

Band 8

Frühe Nutzenbewertung neuer Arzneimittel in Deutschland

Subgruppen

Definition, Analyse und Kriterienkatalog

*In Kooperation mit der Ad-Hoc-Kommission Nutzenbewertung der
Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)*

Frühe Nutzenbewertung neuer Arzneimittel in Deutschland

Subgruppen

Definition, Analyse und Kriterienkatalog

*In Kooperation mit der Ad-Hoc-Kommission Nutzenbewertung der
Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)*

Gesundheitspolitische Schriftenreihe der DGHO
Band 8

Frühe Nutzenbewertung neuer Arzneimittel in Deutschland

Subgruppen

Definition, Analyse und Kriterienkatalog

In Kooperation mit der Ad-Hoc-Kommission Nutzenbewertung der
Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)

Stand: April 2016
ISBN 978-3-9816354-9-2

Herausgeber:

Prof. Dr. med. Carsten Bokemeyer
Prof. Dr. med. Michael Hallek
Prof. Dr. med. Diana Lüftner
Prof. Dr. med. Florian Weißinger

Vorstand der DGHO Deutsche Gesellschaft
für Hämatologie und Medizinische Onkologie
Alexanderplatz 1
10178 Berlin
www.dgho.de
info@dgho.de

Autor:

Prof. Dr. med. Bernhard Wörmann

Titelgestaltung und Satz: unicom Werbeagentur GmbH, Berlin
Druck: copy print Kopie & Druck GmbH, Berlin

Vorwort

Die frühe Nutzenbewertung neu zugelassener Arzneimittel ist ein sinnvolles und funktionierendes Instrument zur Schaffung einer Basis für Preisverhandlungen zwischen Kostenträgern und pharmazeutischen Unternehmen. Inzwischen werden die Festlegungen allerdings auch aus verschiedenen Richtungen für Therapieempfehlungen und die Beeinflussung der Versorgung genutzt.

Dabei haben Subgruppen eine besondere Bedeutung, da sie in vielen Verfahren separat bewertet werden. Im Gesetz zur Neuordnung des Arzneimittelmarktes von 2010 (AMNOG) ist die Identifizierung der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, ausdrücklich vorgesehen. Eine einheitliche Methodik gibt es allerdings nicht, entsprechend ist die Definition und Bewertung distinkter Subgruppen ein häufiger Kritikpunkt in Stellungnahmen.

Die DGHO Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie und die neu gegründete Ad-Hoc-Kommission der Arbeitsgemeinschaft wissenschaftlicher medizinischer Fachgesellschaften (AWMF) haben die Definition und Analyse von Subgruppen als wichtiges Thema für die zukünftige Akzeptanz und Relevanz der frühen Nutzenbewertung identifiziert. Das daraus resultierende Positionspapier benennt zukünftige Arbeitsfelder für die beteiligten Institutionen und Gemeinschaften. Dahinter steht auch das Bemühen, kritische Divergenzen in den parallelen Prozessen von Nutzenbewertung und Leitlinienerstellung zu vermeiden.

Wir laden alle zur konstruktiven Diskussion ein.

Für die DGHO



Prof. Dr. med.
Carsten Bokemeyer
Geschäftsführender
Vorsitzender



Prof. Dr. med.
Michael Hallek
Vorsitzender



Prof. Dr. med.
Diana Lüftner
Mitglied im Vorstand



Prof. Dr. med.
Florian Weißinger
Mitglied im Vorstand

Für die AWMF



Prof. Dr. med.
Rolf Kreienberg
Präsident

Aus dem vorliegenden Band ist ein Positionspapier zur Definition und Analyse von Subgruppen im Rahmen der frühen Nutzenbewertung vom Vorstand der AWMF und dem Vorstand der DGHO verabschiedet worden. Dieses liegt dem Band bei und ist auf der Website der DGHO veröffentlicht.

Beteiligte Fachgesellschaften und Experten

Fachgesellschaft / Institution	Abkürzung	Experte
Arbeitsgemeinschaft der wissenschaftlichen medizinischen Fachgesellschaften	AWMF	Dr. M Nothacker Prof. Dr. H. K. Selbmann
Deutsche Diabetes Gesellschaft	DDG	Prof. Dr. D. Müller-Wieland
Deutsche Gesellschaft für Endokrinologie	DGE	PD Dr. J. Feldkamp
Deutsche Gesellschaft für Epileptologie	DGE	Prof. Dr. H. Hamer
Deutsche Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten (DGVS)	DGVS	Prof. Dr. S. Zeuzem
Deutsche Gesellschaft für Gesundheitsökonomie	DGGÖ	Prof. Dr. R. Busse Prof. Dr. W. Greiner
Deutsche Gesellschaft für Gynäkologie und Geburtshilfe	DGGG	Prof. Dr. M. Beckmann Prof. Dr. E. F. Solomayer
Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie	DGHO	Prof. Dr. B. Wörmann
Deutsche Gesellschaft für Infektiologie	DGI	Prof. Dr. O. Witzke
Deutsche Gesellschaft für Innere Medizin	DGIM	Prof. Dr. T. Sauerbruch
Deutsche Gesellschaft für Kardiologie, Herz- und Kreislaufforschung	DGK	Prof. Dr. B. Nowak
Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie	GMDS	Prof. Dr. U. Mansmann
Deutsche Gesellschaft für Neurologie	DGN	Prof. Dr. H. Hamer
Deutsche Gesellschaft für Pneumologie und Beatmungsmedizin	DGP	Prof. Dr. P. Herth
Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde	DGPPN	Prof. Dr. G. Gründer
Deutsche Gesellschaft für Rheumatologie	DGRh	Prof. Dr. K. Krüger
Deutsche Gesellschaft für Urologie	DGU	Prof. Dr. O. Hakenberg
Deutsche Krebsgesellschaft	DKG	Dr. J. Bruns Prof. Dr. W. Schmiegel
Deutsche Ophthalmologische Gesellschaft	DOG	Prof. Dr. F. Ziemssen

Inhaltsverzeichnis

1. Zusammenfassung	6
2. Einleitung	8
3. Subgruppenbildung in abgeschlossenen Verfahren nach dem AMNOG 2011 – 2015	10
1. Häufigkeit	10
2. Kriterien	11
3. Ergebnisse	11
4. Präspezifikation der Subgruppen	12
5. Zusätzliche Subgruppen des pharmazeutischen Unternehmers oder des IQWiG	13
6. Terminologie	13
4. Methodik der Subgruppenbildung	14
1. European Medicines Agency (EMA)	14
2. Leitlinien – GRADE	15
3. Gemeinsamer Bundesausschuss (G-BA)	15
4. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)	17
5. Problematik der Subgruppenbildung bei der Nutzenbewertung	18
5. Kriterien der wissenschaftlichen medizinischen Fachgesellschaften	19
1. Klinische Relevanz	19
1. Unterschiedliche Standard- und Vergleichstherapie (prädiktive Merkmale)	19
2. Unterschiedlicher Verlauf	20
1. Krankheitsverlauf	20
2. Ansprechen auf die Therapie	20
2. Präspezifikation	21
1. Primärer Endpunkt	21
2. Andere Endpunkte	22
3. Stratifizierung	22
3. Plausible Begründung (Rationale)	23
4. Inkonsistenz und Heterogenität	23
5. Patientenzahl	23
6. Perspektiven	24
6. Fragenkatalog	25
1. Modell	25
2. Beispiel	25
7. Literatur	26
Anhang: Subgruppen in Verfahren der frühen Nutzenbewertung 2011 – 2015	28

1. Zusammenfassung

Die frühe Nutzenbewertung neu zugelassener Arzneimittel schafft seit ihrer Einführung Anfang 2011 die Basis für die Verhandlungen zum Preis neuer Arzneimittel zwischen Krankenkassen und pharmazeutischen Unternehmen. Gleichzeitig liefert sie weitere Informationen zum Nutzen-Schaden-Verhältnis neuer Arzneimittel in Deutschland. Analysen der bisherigen Festlegungen haben einige Kritikpunkte seitens der wissenschaftlichen medizinischen Fachgesellschaften identifiziert. Dazu gehören die Bildung und Bewertung von Subgruppen. Subgruppenanalysen sind bereits bei der Zulassung eines Arzneimittels ein Teil des komplexen Prozesses zur Formulierung der jeweiligen Indikation. Bei der Nutzenbewertung hat die Subgruppenbildung Einfluss auf das Ergebnis des Verfahrens und ggf. auch auf die Patientenversorgung.

