalt der vorgelagerten Nutzenbewertung nicht erfasst werden (z. B. Ermittlung von Nutzwerten, ergänzende Daten für die Ermittlung von Übergangswahrscheinlichkeiten, ggf. vorliegende Daten der klinischen Studie mit Relevanz für den Ressourcenverbrauch) (S. 83-84). So dienen laut IQWiG die gesundheitsökonomischen Modelle „<i>explizit auch dazu, Ergebnisse zur Effektivität und zu den Kosten über den studienbelegten Zeitraum hinaus zu extrapolieren</i>“ (S. 98). Dies wird auch im Rahmen des Umgangs mit der Extrapolation der Ergebnisse aus klinischen Studien bei verschiedenen Szenarien beschrieben (S. 102). Insgesamt wird aus den beschriebenen Vorgaben jedoch nicht deutlich, inwieweit im Rahmen einer KNB auch Modellierungen für die Nutzenseite möglich sind und welche Daten, für die hier zu treffenden Annahmen, akzeptiert werden. Das Ziel der KNB sollte nicht sein, die Nutzenbewertung schlicht zu reproduzieren und um weitere Kostenkomponenten zu ergänzen. Zugleichs sollte ein pragmatischer Ansatz für Modellierungen insb. bei besonderen Therapiesituationen angestrebt werden. <u>Vorgeschlagene Änderung:</u> Es sollte klargestellt werden, dass Modellierungen und Extrapolation der Ergebnisse aus klinischen Studien herangezogen werden können, um als Orientierung bei der Preisfindung unter Berücksichtigung der Unsicherheit aus der best verfügbaren Evidenz zu dienen.
Abschnitt 4.3.2 Intervention und Komparator(en), S. 88	<u>Anmerkung:</u> Im Basisfall der KNB erfolgt laut IQWiG „<i>der Vergleich zwischen dem zu bewertenden Arzneimittel und der für den Nachweis des Zusatznutzens herangezogenen zVT. Um die Bandbreite der Therapiekosten zu berücksichtigen, werden im Rahmen von ergänzenden Analysen inkrementelle Kosten-Effektivitäts-Verhältnisse für weitere von der zVT umfasste Therapieoptionen (z. B. die günstigste und die teuerste Therapie) im Sinne einer Kostenvariation berechnet</i>“.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>werden (Szenarioanalyse zVT-Kosten). Dies kann nur erfolgen, sofern die verschiedenen Therapien als gleichermaßen zweckmäßig anzusehen sind und nicht davon auszugehen ist, dass die Krankheitskosten zwischen den Therapien erheblich variieren“.</i> Wenn mehrere Alternativen der zweckmäßigen Vergleichstherapie seitens des G-BA als gleichermaßen zweckmäßig genannt werden, so gilt diese Vorgabe zunächst nur für den Zusatznutzennachweis (der Zusatznutzen kann laut § 6 Abs. 2a AM-NutzenV gegenüber jeder dieser Therapien nachgewiesen werden). Damit ist jedoch nicht automatisch in allen Fallkonstellationen auszugehen, dass die optionalen Therapien ohne Berücksichtigung der verfügbaren Evidenz auf Endpunktebene inkrementell als gleich effektiv anzusehen sind. Eine ODER-Verknüpfung erfolgt, auch in solchen Fällen, wo ein Vertreter der benannten zVT-Wirkstoffe einen (auch beträchtlichen) Zusatznutzen gegenüber einem anderen Wirkstoff aufweist, so z.B. in den Indikationen: • Plaque-Psoriasis, • multiples Myelom, neudiagnostiziert oder mit mindestens einer vorangegangenen Therapie, • Diabetes Mellitus Typ 2, • Brustkrebs • nicht-kleinzelliger Lungenkarzinom. In solchen Fällen kann erst recht nicht davon ausgegangen werden, dass die von der zVT umfassten Therapien in der inkrementellen Betrachtung jeweils gleich effektiv sind. Vielmehr belegen dann die Bewertungsergebnisse des IQWiG und des G-BA, dass die für den Zusatznutzennachweis benannten Optionen auf Endpunktebene nicht gleich effektiv und beliebig austauschbar sind. Auch bei Abwesenheit von vergleichenden Daten zwischen den von der zVT umfassten Therapien kann zwar nach AM-NutzenV der Zusatznutzennachweis gegenüber jeder der vom G-BA vorgegebenen Therapien erfolgen, jedoch ist nicht unter Verzicht auf eine Evidenzbewertung davon auszugehen, dass diese Therapien je Endpunkt gleich effektiv sind. Aus der Vorgehensweise des IQWiG wird zudem nicht ersichtlich, wie im Falle von anderen zweckmäßigen Vergleichstherapien wie einer patientenindividuellen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Therapie oder Therapie nach Maßgabe des Arztes umgegangen werden soll, da hier die Therapien eben nicht als gleichermaßen zweckmäßig anzusehen sind. Ebenso ist unklar, wie das IQWiG mit Fällen umgehen will, wo keine zugelassene Therapieoption besteht. Hinsichtlich der Festlegung von Komparatoren stellt das IQWiG fest: „In Abhängigkeit vom Auftrag können weitere Komparatoren im Rahmen der KNB berücksichtigt werden. Im Fall, dass in der KNB ein Vergleich mit mehreren Vergleichstherapien und nicht nur mit der gewählten zweckmäßigen Vergleichstherapie stattfinden soll, könnte diese Betrachtung nicht auf Grundlage der bereits vorliegenden Nutzenbewertung erfolgen, sondern ggf. ist eine erneute Nutzenbewertung unter Einbeziehung der weiteren Komparatoren erforderlich. Dies kann insbesondere auch eine KNB von Orphan Drugs betreffen, für die im Rahmen des Auftrags für eine KNB ggf. ein Komparator durch den G-BA erst festzulegen ist.“ Nach §35b Abs. 1 SGB V gilt, dass bei einem möglichen Auftrag des G-BA ist insbesondere festzulegen ist, „für welche zweckmäßige Vergleichstherapie“ die Bewertung erfolgen soll. Ein Vergleich zu einem anderen Komparator als Teil der zuvor vom G-BA festgelegten zweckmäßigen Vergleichstherapie ist somit gesetzlich nicht vorgesehen und sollte entsprechend in den Allgemeinen Methoden auch nicht aufgeführt werden. <u>Vorgeschlagene Änderung:</u> Keine ergänzenden Analysen für zVT-Alternativen, für die zuvor nicht auf der Basis des zu untersuchenden Evidenzkörpers je Endpunkt nachgewiesen wurde, dass die von der zVT umfassten Therapien inkrementell als gleich effektiv anzusehen sind. Klarstellung des Umgangs mit Komparatoren, die ausdrücklich nicht als gleichermaßen zweckmäßig anzusehen sind. Klarstellung des Umgangs mit Fällen, wo keine zugelassene Therapieoption besteht. Im Falle einer KNB nach § 35b SGB V muss die Wahl des Komparators oder der Komparatoren vom jeweiligen Auftrag des G-BA unter der Berücksichtigung der gesetzlichen Vorgaben nach § 35b Abs. 1 SGB V erfolgen („In dem Auftrag ist insbesondere festzulegen, für welche zweckmäßige Vergleichstherapie [...] die Bewertung erfolgen