Patientengruppen können heterogen bezüglich Prognose und Therapiewirkung sein. Homogene Patientenpopulationen erhöhen die Aussagekraft (Power) einer klinischen Studie durch Reduktion der Variabilität, heterogene Patientenpopulationen stärken deren externe Validität. Mittlere Effektaussagen über eine heterogene Gruppe können irreführend sein, da starke Effekte in einer Teilgruppe die fehlende Wirksamkeit oder sogar den Schaden in anderen Subgruppen kompensieren können. Bei Studien zu neuen Arzneimitteln oder anderen Therapieverfahren können deshalb die Ergebnisse von Subgruppenanalysen wertvolle Erkenntnisse zum Verständnis der Wirksamkeit in definierten Patientenpopulationen liefern.

Die Chance von Subgruppenanalysen ist die Identifikation einer Variablen oder einer Kombination von Variablen, die einen signifikanten und klinisch relevanten Einfluss auf den Einsatz des untersuchten Arzneimittels hat, entweder im Sinne einer präferenziellen Gabe oder im Sinne einer Vermeidung. Auf der anderen Seite sind Subgruppenanalysen anfällig für Manipulationen, Verzerrungen und falsche Folgerungen durch spezielle Selektion der Daten, insbesondere wenn die Subgruppen nachträglich (post hoc) in Kenntnis der vorliegenden Ergebnisse definiert wurden.

Auf der Basis einer Analyse der bisherigen Handhabung der Subgruppenbildung im Rahmen der frühen Nutzenbewertung sowie der Methodik von Subgruppenanalysen seitens der European Medicines Agency (EMA), des Gemeinsamen Bundesausschusses (G-BA), des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) und der GRADE Working Group lassen sich relevante Kriterien für wissenschaftlichen medizinischen Fachgesellschaften identifizieren.

Bei der Subgruppenbildung im Rahmen der Nutzenbewertung sind klinisch vor allem zwei Situationen zu unterscheiden:

- Subgruppen mit unterschiedlicher Standard- und Vergleichstherapie (prädiktive Merkmale)
- Subgruppen mit unterschiedlichem Verlauf (prognostische Merkmale)
 - Krankheitsverlauf
 - Ansprechen auf die Therapie

Ein entscheidendes Kriterium ist, dass die Subgruppenbildung in aktuellen Leitlinien oder Empfehlungen von Fachgesellschaften angewandt wird. Damit besteht ein begründeter Konsens über deren klinische Relevanz.

Subgruppenanalysen im Rahmen klinischer Studien sollen präspezifiziert sein, also a priori festgelegt werden. Dies betrifft die exakte Beschreibung des jeweiligen Merkmals, den Bezug auf den Endpunkt und die Methodik der Analyse. Post-Hoc-Analysen (a posteriori) haben grundsätzlich ein erhöhtes Verzerrungspotenzial. Eine Ausnahme können prospektiv-retrospektive Studien sein, bei denen Untersuchungsmaterial/-daten vor Therapiebeginn ergebnisoffen asserviert wurden.

Wir schlagen vor, dass alle Dossiers im Rahmen der Nutzenbewertung eine tabellarische Übersicht enthalten, in der der Status der Subgruppenfestlegung (a priori vs post hoc) und die Art ihrer Verwendung (Stratifikation, Auswertung o.a.) transparent dargestellt werden.

Die meisten Präspezifikationen von Subgruppen beruhen auf biologischen, pathophysiologischen, pharmakologischen, klinischen oder sozialen Parametern. Von besonderer Bedeutung ist, wenn rein statistische Tests auf Homogenität bzw. Interaktion einen Unterschied im Ansprechen auf ein neues Arzneimittel anzeigen.

Zur Untersuchung der Homo- bzw. Heterogenität von Daten sollen Interaktionstests eingesetzt werden. Als Signifikanzschwelle schlagen wir für Heterogenitäts- bzw. Interaktionstests eine Schwelle für den p-Wert von 0,1 vor. Dies erhöht die Power zum Nachweis von Interaktionen.

Regelhaft orientiert sich die Stichprobengröße von Zulassungsstudien am primären Endpunkt und ist nicht darauf angelegt, Unterschiede bzw. ggf. Nicht-Unterlegenheit in Subgruppen zu belegen. Kritisch ist die Bildung sehr kleiner Subgruppen. Zahlenmäßig liegen diese Subgruppen oft in der absoluten Häufigkeit von Erkrankungen mit Orphan-Drug-Status.

Wir schlagen vor, die Kriterien zur Subgruppenbildung transparent und einheitlich in die Stellungnahmen der Fachgesellschaften aufzunehmen. In dem Fragenkatalog wird jeder Parameter seitens der klinischen Experten nach den Kriterien „zutreffend“, „nicht zutreffend“, „unklar“ oder „keine Angaben“ bewertet.

Kriterien für Subgruppen	G-BA					pU		IQWiG	
	Subgruppe 1	2	3	4	...	1	...	1	...
klinisch relevant · eigener Therapiestandard in aktuellen Leitlinien · unterschiedliche Wirksamkeit der Therapie									
Präspezifikation · für den primären Endpunkt · für andere Endpunkte · Stratifizierung oder separate Studie									
Rationale (plausible Begründung)									
positiver Homogenitäts- bzw. Interaktionstest									
relative und absolute Häufigkeit									

Fragen werden mit ja, nein oder dem Hinweis auf unklare bzw. fehlende Angaben beantwortet. Der Fragenkatalog soll neben einer Vereinheitlichung in den Stellungnahmen der Fachgesellschaften zusätzliche Transparenz in dem wichtigen Gebiet der Bildung und Bewertung von Subgruppen schaffen.

2. Einleitung

Die frühe Nutzenbewertung neu zugelassener Arzneimittel schafft seit ihrer Einführung Anfang 2011 die Basis für die Verhandlungen zum Preis neuer Arzneimittel zwischen Krankenkassen und pharmazeutischen Unternehmen. Gleichzeitig liefert sie weitere Informationen zum Nutzen-Schaden-Verhältnis neuer Arzneimittel in Deutschland. Bis Ende 2015 wurden mehr als 150 Verfahren vollständig, d. h. mit Vorlage des Dossiers seitens des pharmazeutischen Unternehmers, Bericht von IQWiG oder G-BA, Stellungnahmen, Anhörung und Beschluss des G-BA zur Festlegung des Zusatznutzens, abgeschlossen. Analysen der bisherigen Festlegungen haben einige Kritikpunkte identifiziert, die von den wissenschaftlichen medizinischen Fachgesellschaften wiederholt in den Stellungnahmen zu den Verfahren vorgebracht wurden. Dazu gehören die Definition und der Umgang mit Subgruppen.

Patientengruppen können heterogen bezüglich Prognose und Therapiewirkung sein. Homogene Patientenpopulationen erhöhen die Aussagekraft (Power) einer klinischen Studie durch Reduktion der Variabilität, heterogene Patientenpopulationen klinischer Studien stärken deren externe Validität. Mittlere Effektaussagen über eine heterogene Gruppe können irreführend sein, da starke Effekte in einer Teilgruppe, fehlende Wirksamkeit oder sogar Schaden in anderen Subgruppen kompensieren können. Bei Studien zu neuen Arzneimitteln oder anderen Therapieverfahren können deshalb die Ergebnisse von Subgruppenanalysen wertvolle Erkenntnisse zum Verständnis der Wirksamkeit in definierten Patientenpopulationen liefern. Die Chance von Subgruppenanalysen ist die Identifikation einer patientenbedingter und behandlungsspezifischen Variablen bzw. einer Kombination von Variablen, die signifikanten und klinisch relevanten Einfluss auf den Einsatz des untersuchten Arzneimittels hat, entweder im Sinne einer präferenziellen Gabe oder im Sinne einer Vermeidung. Auf der anderen Seite sind Subgruppenanalysen anfällig für Manipulationen, Verzerrungen und falsche Folgerungen, insbesondere wenn die Subgruppen nachträglich (post hoc) in Kenntnis der vorliegenden Ergebnisse definiert wurden. Ein Zitat von Feinstein beschreibt die beiden entgegengesetzten Positionen treffend [7]:

“The essence of tragedy has been described as the destructive collision of two sets of protagonists, both of whom are correct. The statisticians are right in denouncing subgroups that are formed post hoc from exercises in pure data dredging. The clinicians are also right, however, in insisting that a subgroup is respectable and worthwhile when established a priori from pathophysiological principles.”

Kriterien zum Umgang mit Subgruppen sind abhängig vom Zwecke des jeweiligen Verfahrens. Der Fragenkatalog beim Design einer Studie ist nicht identisch mit demjenigen bei der Datenanalyse einer abgeschlossenen Studie, bei Metaanalysen, bei der Heranziehung von Ergebnissen für eine Zulassung, bei Leitlinien oder bei einer Nutzenbewertung. Es gibt bisher keine allgemein akzeptierten, methodischen Festlegungen zu Subgruppen. Allerdings lassen sich aus der Vielfalt von Publikationen, Reviews und Leitlinien bzw. Methodikpapieren gemeinsame, übergeordnete und situationsangepasste Kriterien identifizieren.

Subgruppenbildung im Rahmen der Nutzenbewertung neuer Arzneimittel findet auf dem Weg der Einführung eines neuen Arzneimittels relativ spät statt. Sie folgt den Entscheidungen zum Design von Zulassungsstudien, den Entscheidungen der Zulassungsbehörde zur Definition der Zulassungsindikation(en), und ggf. auch den Empfehlungen von wissenschaftlichen medizinischen Fachgesellschaften zum Einsatz der neuen Wirkstoffe. Gleichzeitig geht die Nutzenbewertung den Entscheidungen von Kostenträgern über die Preisbildung und die Kostenerstattung voran. Im deutschen Verfahren der frühen Nutzenbewertung nach dem Arzneimittelmarktneuordnungsgesetz (AMNOG) ist die Analyse von Subgruppen im Gesetz und in der Verfahrensordnung implementiert [19]:

Gesetz

zur Neuordnung des Arzneimittelmarktes in der gesetzlichen Krankenversicherung (Arzneimittelmarktneuordnungsgesetz – AMNOG)

§35a:

... Die Nutzenbewertung erfolgt aufgrund von Nachweisen des pharmazeutischen Unternehmers..., die insbesondere folgende Angaben enthalten müssen:

1. ...
4. Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht.

Dies wird in den Anforderungen an das Dossier der pharmazeutischen Unternehmer folgendermaßen umgesetzt [1]:
Modul 4

4. 2. 5. 5. Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. -stadium
- Zentrums- und Ländereffekte

Die medizinische Wissenschaft ist auf allen Ebenen der Entwicklung von neuen Arzneimitteln beteiligt, von den Grundlagen über klinische Studien bis zu Leitlinien und Projekten der Versorgungsforschung. Empfehlungen der wissenschaftlichen medizinischen Fachgesellschaften zum Umgang mit Subgruppen im Rahmen der Nutzenbewertungen sollen konsistent mit den Empfehlungen und dem Vorgehen in den anderen Verfahren sein.

In praxi haben Beschlüsse des G-BA starken Einfluss auf das Verschreibungsverhalten der Ärzte. Die Rechtsauffassung ist unklar, ob der final verhandelte Preis für alle Subgruppen gilt oder nur für die Subgruppen, für die vom G-BA ein Zusatznutzen festgelegt worden ist. Daher ist es die Verantwortung des G-BA, Subgruppen so zu definieren, wie sie im praktischen Alltag auch eingeteilt und diagnostiziert werden. Ein Beispiel: eine Leberbiopsie bei Patienten mit chronischer Hepatitis C wird heute unter Gesichtspunkten von Kosten und Nutzen nicht mehr routinemäßig durchgeführt. In den Studien erfolgt die Abschätzung des Fibrosestadiums heute mittels nichtinvasiver Verfahren (Elastographie), die wiederum in der Versorgungspraxis aber nicht verfügbar ist. Eine Subgruppe „kompensierte Leberzirrhose“ mit unterschiedlicher Zusatznutzenbewertung ist daher für den klinisch tätigen Arzt nicht nachvollziehbar.

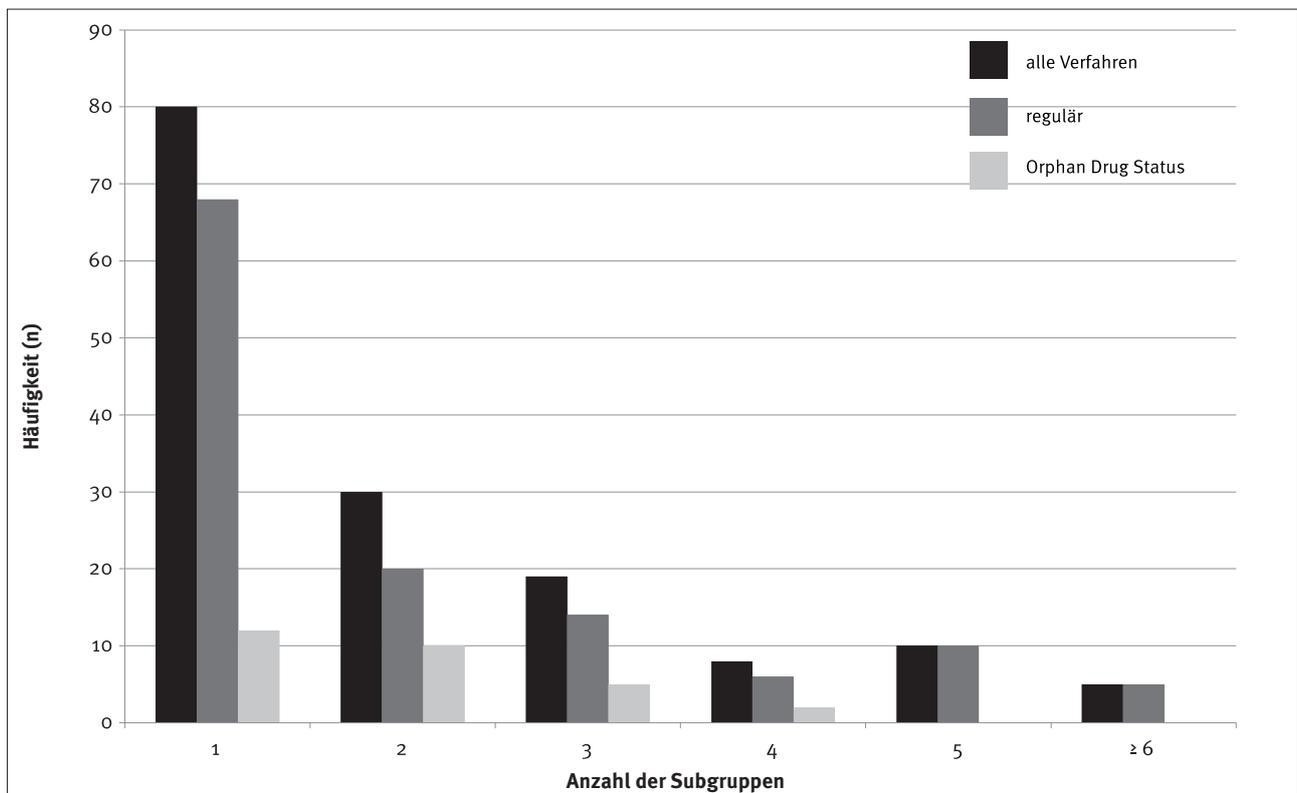
Im Folgenden wird zuerst die Handhabung der Subgruppenbildung in bis Ende 2015 vollständig abgeschlossenen Verfahren des G-BA im Rahmen der frühen Nutzenbewertung analysiert. Es folgen Zusammenfassungen zur Methodik der Subgruppenanalyse seitens der European Medicines Agency (EMA), des Gemeinsamen Bundesausschusses (G-BA), des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) und der GRADE Working Group. Darauf aufbauend werden relevante Kriterien für wissenschaftliche medizinische Fachgesellschaften in Bezug auf die frühe Nutzenbewertung identifiziert und begründet.