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	soll“). Die Einbeziehung weiterer Komparatoren als die der zweckmäßigen Vergleichstherapie sollte gestrichen werden.
Abschnitt 4.3.3 Gesundheitsökonomische Zielgröße und 4.6 Nutzwerte (im Fall einer Kosten-Nutzwert-Analyse), S. 85, 88-89 und S. 91	<u>Anmerkung:</u> Das IQWiG sieht die Möglichkeit einer Kosten-Nutzwert-Analyse als eine Form der Kosten-Effektivitäts-Analyse vor. Die Kosten sollen in Relation zu qualitätsadjustierten Lebensjahren (QALYs) gesetzt werden, wobei die Bewertung von Gesundheitszuständen in Form von Nutzwerten mit der Restlebenserwartung oder Verweildauer pro Gesundheitszustand kombiniert werden. Die Nutzwerte können dabei auch durch indirekte Verfahren z. B. basierend auf dem EQ-5D, SF-6D sowie HUI 2 oder 3 ermittelt werden. Das IQWiG stellt fest, dass in der Literatur verschiedene Limitationen des QALY-Konzepts diskutiert werden, insbesondere ethische, methodische sowie konzeptionelle Einschränkungen. Dabei wird auf verschiedene Literaturquellen verwiesen. Zugleich merkt das IQWiG an, dass das QALY-Konzept dennoch die derzeit am umfassendsten angewandte Methodik darstelle, „die explizit und mit einem theoretisch begründeten Konstrukt die Bewertung der Effektivität von Therapien in gesundheitsökonomischen Evaluationen unter Berücksichtigung von Präferenzen ermöglicht“. Die erwähnten ethischen (z.B. Akzeptanz und Verwendung bei seltenen Erkrankungen) und methodischen Limitationen (z.B. Änderungssensitivität bei der Nutzwertenerhebung) werden jedoch nicht erwähnt, dies obwohl sie von hoher Relevanz für die theoretische Anwendung in Deutschland wären und zu erheblichen Verzerrungen bei der Informationsbereitstellung für Preisverhandlungen bzw. damit zu Fehlallokationen bei daraus folgenden Erstattungen von Arzneimitteln führen können. Auf S. 85 gibt das IQWiG vor: „Nutzwerte basierend auf Bewertungen durch die Allgemeinbevölkerung sind als dann sinnvoll, wenn sich die Bewertungen nicht von denen durch Patientinnen und Patienten unterscheiden“. Dem wäre zu entnehmen, dass Nutzwerte basierend auf Bewertungen durch die Allgemeinbevölkerung nicht akzeptiert werden, wenn Unterschiede vorliegen. Dabei werden die Unterschiede nicht weiter präzisiert. Zugleich wird außer Acht gelassen, dass solche Fallkonstellationen in der Praxis kaum realisierbar wären. Zugleich wird auf S. 92 angegeben: „Wenn davon auszugehen ist, dass Unterschiede existieren, besteht für die Ergebnisse von Analysen auf Grundlage von

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<i>Bewertungen durch die Allgemeinbevölkerung die Unsicherheit, dass die Ergebnisse auf Basis von Bewertungen durch Patientinnen und Patienten anders ausfallen könnten. Dies ist bei der Einordnung der Ergebnisse der KNB zu berücksichtigen.“ Hieraus ist abzuleiten, dass auch Analysen auf Grundlage von Bewertungen durch die Allgemeinbevölkerung grundsätzlich akzeptiert werden.</i> <u>Vorgeschlagene Änderung:</u> Klarstellung zur Relevanz der Instrumente EQ-5D, SF-6D sowie HUI 2 oder 3 auch für die Nutzenbewertung nach §35a SGB V. Die ethischen, methodischen sowie konzeptionellen Limitationen des QALY-Konzepts und auch eine mögliche praktische Umsetzbarkeit in Deutschland sollten im Methodenpapier detaillierter diskutiert und dargestellt werden. Die Verwendung der Nutzwerte basierend auf Bewertungen durch die Allgemeinbevölkerung sollte präziser beschrieben werden.
Abschnitt 4.3.4 Perspektive, S. 89	<u>Anmerkung:</u> Im Referenzfall soll die KNB aus der Perspektive der GKV-Versichertengemeinschaft durchgeführt werden. In Abhängigkeit vom Auftrag kann ergänzend auch die (reine) GKV-Perspektive, die Sozialversicherungsperspektive bzw. die Perspektive einzelner Sozialversicherungsträger sowie die gesellschaftliche Perspektive Berücksichtigung finden. <u>Vorgeschlagene Änderung:</u> Die Wahl der Perspektive sollte flexibel ausgewählt werden und je nach Datenlage, Besonderheiten der Erkrankung und Auftrag als Referenzfall angepasst werden können.
Abschnitt 4.5 Effektivität sowie weitere klinische Inputparameter und 4.6 Nutzwerte (im Fall einer Kosten-Nutzwert-Analyse), S. 90-92	<u>Anmerkung:</u> Die Kosten für die Angehörigen können nach Abschnitt 4.7 erfasst werden. Für die Bewertung der Effektivität bzw. Nutzwerte ist diese Option nicht aufgeführt. <u>Vorgeschlagene Änderung:</u> Die Perspektive der Angehörigen sollte bei den Auswirkungen von Interventionen hinsichtlich der Parameter für den Nutzen ebenfalls berücksichtigt werden können. Dies

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	sollte in den Abschnitten 4.5 und 4.6 entsprechend ergänzt werden.
Abschnitt 4.7.2 Schritte zur Kostenbestimmung, S. 96	<u>Anmerkung:</u> Das IQWiG formuliert für die Bewertung der Ressourcen wie folgt: „Für die Bewertung der Kostenkomponenten [...] je nach Kostenbereich, mit regulierten Preisen (z. B. <i>Diagnosis-related Groups</i> für den stationären Sektor) oder verhandelten Preisen (Arzneimittelpreise).“ Bei der hier dargestellten Kostenbetrachtung bleibt unberücksichtigt, dass nicht alle Arzneimittel über einen verhandelten Erstattungsbetrag verfügen. <u>Vorgeschlagene Änderung:</u> Der Satz sollte entsprechend angepasst werden: „... oder beispielsweise verhandelten Preisen (Arzneimittelpreise).“
Abschnitt 4.8 Inflationsbereinigung und Diskontierung, S. 98	<u>Anmerkung:</u> Bei Perioden > 1 Jahr, will das IQWiG nach dem 1. Jahr im Basisfall mit einer konstanten Rate von 3 % pro Jahr auf die aktuelle Periode diskontieren. Eine konstante Rate von 3% erscheint vor dem Hintergrund der sich verändernden Rahmenbedingungen, der verschiedenen Angaben in der Literatur und Besonderheiten von Therapiesituationen nicht adäquat. <u>Vorgeschlagene Änderung:</u> Die Wahl der jeweiligen Inflationsbereinigung und Diskontierungsraten sollte je nach Situationen begründet werden.
Abschnitt 4.9 Gesundheitsökonomisches Studiendesign, S. 102	<u>Anmerkung:</u> Für die Extrapolation soll in Szenarien u.a. angenommen werden, „dass der in den klinischen Studien beobachtete Effekt/Unterschied zwischen den betrachteten Therapien über das Ende der klinischen Studie hinaus anhält und danach abfällt oder endet.“ Unberücksichtigt bleibt dabei die Möglichkeit einer Zunahme des Effekts. <u>Vorgeschlagene Änderung:</u> Ergänzung des Satzes „dass der in den klinischen Studien beobachtete Effekt/Unterschied zwischen den betrachteten Therapien über das Ende der klinischen Studie hinaus sich verstärkt, anhält, danach abfällt oder endet.“
Abschnitt 4.10 Ergebnisdarstellung und	<u>Anmerkung:</u>