3. Subgruppenbildung in abgeschlossenen Verfahren nach dem AMNOG 2011 – 2015

3.1. Häufigkeit

Der G-BA hat in fast der Hälfte der Verfahren Subgruppen für die zweckmäßige Vergleichstherapie und für die Bestimmung des Zusatznutzens festgelegt. **Abbildung 1** zeigt die Häufigkeit von Subgruppen in den abgeschlossenen Verfahren. Mindestens 2 Subgruppen wurden in 72 Verfahren (47,4%) gebildet. Vor allem in den Verfahren zu neuen Antidiabetika und zu neuen Arzneimitteln zur Therapie der chronischen Hepatitis C wurden >5 Subgruppen gebildet.

Abbildung 1: Anzahl von Subgruppen in vollständig abgeschlossenen Verfahren (2011 – 2015)



Die Abbildung zeigt, dass Subgruppen auch in Verfahren für Arzneimittel im Orphan-Drug-Status gebildet werden. Der relative Anteil liegt mit 59% (17/29) bei diesen Verfahren für Arzneimittel für seltene Erkrankungen sogar höher als in den anderen Verfahren.

Die Häufigkeit der Subgruppenbildung schwankte in den vergangenen Jahren und lag im Jahr 2015 höher als in allen vorangegangenen Jahren. Die Auswertung auf der Basis der bis Ende 2011 vollständig abgeschlossenen Verfahren ist in **Tabelle 1** zusammengestellt.

Tabelle 1: Subgruppenbildung in vollständig abgeschlossenen Verfahren 2011 – 2015

Jahr	alle	>1 Subgruppe	% Verfahren mit Subgruppenbildung
2011/2012	27	11	41
2013	34	18	53
2014	38	16	42
2015	53	30	57

3.2. Kriterien

Tabelle A im Anhang listet die Charakteristika aller Subgruppen auf. Es zeigt sich eine recht große Heterogenität, sowohl in den Kriterien als auch in der Zahl der Subgruppen. Dabei wurden vor allem folgende Parameter berücksichtigt (alphabetische Reihenfolge):

- Allgemeinzustand
- Alter
- histologische, immunologische, molekulargenetische Charakterisierung
- Komorbidität
- Kontraindikationen gegenüber Standardtherapie
- Krankheitsstadium
- Krankheitsschwere, -aktivität, -symptomatik
- Therapieziel
- Vortherapie

Bei einigen, bisher abgeschlossenen Verfahren waren die Subgruppenmerkmale für dieselbe Indikation nicht identisch, z. B. bei Antidiabetika, bei Arzneimittel für Patienten mit chronischer Hepatitis C und bei onkologischen Indikationen wie dem kastrationsresistenten Prostatakarzinom, siehe Anhang Tabelle A. Auch bei erneuter Bewertung eines Arzneimittels nach Fristablauf oder bei Orphan Drugs nach Überschreiten der kritischen Umsatzgrenze sind die Merkmale im ersten und im zweiten Verfahren nicht durchgehend identisch [39]. Das kann „im lernenden System“ der frühen Nutzenbewertung auch als Reaktion auf vorhergehende kritische Diskussionen interpretiert werden.

3.3. Ergebnisse

Die Bildung von Subgruppen kann einen Einfluss auf die Festlegung des Zusatznutzens haben. In [Abbildung 2](#) sind die Ergebnisse aller vollständig abgeschlossenen Verfahren bis Ende 2015 dargestellt. [Abbildung 3](#) differenziert die Ergebnisse nach der Anzahl der Subgruppen.

Abbildung 2: Festlegung des Zusatznutzens in vollständig abgeschlossenen Verfahren 2011 – 2015

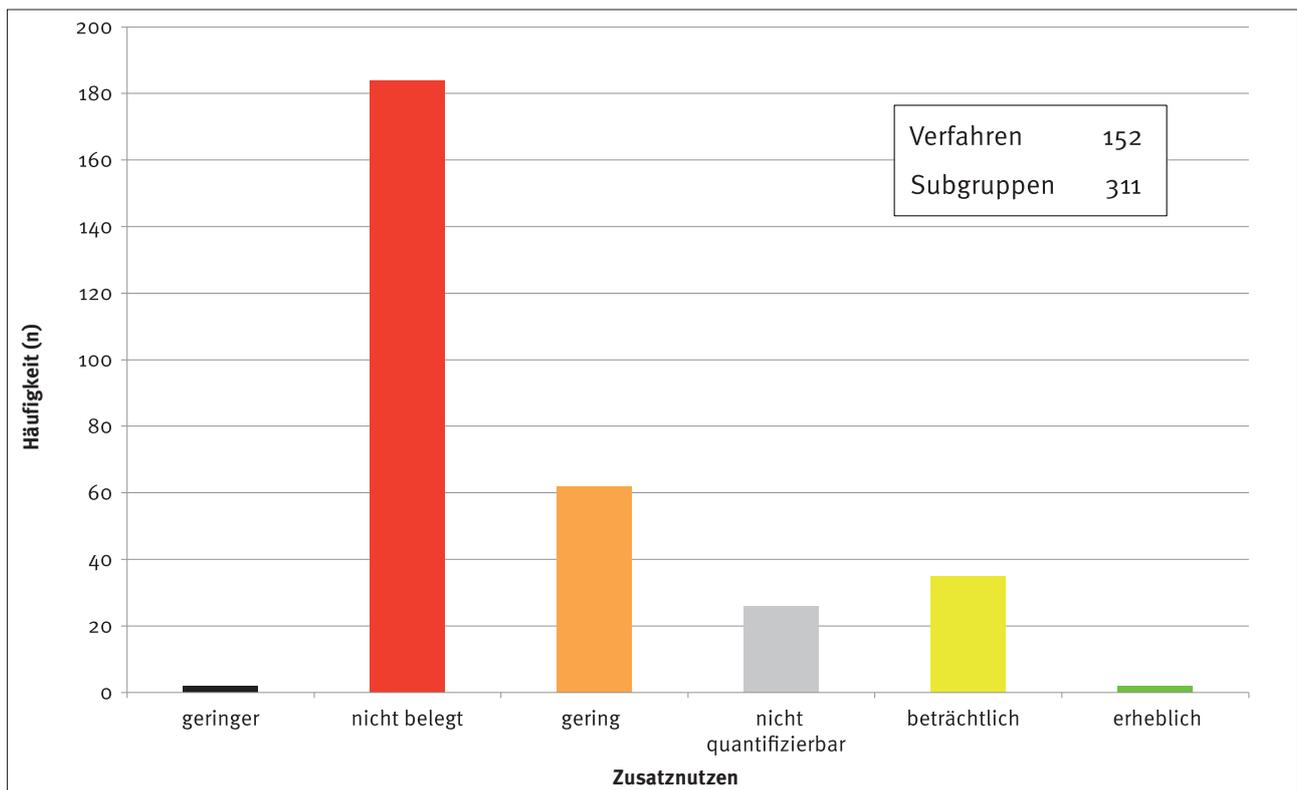
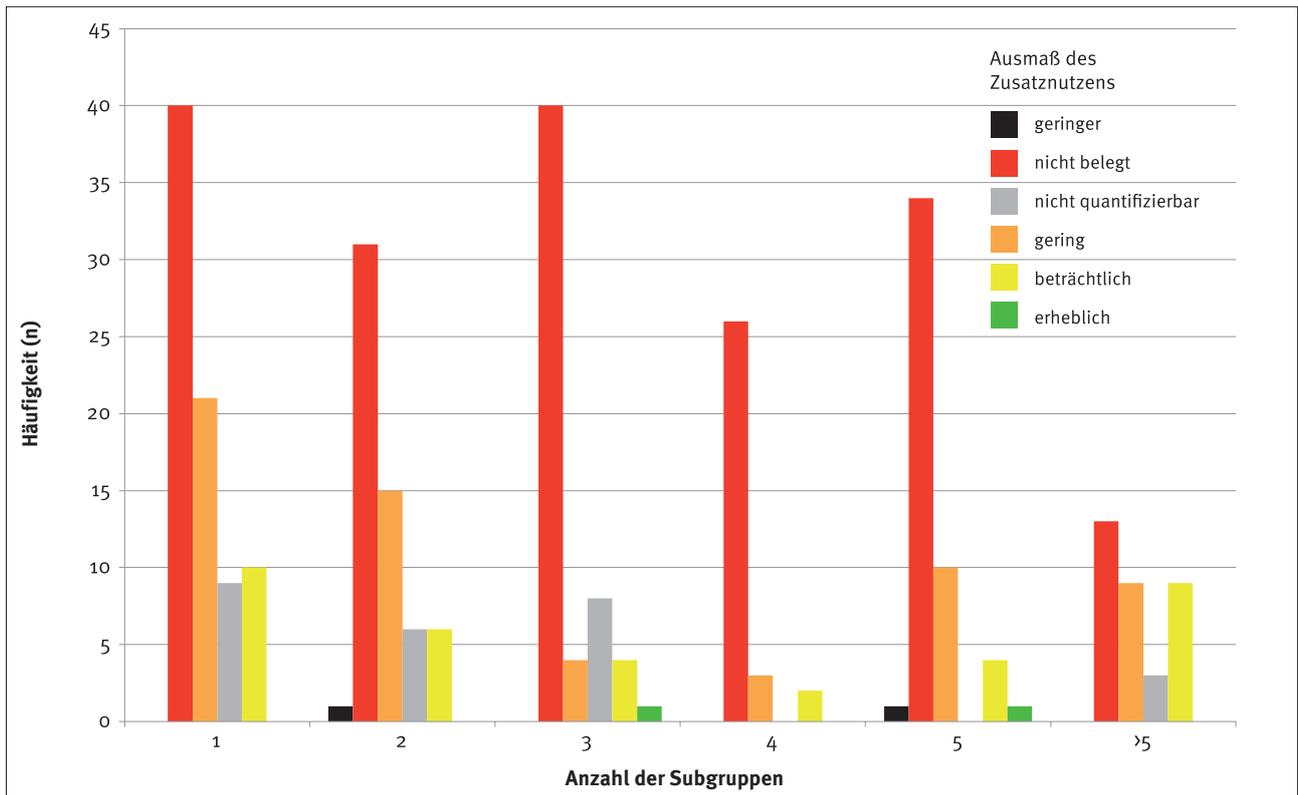