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
4.12 Einordnung des ICERs, S. 103-104 und S. 107	Der Abschnitt 4.6 „Ergebnisdarstellung als Effizienzgrenze“ entfällt im neuen Entwurf der Methoden. Anstatt einer Effizienzgrenze soll nun die Berechnung des ICER zwischen den betrachteten Therapiealternativen das wesentliche Ergebnis darstellen und tabellarisch bzw. grafisch dargestellt werden. Die dabei erfolgte Abkehr vom Konzept der Effizienzgrenze ist grundsätzlich nachvollziehbar. Zudem stellt das IQWiG fest: <i>„Die von der zVT umfassten Therapieoptionen können unterschiedlich hohe Therapiekosten aufweisen. Um diese Bandbreite zu berücksichtigen, muss daher nicht nur das ICER für den Vergleich zwischen dem zu bewertenden Arzneimittel und der zum Nachweis des Zusatznutzens ausgewählten Vergleichstherapie berechnet werden, sondern auch für die Vergleiche zwischen dem zu bewertenden Arzneimittel und den anderen von der zVT umfassten Therapien (Szenarioanalyse zVT-Kosten). Diese ergänzenden Analysen erfolgen auf Grundlage der Therapiekosten der anderen von der zVT umfassten Therapien (z. B die günstigste und die teuerste Therapie). Diese Szenarioanalyse zVT-Kosten findet unter den Annahmen statt, dass die inkrementelle Effektivität des zu bewertenden Arzneimittel im Vergleich zu den verschiedenen von der zVT umfassten Therapien jeweils gleich ist und die Krankheitskosten nicht erheblich voneinander abweichen.“</i> Die Feststellung, dass die inkrementelle Effektivität des zu bewertenden Arzneimittels im Vergleich zu den verschiedenen von der zVT umfassten Therapien jeweils gleich ist, kann (wie oben bereits erläutert) nicht ohne explizit zu bewertender Evidenzgrundlage auf der Endpunktebene getroffen werden. Wie bereits ausgeführt und mit Beispielen aufgezeigt, ist nicht automatisch davon auszugehen, dass die als zVT für den Zusatznutzennachweis vorgegebenen Therapiealternativen auch als inkrementell gleich effektiv anzusehen sind. Eine ODER-Verknüpfung erfolgt seitens des G-BA teils sogar in solchen Fällen, obwohl ein Vertreter der benannten Wirkstoffe einen Zusatznutzen (auch beträchtlichen Zusatznutzen) gegenüber dem anderen aufweist. In solchen Fällen kann nicht davon ausgegangen werden, dass mehrere Alternativen der Vergleichstherapie gleich effektiv sind bzw. die inkrementelle Effektivität je Endpunkt des zu bewertenden Arzneimittel im Vergleich zu den verschiedenen von der zVT umfassten Therapien jeweils gleich ist. Vielmehr zeigen die Bewertungsergebnisse des IQWiG und des G-BA, dass die für den Zusatznutzennachweis benannten Optionen auch

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	unterschiedliche Effektivität aufweisen können und damit in der inkrementellen Betrachtung nicht gleichgestellt werden dürfen. <u>Vorgeschlagene Änderung:</u> Es sollte klargestellt werden, dass keine Szenarioanalyse für zVT-Kosten erfolgt, wenn nicht auf der Grundlage der nach Standards der Evidenz-basierten Medizin zu untersuchenden Studiengrundlage bewertet wurde, dass die Komparatoren auf der jeweiligen Endpunktebene inkrementell als gleich effektiv zu bewerten sind.
Abschnitt 9.1.1 Kriterien für den Einschluss von Studien, S. 159	<u>Anmerkung:</u> Das IQWiG fasst die Kriterien für den Einschluss von Studien neu. Liegen demnach Analysen für relevante Teilpopulationen von Studien nicht vor, wird die Studie dann eingeschlossen, falls auf mindestens 80 % die Einschlusskriterien für die Population, Prüfintervention bzw. Vergleichsintervention gelten. Falls geeignete Informationen über eine Effektmodifikation durch das betreffende Einschlusskriterium (Population bzw. Interventionen) vorliegen, muss über die Berücksichtigung der Studie anhand der Stärke der Effektmodifikation und dem Anteil der Patientinnen und Patienten, die das Einschlusskriterium nicht erfüllen bzw. dem Grad der Abweichung der Interventionen entschieden werden. Aus den neuen Vorgaben wird insg. nicht ersichtlich, ob (wie aktuell in den Aktuellen Methoden 6.1) auch Studien, bei denen das Einschlusskriterium bei weniger als 80 % erfüllt ist, eingeschlossen werden können, wenn Analysen der relevanten Teilpopulation vorliegen oder wenn hinreichend sicher plausibel oder nachgewiesen ist, dass die in dieser Studie gewonnenen Ergebnisse auf die Zielpopulation der Fragestellung anwendbar sind. Diese Vorgabe sollte auch weiterhin beibehalten werden, um die Beurteilung hinsichtlich der Berücksichtigung von Studien nicht allein vom arbiträren Schwellenwert und ohne Berücksichtigung von zufälligen Effekten für einzelne Endpunkte abhängig zu machen. Zugleich ist anzuregen, die Ergebnisse des IQWiG-Arbeitspapiers „Untersuchung der statistischen Eigenschaften von Verfahren zur Übertragbarkeit von Studienergebnissen auf Teilpopulationen“ im Methodenpapier zu berücksichtigen. Die Problematik des Power-Verlusts durch die Aufteilung und Berücksichtigung von Teilpopulationen einer Studie bleibt anderenfalls weiterhin gänzlich ohne Berücksichtigung.