Abbildung 3: Festlegung des Zusatznutzens nach der Anzahl der Subgruppen

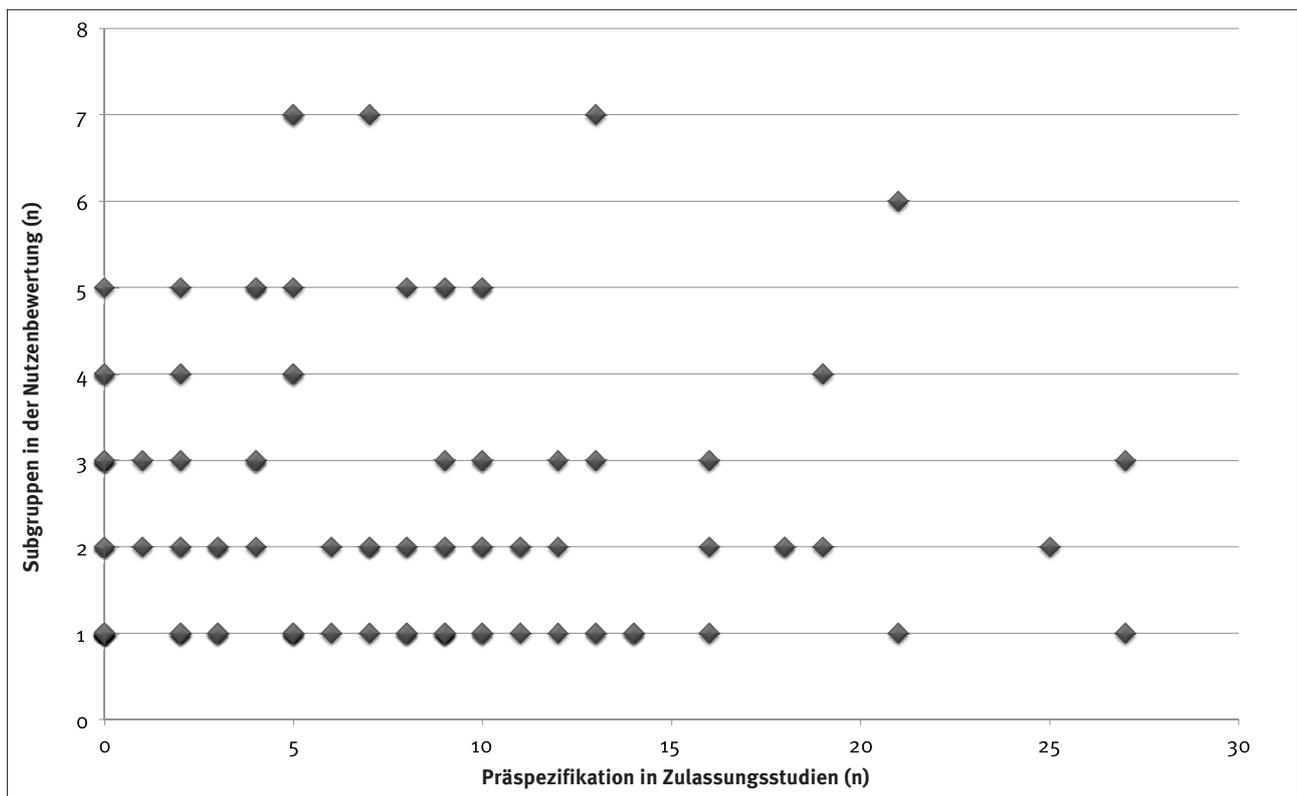


Ein beträchtlicher, aber auch ein geringer Zusatznutzen wurde häufiger bei Verfahren mit 1-2 oder mit >5 Subgruppen festgestellt. Bei Bildung von 3-4 Subgruppen wurde häufiger kein Zusatznutzen festgestellt.

3.4. Präspezifikation

Im Dossier sind Subgruppenanalysen zu Geschlecht, Alter, Krankheitsschwere bzw. -stadium und zu möglichen Zentrums- bzw. Ländereffekten vorgesehen. Gleichzeitig muss der pharmazeutische Unternehmer angeben, welche Subgruppenanalysen präspezifiziert waren, d. h. a priori vorgesehen waren. [Tabelle A](#) im Anhang macht die extreme Diversität im Umgang mit den Subgruppen deutlich. Die Zahl der präspezifizierten Subgruppen reicht von 0 bis >20. [Abbildung 4](#) bildet dies graphisch ab.

Abbildung 4: Zahl der präspezifizierten und Zahl der vom G-BA festgelegten Subgruppen



Häufig legt der G-BA im Unterschied zum Design der Zulassungsstudie(n) eine deutlich niedrigere Zahl von Subgruppen fest. In einigen Fällen wurden mehr Subgruppen festgelegt. Die Definition der Subgruppenmerkmale durch den G-BA kann sich von der Definition der Zulassungsstudie(n) unterscheiden.

3.5. Zusätzliche Subgruppen des pharmazeutischen Unternehmers oder des IQWiG

Der pharmazeutische Unternehmer und das IQWiG können die Bildung zusätzlicher Subgruppen vorschlagen [39]. Grundlage für Vorschläge des IQWiG sind Homogenitäts- und Interaktionstests. Grundlage für Vorschläge der pharmazeutischen Unternehmer sind weitere Analysen der Daten aus der Zulassungsstudie, z. T. mit präspezifizierten Parametern, z. T. post hoc.