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	<u>Vorgeschlagene Änderung:</u> Es sollte präzisiert werden, dass auch Studien, bei denen das Einschlusskriterium bei weniger als 80 % erfüllt ist, eingeschlossen werden können, wenn Analysen der relevanten Teilpopulation vorliegen oder wenn hinreichend nachgewiesen ist, dass die in dieser Studie gewonnenen Ergebnisse auf die Zielpopulation der Fragestellung anwendbar sind. Es sollte zudem klargestellt werden, dass bei mind. 80 % der Erfüllung der Einschlusskriterien kein Ausschluss der Studien auf der Grundlage von Subgruppenanalysen erfolgt. Die Problematik des Power-Verlusts durch die Aufteilung und Berücksichtigung von Teilpopulationen einer Studie sollte im Methodenpapier erwähnt und mit einer möglichen Vorgehensweise kommentiert werden.
Abschnitt 9.3.2 zu statistischer Signifikanz und Konfidenzintervallen, S. 170	<u>Anmerkung:</u> In den Anpassungen stellt das IQWiG fest, dass das Problem des multiplen Testens in systematischen Übersichten nicht formal (zuvor „nicht komplett“) lösbar ist, dennoch zumindest bei der Interpretation der Ergebnisse beachtet werden sollte. Wenn sinnvoll und möglich, wendet das IQWiG dabei „Methoden zur Adjustierung für multiples Testen an“. Der folgende präzisierende Satz entfällt dabei „Im Rahmen der Nutzenbewertung (siehe Abschnitt 3.1) versucht das Institut, den Fehler 1. Art bezüglich der Aussagen zu jedem einzelnen Nutzenaspekt getrennt zu kontrollieren.“ Stattdessen hält das IQWiG fest, dass die Methoden „z. B. für die Bewertung von Ergebnissen zu einem Endpunkt, der mithilfe verschiedener Operationalisierungen untersucht worden ist“ angewandt werden können. Es ist zuzustimmen, dass für den Kontext der Fragestellungen der Nutzenbewertung (in diesem Fall von Arzneimitteln) ein formales Vorgehen zum Umgang mit dem multiplen Testen inhaltlich nicht sinnvoll ist. Ebenso sollte, wie vom IQWiG weiterhin vorgeschlagen, eine zusammenfassende Bewertung über alle Endpunkte hinweg nicht auf quantitative Weise durchgeführt und auch keine formalen Methoden zur Adjustierung für multiples Testen angewendet werden. Bei der Bewertung von Endpunkten, die ggf. mithilfe verschiedener Operationalisierungen untersucht worden sind, sollte der Fokus hingegen auf die Aussagekraft und Eignung der jeweiligen Operationalisierung oder deren ergänzenden Charakter gelegt werden. Eine Anwendung von Methoden zur Adjustierung für multiples Testen bei bloßem

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Vorliegen von mehreren Operationalisierungen ist dabei nicht sinnvoll. <u>Vorgeschlagene Änderung:</u> Es sollte präzisiert werden, dass Methoden zur Adjustierung für multiples Testen nicht automatisch beim Vorliegen von mehreren Operationalisierungen eines Endpunktes angewandt werden sollen. Zudem sollte klargestellt werden, dass die Darstellung von Sensitivitätsanalysen nicht zur Adjustierung für multiples Testen führen. Das Vorgehen einer möglichen Änderung der üblichen Signifikanzschwelle sollte genauer beschrieben und an Beispielen präzisiert werden.
Abschnitt 9.3.12 zur Auswirkung systematisch verkürzter Beobachtungszeiten bei der Erhebung patientenberichteter Endpunkte (PROs)	<u>Anmerkung:</u> Das IQWiG ergänzt die Beschreibung zum Umgang mit variablen Beobachtungszeiten. Insb. wird dabei auf die Situationen mit systematisch verkürzter Datenerhebung eingegangen. Dabei sind laut IQWiG sogenannte „dauerhafte“ Veränderungen bei sich deutlich zwischen den Behandlungsarmen unterscheidenden Beobachtungsdauern nicht sinnvoll interpretierbar. Darüber hinaus sei laut IQWiG auch nicht sachgerecht, von einer „dauerhaften“ Veränderung zu sprechen, wenn die Beobachtungszeit nur einen verkürzten und ggf. sehr kleinen Anteil der gesamten Studiendauer abdeckt. Auswertungen als „dauerhafte“ Veränderungen sollen daher im Falle einer verkürzten oder sich zwischen den Behandlungsarmen deutlich unterscheidenden Beobachtungsdauer in der Regel nicht für die Nutzenbewertung herangezogen werden. Neu ist zudem, dass „im Falle systematisch verkürzter Beobachtungsdauer die Ergebnisse davon betroffener Endpunkte (z. B. unerwünschte Ereignisse oder PROs) die Gesamtaussage zum Nutzen oder Zusatznutzen einer medizinischen Intervention in der Regel nicht dominieren“ können. Aus methodischer Sicht ist grundsätzlich nachvollziehbar, dass bei relevant unterschiedlichen Beobachtungszeiten zwischen den Studienarmen die Auswertungen als Zeit bis zur bestätigten oder dauerhaften Veränderung (Verbesserung und/oder Verschlechterung) die Interpretierbarkeit der Ergebnisse verzerrt sein kann. In solchen Fällen kann eine Responderanalyse als Zeit bis zur erstmaligen Veränderung eine besser geeignete

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	Operationalisierung darstellen oder die Analysen für eine dauerhafte Veränderung ergänzen. Beim Umgang mit unterschiedlichen Beobachtungszeiten hat inzwischen der G-BA auf seiner Webseite (Antworten auf häufig gestellte Fragen zum Verfahren der Nutzenbewertung) pragmatische Hinweise zu diesem Thema dargestellt, die einen sinnvolleren Umgang als der Vorschlag des IQWiG darstellen. Liegen demnach zu einem Endpunkt keine relevant unterschiedlichen Beobachtungszeiten zwischen den Studienarmen vor, ist in der Regel die Auswertung der Zeit bis zu einer bestätigten oder dauerhaften Veränderung grundsätzlich inhaltlich sinnvoll. Der pharmazeutische Unternehmer habe im Dossier zu begründen, wieso eine Auswertung zur Zeit bis zur erstmaligen, Zeit bis zur bestätigten oder Zeit bis zur dauerhaften Veränderung inhalts- und kontextbezogen geeignet ist. Als nicht sinnvoll wird hingegen der Vorschlag des IQWiG erachtet, dass bei sog. „<i>systematisch verkürzter Beobachtungsdauer</i>“ die Ergebnisse betroffener Endpunkte die Gesamtaussage „<i>nicht dominieren</i>“ könnten. Damit werden die patientenberichteten Endpunkte regelhaft ohne Berücksichtigung der jeweiligen Therapiesituation im Bewertungsschema degradiert. Die Beobachtungsdauer der PROs in den klinischen Studien ist jedoch stets kontextabhängig. Eine Erfassung für die gesamte Studiendauer bzw. über die gesamte Behandlungsstrategie unabhängig von einem Behandlungsabbruch oder -wechsel ist weder praktikabel noch im Sinne der eigentlichen Fragestellung der Nutzenbewertung und sinnvoll interpretierbarer Ergebnisse. Ebenso ist vor allem bei onkologischen Erkrankungen eine zu langwierige PRO-Erhebung nach Progress fragwürdig, auch vor dem Hintergrund der Anforderungen an die Zulassungsstudien, wo eine klare und hypothesengestützte Begründung für die PRO-Erhebung gefordert wird bzw. eine hohe Rücklaufquote zu erwarten ist [1]. Eine Einstufung der Beobachtungsdauer als systematisch oder nicht systematisch verkürzt kann daher nicht sinnvoll operationalisiert werden. Zugleich bleibt bei dem vom IQWiG angestrebten Vorgehen einer automatischen Degradierung der PROs unberücksichtigt, dass gerade längere Beobachtungsdauern zunehmend mit einem deutlich höheren Verzerrungspotenzial einhergehen können, insb. aufgrund

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
	sinkender Rücklaufquoten (auf unter 70%) oder beim erlaubten Therapiewechsel. <u>Vorgeschlagene Änderung:</u> Der Begriff einer „systematisch verkürzten Beobachtungsdauer“ ist in diesem Kontext irreführend und sollte nicht verwendet werden. Es sollte klargestellt werden, dass so lange zu einem Endpunkt keine relevant unterschiedlichen Beobachtungszeiten zwischen den Studienarmen vorliegen, die Operationalisierung als Zeit bis zu einer bestätigten oder dauerhaften Veränderung inhaltlich sinnvoll sein kann. Die Vorgabe in der Version 6.1 („Liegen bei den zu vergleichenden Gruppen unterschiedlich lange Beobachtungsdauern vor, ohne dass dies in der Datenanalyse adäquat berücksichtigt worden ist, so kann dies dazu führen, dass die entsprechenden Ergebnisse für die Nutzenbewertung nicht verwertbar sind“) deckt dies bereits hinreichend ab und sollte nicht verändert werden. Die Beobachtungsdauer sollte nicht als Kriterium zu einer automatischen Abwertung des Stellenwerts von patientenberichteten Endpunkten bei der Gesamtaussage zum Nutzen oder Zusatznutzen führen.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

1. Committee for Medicinal Products for Human Use (CHMP): Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies, EMA/CHMP/292464/2014, 1 April 2016.
- 2.