Diese Vorschläge werden in den Stellungnahmen und in der Anhörung oft intensiv diskutiert, aber selten in die Festlegungen des Zusatznutzens aufgenommen. Ausnahme ist das Verfahren zu Afatinib, einem gezielten Tyrosinkinase-Inhibitor zur Therapie des nichtkleinzelligen Lungenkarzinoms mit EGFR-aktivierenden Mutationen. Hier wurde vom IQWiG im ersten Verfahren der Hinweis auf eine Effektmodifikation auf das Gesamtüberleben durch die Art der molekularen Aberration festgestellt und vom G-BA in die Festlegung des Zusatznutzens aufgenommen. Im zweiten Verfahren nach Fristablauf führten diese molekularen Subtypen zusammen mit zwei klinischen Kriterien zur Definition von insgesamt 5 Subgruppen.

3.6. Terminologie

In den Verfahren der frühen Nutzenbewertung werden vom G-BA und vom IQWiG manchmal zwei Begriffe verwendet.

- Subpopulationen: Patientengruppen, die sich aus dem Wortlaut des zugelassenen Anwendungsgebietes ergeben
- Subgruppen: Teilpopulationen einer Studie, die sich aufgrund einer potenziellen Effektmodifikation, d. h. unterschieden im Ausmaß eines Effekts, ergeben.

4. Methodik der Subgruppenbildung

4.1. European Medicines Agency

Die European Medicines Agency (EMA) hat 2014 ihren Entwurf für eine „Guideline on the investigation of subgroups in confirmatory clinical trials“ publiziert [6]. Die EMA differenziert die beiden Situationen

- Studienplanung
- Datenanalyse

und differenziert drei Arten der Subgruppenanalyse

- konfirmatorisch
- „hervorgehoben (key)“ exploratorisch
- exploratisch.

Der Entwurf der Leitlinie diskutiert, wie durch eine Maximierung von a priori festgelegten Subgruppenanalysen eine Minimierung a posteriori festzulegender Subgruppenanalysen erfolgen kann. Es findet sich darin leider keine schlüssige Strategie, dies auch umsetzen zu können. Wichtig ist der Aspekt, dass nur Variablen, die vor der Randomisierung gemessen wurden, für die Subgruppenbildung herangezogen werden können.

In der Zusammenfassung der EMA werden 6 zentrale Punkte formuliert:

1. Die Untersuchung von Behandlungseffekten in klar definierten Subgruppen der Studienpopulation ist integraler Bestandteil der Studienplanung, der Analyse und von Schlussfolgerungen, die der Betrachtung des primären Endpunktes folgen. Die Guideline soll bei der Planung und der Präsentation dieser Analysen und beim Verständnis der Faktoren helfen, die bezüglich der Plausibilität der Ergebnisse zu diskutieren sind.
2. Je homogener die Studienpopulation in Bezug auf das Basisrisiko und das Therapieansprechen ist, umso geringer ist die Notwendigkeit exploratorischer Subgruppenanalysen für die Zulassung. Je heterogener die Studienpopulation ist, umso größer ist die Bedeutung von Subgruppenanalysen zur Einschätzung, ob der voraussichtliche Gesamteffekt breit anwendbar ist und sich die Bewertung des Nutzen-Schaden-Verhältnisses auf die gesamte, beantragte Indikation übertragen lässt. Untersuchungen zur Heterogenität sollen Kovariablen berücksichtigen und Subgruppenanalysen enthalten.
3. Aufgrund der methodischen Schwierigkeiten bei multiplen Analysen müssen exploratorische Untersuchungen mit Vorsicht angegangen werden und die gesamte verfügbare Evidenz berücksichtigen, nicht nur die Punktschätzer der individuellen Subgruppenanalysen. Trotz der statistischen Schwierigkeiten können auch das Unterlassen und das Ignorieren von Subgruppenanalysen zu falschen Entscheidungen führen.
4. Die Berichtersteller sollen in den Studienprotokollen Unterlagen zum erwarteten Ausmaß der Heterogenität der Studienpopulation sowohl in Bezug auf prognostische Faktoren für den Krankheitsverlauf als auch auf prädiktive Faktoren für das Therapieansprechen finden. Eine Strategie, die einfach Homogenität einer Population in Bezug auf den Behandlungseffekt annimmt, ohne Diskussion und ohne weitere Untersuchungen, ist nicht ausreichend. Analog ist es auch nicht ausreichend, alle Subgruppenanalysen mit Hinweisen auf ein heterogenes Ansprechen als zufällig zu verwerfen. Der Gewinn dieser zusätzlichen Diskussion soll darin liegen, die *a priori* Diskussion über bestimmte Subgruppen zu intensivieren und dadurch *a posteriori* Diskussionen zu vermeiden in dem Versuch, eine rationale Betrachtung von Subgruppen zu fördern und das Risiko falscher Schlussfolgerungen zu reduzieren. Wenn korrekt durchgeführt, sollte sich die Analyse auf eine gut begründete, präspezifizierte Strategie stützen und der Bedarf an Daten-gesteuerten Untersuchungen minimiert werden.
5. Die Konsistenz der Daten in relevanten Subgruppen muss in dem Analysebericht diskutiert werden. Dabei spielen Forest-Plots mit graphischer Darstellung von Behandlungseffekten in einer Serie von Subgruppen und statistische Methoden zur Erfassung der Heterogenität von Behandlungseffekten in Subgruppen eine wichtige Rolle, um Signale dafür zu erhalten, ob sich der Gesamteffekt der Behandlung auf die gesamte Studienpopulation bezieht. Klinische und pharmakologische Kenntnisse sind erforderlich, um die Glaubwürdigkeit und die Relevanz der Subgruppenergebnisse zu bewerten. Eine ganze Reihe von Faktoren beeinflussen die Glaubwürdigkeit von Subgruppenergebnissen, einschließlich der „biologischen Plausibilität“, der Reproduzierbarkeit der Evidenz und der Evidenzstärke in den Studien. Für die Heterogenität sollen glaubwürdige Erklärungen gesucht werden. Zum Beleg der Glaubwürdigkeit können multiple Analysen und Präsentationen der Daten erforderlich sein.

6. In dieser Guideline wird eine Strategie zur Erfassung der Glaubwürdigkeit von Subgruppenergebnissen in unterschiedlichen Situationen vorgestellt. Dazu gehören zentrale Überlegungen für den Wechsel von einer Analyse der Gesamtpopulation zu einer Subgruppenanalyse. Normalerweise werden Subgruppenanalysen nicht zur Rettung einer Studie mit negativem Ausgang führen.

4.2. Leitlinien – GRADE

Eine geeignete Plattform für wissenschaftliche Fachgesellschaften für Empfehlungen zur Einschätzung der Qualität klinischer Studien ist GRADE (Grading of Recommendations, Assessment, Development and Evaluation) [22]. Ziel von GRADE ist die Verbindung einer Evidenzbeurteilung mit der daraus folgenden Ableitung von Empfehlungen in klinischen Leitlinien anhand definierter Kriterien. Seit 2004 hat eine internationale Expertengruppe ein umfassendes System zu den verschiedenen Inhalten der Qualitätsbeurteilung von Evidenzprofilen über die verschiedenen Aspekte der Ergebnisverzerrung bis zur Vorbereitung von Zusammenfassungen publiziert. Elemente des GRADE Systems finden zunehmend Anwendung in internationalen und deutschen klinischen Leitlinien. In seiner Gesamtheit und der konzeptionellen Stringenz bietet sich dieses System besonders bei Themen an, in denen dieselben Studien Grundlage von Empfehlungen im Rahmen von Leitlinien als auch Grundlage einer Nutzenbewertung sind.

Das Thema der Subgruppenbildung wurde von der GRADE Working Group und ihren Initiatoren im Rahmen der Empfehlungen zum Umgang mit Inkonsistenz von Studiendaten bearbeitet [21, 23]. Aufgrund von Beispielen, in denen Ergebnisse von Subgruppenanalysen in späteren Studien nicht bestätigt wurden, wird grundsätzlich eine vorsichtige Interpretation der Ergebnisse von Subgruppenanalysen empfohlen.