A.2 Stellungnahmen von Privatpersonen

A.2.1 Krones, Tanja

Stellungnahme zum Entwurf der Allgemeinen Methoden

Diese und die folgenden Seiten dieses Dokuments oder Ihre selbst erstellten und beigefügten Anlagen werden in dem Dokument „Dokumentation und Würdigung der Anhörung zum Entwurf der Allgemeinen Methoden Version 7.0“ auf der Internetseite des IQWiG veröffentlicht.

Name, Vorname; Titel des/der Stellungnehmenden <i>Bitte nutzen Sie pro Person 1 Zeile.</i>
Krones, Tanja; Prof. Dr. med. Dipl. Soz.
Die Abgabe der Stellungnahme erfolgt (bitte ankreuzen)
☐ im Namen folgender Institution / Organisation:
☒ als Privatperson(en)

Die folgende Vorlage für eine Stellungnahme kann beliebig durch eigene Anlagen ergänzt oder ersetzt werden. Diese unterliegen denselben Vorgaben und der Veröffentlichung, wie oben dargestellt.

Stellungnahme zu allgemeinen Aspekten (optional)

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

Gern nehme ich zu dem Entwurf des Methodenpapiers 7.0 als Ethikerin, die auch im Bereich der evidenzbasierten Medizin und HTA arbeitet und forscht, Stellung. International wurde die Integration ethischer Aspekte in HTA allgemein in den letzten 10 Jahren weiter vorangetrieben. Weitgehender Konsens besteht darin, dass ethische Aspekte in allen Schritten des HTA-Prozesses, von der Priorisierung der Forschungsfrage, über das «Core Assessment» von Wirksamkeit und Wirtschaftlichkeit bis hin zum Appraisal der Ergebnisse mit Empfehlung an die Politik eine zentrale Rolle spielen. Dies wird sehr eindrücklich im aktuell gültigen HTA Core Model der EUNETHTA (1) ausgeführt. Beispielsweise in Norwegen, Schweden und der Schweiz werden die ethischen Aspekte von Gesundheitstechnologien, inklusive dem Assessment von Medikamenten, mittlerweile systematisch von Beginn an berücksichtigt:

Siehe

Norwegen

Norwegian agency (National Institute for Public Health)

<https://www.fhi.no/nettpub/metodeboka/metodevurdering/etiske-aspekter/>

Schweden

SBU <https://www.sbu.se/en/>; sowie (4) und (7)

Schweiz (1)

Hierbei werden meist sowohl material-inhaltliche als auch prozedurale Methoden und Prozesse definiert, um auch im Rahmen der Wirksamkeits- und Wirtschaftlichkeitsprüfung ethische Aspekte in den Bewertungen zu berücksichtigen. Verwendet werden hierbei meist zum einen prinzipienethische Grundsätze (meist mit Referenz zu dem Ansatz von Tom Beauchamp und James Childress «Principles of Biomedical Ethics», (8. Auflage, zuletzt 2019) und zum anderen der «socratic approach», der speziell für HTA entwickelt und weiterentwickelt wurde. (5-6)

Hierbei hat auch eine ehemalige Mitarbeiterin des IQWiG, Sigrid Droste, massgeblich beigetragen, die zudem auch ein wesentliches Methodenpapier zur systematischen Recherche ethischer Aspekte publiziert hat, welches in den oben genannten HTA Organisationen teils als Goldstandard referenziert wird (2).

Auch im aktuellen Entwurf des Methodenpapiers findet sich ein Prozess zur Integration ethischer Aspekte. Dies jedoch allein in der Bewertung der ThemenCheckberichte (6.5.). Unter 6.5.3. wird ausgeführt:

«In der Medizin, aber auch in der Public-Health- und Versorgungsforschung sind für eine ethische Bewertung die 4 Prinzipien des Wohltuns und Nutzens, des Nichtschadens, des Respekts vor der Autonomie der Patientin oder des Patienten und der Gerechtigkeit nach Beauchamp und Childress [48] weitverbreitet [214,326]. Für die Analyse ethischer Aspekte von medizinischen Interventionen wurde dieser Ansatz ebenfalls häufig verwendet [47,507]. In den vergangenen Jahren wurden auch weitere methodische Ansätze eingesetzt. Dazu zählen z. B. der sokratische Ansatz, das Social Shaping of Technology, die Virtue Ethics oder auch der trianguläre Ansatz [28,368]. Je nach den zu bearbeitenden Fragestellungen eignen sich die verfügbaren methodischen Konzepte in unterschiedlicher Weise.

Der sokratische Ansatz liefert detaillierte Ergebnisse und wurde bei vielen unterschiedlichen medizinischen Interventionen eingesetzt [207,367,496]. Der Fragenkatalog von Hofmann 2005 [365,366] basiert auf dem sokratischen Ansatz und soll regelhaft für die ThemenCheck-Berichte

Sofern Sie allgemeine Anmerkungen zum Entwurf der Allgemeinen Methoden haben, tragen Sie diese bitte hier ein. Spezifische Aspekte adressieren Sie bitte in der folgenden Tabelle.

eingesetzt werden. Bei besserer Eignung anderer methodischer Ansätze können bei entsprechender Begründung im ThemenCheck-Berichtsprotokoll auch diese eingesetzt werden.»