GRADE schlägt 7 Kriterien zur Beurteilung der Glaubwürdigkeit von Subgruppenanalysen vor [23]:

- Wurde die Subgruppenvariable präspezifiziert (im Gegensatz zur Spezifizierung nach Randomisierung)?
- Wurde die Differenz zwischen den Subgruppen durch Vergleiche innerhalb der Studie(n) suggeriert (im Gegensatz zu Vergleichen zwischen Studien)?
- Suggestiert die statistische Analyse, dass Zufall unwahrscheinlich ist als Erklärung für die Subgruppenunterschiede?
- Ging die Hypothese der Analyse voraus und bestätigte die Analyse die vermutete Richtung des Effektes (im Gegensatz zu einer Hypothese, die der Analyse folgt)?
- War die spezifische Subgruppenanalyse eine von mehreren Subgruppenanalysen?
- Ist der Subgruppenunterschied konsistent in anderen Studien und bezüglich wichtiger Endpunkte?
- Bestätigt externe Evidenz (biologisch oder soziologische Rationale) den vermuteten Subgruppenunterschied?

Die GRADE-Methodik ist Leitlinien-orientiert mit einem Schwerpunkt auf Metaanalysen zu jeweils definierten Endpunkten einer spezifischen klinischen Fragestellung. Dies trifft auf die Datenlage in Mehrzahl der Verfahren im Rahmen der frühen Nutzenbewertung nicht zu. Die zugrundeliegenden Überlegungen und einige der Fragen sind aber auf die Nutzenbewertung übertragbar [35, 36].

4.3. Gemeinsamer Bundesausschuss (G-BA)

Der G-BA hat für den Umgang mit Subgruppen keine eigene Methodik publiziert. Inhaltlich orientiert er sich an den methodischen Standards, wie sie von P. M. Rothwell in mehreren Übersichten im Jahr 2005 publiziert wurden [30, 31]. Auf der Basis der eigenen wissenschaftlichen Arbeiten in der klinischen Schlaganfallforschung und der Analyse des Umgangs mit Subgruppen in randomisierten klinischen Studien entwickelte Rothwell zwei Gruppen von Fragen als Leitlinie für den Umgang mit Subgruppen. Da die große Mehrzahl der relevanten Themen in der Subgruppenforschung in diesen Vorschlägen enthalten ist, werden sie hier dargestellt.

Design klinischer Studien

- Subgruppen sollen vor Beginn der Studie definiert werden und auf eine kleine Anzahl klinisch relevanter Fragen begrenzt sein.
- Beratung durch Experten ist erforderlich um sicherzustellen, dass alle relevanten Risikofaktoren und andere Daten erfasst werden.
- Die Richtung und die Größe des erwarteten Subgruppeneffektes soll vor Beginn der Studie festgelegt werden.
- Die genauen Definitionen und Kategorien der Subgruppenvariablen sollen explizit im Studienprotokoll definiert werden, um Daten-getriebene Post-Hoc-Analysen von Variablen oder Kategorien zu vermeiden. Für kontinuierliche oder hierarchische Variablen sollen die Grenzwerte (Cut-off Points) präspezifiziert werden.
- Bei wichtigen Subgruppenvariablen soll eine Stratifikation vor der Randomisierung erwogen werden.
- Wenn wichtige Behandlungseffekte für Subgruppen antizipiert werden, sollen die Studien idealerweise auch in den Patientenzahlen so angelegt werden, dass diese Effekte zuverlässig nachgewiesen werden können.
- Regeln zum Studienabbruch sollen vorhersehbare Subgruppeneffekte und nicht nur den Gesamteffekt der Behandlung mit berücksichtigen.
- Wenn der relative Behandlungseffekt voraussichtlich mit vor Therapie bekannten Risikofaktoren korrelieren wird, soll die geplante Analyse eine Stratifikation nach diesem Risiko beinhalten. Risiko-Score oder – Modell sollen vorher festgelegt werden, um die relevanten Basisdaten in der Studie zu dokumentieren.

Analyse und Bericht (Veröffentlichung)

- Das oben skizzierte Design soll im Methodenteil zusammen mit Details über die Auswahl der Subgruppen (wie, warum) dargestellt werden.
- Die Signifikanz des Behandlungseffektes in individuellen Subgruppen soll nicht berichtet werden; die Raten falsch negativer oder falsch positiver Ergebnisse sind extrem hoch. Der einzig, statistisch zuverlässige Ansatz ist der Test auf Interaktionen.
- Alle durchgeführten Subgruppenanalysen sollen berichtet werden, d. h. nicht nur die Zahl der Subgruppenvariablen sondern auch die Zahl der pro Subgruppe analysierten Endpunkte, die unterschiedliche Länge der Nachbeobachtung etc.
- Die Signifikanz der Interaktionsanalysen für Behandlungseffekte in präspezifizierten Subgruppen soll angepasst werden, wenn multiple Subgruppenanalysen durchgeführt werden.
- Subgruppenanalysen sollen als Reduktionen des absoluten und des relativen Risikos berichtet werden. Wenn relevant, soll die statistische Signifikanz der Unterschiede in der Reduktion des absoluten Risikos untersucht werden.
- Idealerweise soll nur ein Endpunkt untersucht werden und dies soll üblicherweise der primäre Studienendpunkt sein, unabhängig davon, ob dieser Endpunkt ein klinisch relevanter, zusammengesetzter Endpunkt ist.
- In den Studienarmen sollen die Subgruppen auf ihre Vergleichbarkeit in Bezug auf prognostische Faktoren untersucht werden.
- Wenn multiple Effekte in den Interaktionstests identifiziert werden, sind weitere Analysen zur Unabhängigkeit dieser Effekte erforderlich.

Auswertung

- Berichte über signifikante Behandlungseffekte in individuellen Subgruppen sollen ignoriert werden, insbesondere über einen fehlenden Nutzen in einer spezifischen Subgruppe bei Vorliegen eines Nutzens in der Gesamtgruppe. Ausnahme ist der Nachweis eines signifikanten Effektes im Interaktionstest.
- Echte, unerwartete Behandlungseffekte in Interaktionstest sind selten (wenn klinische Experten bei der Definition potenziell relevanter Subgruppen am Studiendesign beteiligt waren). Deshalb sollen scheinbare, post hoc entdeckte Interaktionen mit Vorsicht interpretiert werden. In dieser Situation gibt es keinen zuverlässigen Test auf Signifikanz.
- Präspezifizierte Subgruppen sind nicht per se stichhaltig und sollen mit Vorsicht interpretiert werden. Die Rate falsch positiver Interaktionstests bei Behandlungseffekten in Subgruppen liegt bei 5% pro Subgruppe.

- Der beste Test zur Stichhaltigkeit einer statistisch signifikanten Interaktion ist die Reproduzierbarkeit in anderen Studien.
- Nur wenige Studien sind auf den Nachweis von Subgruppeneffekten angelegt. Daraus resultiert eine hohe Rate falsch negativer Ergebnisse in den Interaktionsanalysen von Behandlungseffekten in Subgruppen.

Subgruppen-Analyse, Teil 2: Die 4 wesentlichen klinischen Indikationen [30]

Potenzielle Heterogenität der Behandlungseffekte aufgrund von Risikofaktoren

- Unterschiede durch Risikofaktoren der Behandlung
- Unterschiede durch Risikofaktoren ohne Behandlung

Potenzielle Heterogenität der Behandlungseffekte aufgrund der Pathophysiologie

- Multiple pathophysiologische Faktoren als Ursache des klinischen Syndroms
- Unterschiede im biologischen Ansprechen aufgrund eines einzelnen pathophysiologischen Faktors
- Genetische Variabilität

Klinisch relevante Fragen zur praktischen Durchführung der Behandlung

- Wird der Nutzen durch die Schwere der Erkrankung beeinflusst?
- Wird der Nutzen durch das Stadium im natürlichen Krankheitsverlauf beeinflusst?
- Wird der Nutzen durch den Zeitpunkt der Behandlung nach einem klinischen Ereignis beeinflusst?
- Wird der Nutzen durch Komorbidität beeinflusst?