Ich möchte hiermit beantragen, dass dieser Ansatz, analog zu internationalen Entwicklungen in HTA, auch regelhaft für die Nutzenbewertung medizinischer Verfahren, sowohl bei medikamentösen wie nicht-medikamentösen Verfahren angewendet wird. In der internationalen Debatte ist mittlerweile unstrittig, dass ethische Aspekte nicht allein im Appraisal, sondern bereits im Assessment berücksichtigt werden müssen, um die immer einflussenden Wertungen in der Faktenanalyse angemessen zu evaluieren und die Analyse damit noch valider und transparenter zu machen. Die systematische Berücksichtigung und Berichterstattung zu den ethischen Aspekten im Assessmentprozess ist meines Erachtens auch durch den Prüfauftrag des IQWiG zuhanden des G-BA gedeckt, da, wie oben beschrieben und zum Beispiel durch EUNETHTA ausführlich begründet, die Nutzenbewertung und Wirtschaftlichkeitsprüfung wesentlich in ihrer Aussage verbessert und so im Appraisal und den politischen Entscheidungen einen noch wertvolleren Beitrag leisten können.

Denkbar hierbei ist ein mehrstufiges Verfahren mit initialem Screening, ob im Rahmen der Wirksamkeits- oder Wirtschaftlichkeitsprüfung bei medikamentösen oder nicht-medikamentösen Verfahren ethisch relevante Aspekte besonders ins Auge fallen (z.B. weil die Patient:innenpopulation besonders vulnerabel ist, relevante Evidenzlücken bestehen, oder sich Gerechtigkeitsfragen durch die Wirtschaftlichkeitsanalyse stellen) und, falls ja, einer vertieften ethischen Analyse auf der Basis des im Methodenpapier unter 6.5.3. unter Berücksichtigung der Methode von Droste et al (2010) zur systematischen Literaturrecherche durchzuführen. So schlägt die SBU (2021, ohne Seitenangabe) beispielsweise vor

«Summary -to consider:

- Consider the ethics as early as during the pilot study.
- Early in the project, discuss the ethical aspects that have been spontaneously identified, i.e., *before* the guideline is incorporated into the project.
- Different ethical aspects can be of varied importance in different projects.
- Ethical expertise may be needed to a different extent depending on the complexity of ethical issues.
- If an in-depth ethical analysis is carried out according to a specific analytical model (e.g., the actor model), this should be made clear. Regardless of the model, the analysis should apply to Swedish context and regulations and be limited to the specific question.»

Ich bin überzeugt davon, dass diese -grundlegende- Änderung die Qualität und Nutzbarkeit der Berichterstattung und damit dem gesamten Prozess der Nutzenbewertung und Prüfung der Refinanzierung durch die gesetzlichen Krankenkassen wesentlich verbessern würde.

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Stellungnahme zu spezifischen Aspekten (optional)

Kapitel/ Abschnitt (Seite) im Entwurf der Allgemeinen Methoden	Stellungnahme mit Begründung sowie vorgeschlagene Änderung <i>Die Begründung sollte durch entsprechende Literatur belegt werden. Die Literaturstellen, die zitiert werden, müssen eindeutig benannt und im Anhang im Volltext beigefügt werden.</i>
z. B. Abschnitt 3.4 (S.16)	<u>Anmerkung:</u> <u>Vorgeschlagene Änderung:</u>
	<u>Anmerkung:</u> <u>Vorgeschlagene Änderung:</u>

(Bitte fügen Sie weitere Zeilen an, falls dies notwendig sein sollte.)

Literaturverzeichnis

Bitte verwenden Sie zur Auflistung der zitierten Literatur möglichst eine nummerierte Referenzliste und behalten Sie diese Nummerierung bei der Benennung der Dateien (Volltexte) bei.

- 1) BAG (2022) Operationalisierung der WZW-Kriterien vom 31.03.2022, g Itig ab 01.09.2022.pdf
- 2) Droste S, Dintsios CM, Gerber A. Information on ethical issues health technology assessment. How and where to find them. Int J Technol Assess Health Care 2010; 26: 441-9.
- 3) EUNETHTA <https://www.eunetha.eu/hta-core-model/>
- 4) Heintz E, L Lintamo, Hultcrantz M, Jacobson S, R Levi, Munthe, et al. Framework for systematic identification of ethical aspects of healthcare technologies: the SBU approach. Int J Technol Assess Health Care 2015; 31 (3): 124-130.
- 5) Hofmann B, Droste S, Oortwijn W, Cleemput I, Sacchini D. Harmonization of ethics in health technology assessment: a revision of the Socratic approach. Int J Technol Assess Health Care. 2014 Jan;30(1):3-9. doi: 10.1017/S0266462313000688. PMID: 24499630.
- 6) Hofmann, B. Toward a procedure for integrating moral issues in health technology assessment. Int J Technol Assess Health Care. 2005;21:312–318.
- 7) SBU ethical-aspects-of-health-care-interventions_a-guideline_final-003sa.pdf