Unterdurchschnittlicher Einsatz der Behandlung in der klinischen Routine aufgrund von Unsicherheit über den Nutzen

- in bestimmten Gruppen, z. B. bei älteren Patienten
- Begrenzung der Behandlung durch einen engen Korridor relevanter physiologischer Variablen, z. B. Schwellen von Cholesterin- oder Blutdruckwerten.

Diese Aufstellung von Rothwell umfasst fast das gesamte Spektrum der relevanten und kritischen Themen. Inhaltlich fehlt nur der Umgang mit sogenannten prospektiv-retrospektiven Analysen, bei denen Untersuchungsmaterial im Rahmen randomisierter Studien prospektiv asserviert, aber mit neuen z. B. molekulargenetischen Methoden später analysiert und damit formal retrospektiv in seinem Einfluss auf den Behandlungseffekt analysiert wurde. Dieser Ansatz ist erst in den letzten Jahren vermehrt eingesetzt worden.

Allerdings decken die verschiedenen Kriterien von Rothwell sehr unterschiedliche Phasen klinischer Studien von der Evidenzgenerierung bis zur Post-Hoc-Analyse ab. Sie sind nicht spezifisch und nicht fokussiert auf die Situation der Nutzenbewertung. Ebenso fehlt eine Hierarchisierung der Inhalte, z. B. auf der Basis der Validierung der unterschiedlichen Kriterien. Aus der Möglichkeit zur Auswahl aus den verschiedenen Bereichen ohne verbindliche Festlegung obligat versus fakultativ zu erfüllender Kriterien ergibt sich ein Risiko für willkürliche Definition von Subgruppen im jeweiligen Einzelverfahren.

4.4. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) hat in seine Methodik für die Nutzenbewertung auch die Subgruppenanalyse aufgenommen [24, 25]. Einleitend weist es auf Publikationen hin, in denen Subgruppenanalysen sehr kritisch diskutiert werden [2, 28]. Hervorgehoben werden vor allem drei kontroverse Aspekte [25, Kapitel 8.1.6]:

- kein Beweischarakter: Subgruppenanalysen werden selten *a priori* geplant und sind selten Bestandteil des Studienprotokolls (ggf. auch von Amendments). Wenn Subgruppenanalysen bezüglich mehr oder weniger arbiträrer subgruppenbildender Merkmale post hoc stattfinden, können ihre Ergebnisse nicht als methodisch korrekte Prüfung einer Hypothese betrachtet werden.
- multiples Testen: Bei der Analyse zahlreicher Subgruppen besteht eine manchmal recht hohe Wahrscheinlichkeit, dass die Ergebnisse irgendeiner Subgruppe statistische Signifikanz erreichen, obwohl es sich in Wahrheit um ein zufälliges Ergebnis handelt.

- geringe Power: Oft erreichen Subgruppen nicht die Stichprobengröße, die für das (inferenzstatistische) Aufdecken moderater Unterschiede nötig ist, so dass auch beim Vorliegen tatsächlicher Effekte mit nicht signifikanten Ergebnissen zu rechnen ist. Anders ist die Situation, wenn schon bei der Fallzahlkalkulation eine für die Subgruppenanalyse angemessene Power berücksichtigt und ein entsprechend erhöhter Stichprobenumfang geplant wurde. Der statistische Nachweis unterschiedlicher Effekte zwischen verschiedenen Subgruppen muss anhand eines adäquaten Homogenitäts- bzw. Interaktionstests vorgenommen werden. Der Befund, dass in einer Subgruppe ein statistisch signifikanter Effekt beobachtet wurde, in einer anderen aber nicht, kann (inferenzstatistisch) nicht als das Vorliegen eines Subgruppeneffekts interpretiert werden. Der geringeren Aussagekraft (Power) wird durch das Heben des Signifikanzniveaus begegnet, wodurch sich die Aussagekraft erhöht.

Als Goldstandard für Subgruppenanalysen werden Auswertungen von a priori gebildeten Subgruppen bezeichnet, wenn eine anhand der Subgruppen stratifizierte Randomisierung und eine für die Datenanalyse adäquate statistische Methode (Homogenitätstest, Interaktionstest) vorgenommen bzw. angewandt werden [5].

In der Praxis übernimmt das IQWiG in seinen vom G-BA in Auftrag gegebenen Berichten zu den Dossiers der pharmazeutischen Unternehmer regelhaft die vorgegebenen Subgruppen. Grundlage zusätzlicher Vorschläge für die Bildung von Subgruppen sind Homogenitäts- bzw. Interaktionstests in Bezug auf die Mortalität oder andere, vom IQWiG als relevant angesehen Endpunkten. Aus der Höhe der statistischen Signifikanz wurde folgende Operationalisierung zur Klassifikation der Effekte festgelegt:

- p-Wert zwischen 0,05 und 0,2: Hinweis auf unterschiedliche Effekte zwischen Subgruppen
- p-Wert <0,05: Beleg auf unterschiedliche Effekte zwischen Subgruppen.

Für den Umgang mit den Ergebnissen zu Subgruppenanalysen in Bezug auf unterschiedliche Endpunkte schlägt die IQWiG-Methodik vor, ggf. eine Hierarchisierung vorzunehmen. Dabei stützen sich die Vorschläge zum Zusatznutzen in erster Linie auf Belege bezüglich der höher gewichteten Zielgrößen. Gleichzeitig weist das IQWiG-Methodenpapier auf den gesetzlichen Auftrag von Subgruppenanalysen hin, auch wenn diese nicht a priori spezifiziert wurden.

4.5. Problematik der Subgruppenbildung bei der Nutzenbewertung

In einer aktuellen Veröffentlichung wird die Problematik der Subgruppenbildung beim Prozess der Nutzenbewertung im Rahmen des AMNOG analysiert [29]. Auf der Basis der praktizierten Methodik und Systematik von IQWiG und G-BA werden wesentliche Problemfelder rund um Subgruppenanalysen, Subpopulationen sowie Inkongruenzen zwischen den Vorgaben der frühen Nutzenbewertung und der Zulassung aufgezeigt. Dabei wird auf eine in Fachkreisen konsentiertere Schnittmenge an Regeln zum Umgang mit Subgruppenanalysen hingewiesen. Diese finden sich auch in einem Gutachten des Medizinischen Dienstes der Spitzenverbände der Krankenkasse e.V. von 2004 [8]. Notwendige Kriterien auf Basis eines systematischen Reviews zur internationalen methodischen Literatur für die Interpretation von Subgruppenanalysen sind:

- a priori Planung der Subgruppenanalysen und Spezifizierung der Effektrichtung
- Durchführung von Interaktionstests (inkl. Interpretation des p-Werts und der Effektunterschiede)
- Berücksichtigung der Multiplizität hinsichtlich der Anzahl der Subgruppen
- Subgruppenmerkmale als Baseline-Eigenschaft und Faktor einer stratifizierten Randomisierung
- Konsistenz der Ergebnisse über mehrere Endpunkte
- Medizinische Plausibilität und Vergleichbarkeit mit anderen Studienergebnissen

Subgruppenanalysen kann nur dann ausreichend Glauben geschenkt werden, wenn die Anzahl präspezifizierter Subgruppen gering bleibt, die Effekte groß, konsistent und auch aus klinischer Sicht plausibel sind.

In der praktischen Umsetzung sind die IQWiG-Vorgaben, dass für alle Teilindikationen sowie untersuchte primäre, sekundäre oder tertiäre Endpunkte jeweils Subgruppenanalysen für alle potenzielle Effektmodifikatoren durchgeführt werden müssen. Neben einer mehrfachen Aufteilung der Zielpopulation hinsichtlich der untersuchten Fragestellung werden auch mehrschichtige Subgruppenanalysen praktiziert. Dies kann zu einer Subpopulation-Überlagerung auf mehreren Ebenen führen (z.B. beim Verfahren zu Afatinib mit Vorbehandlungsstatus, ECOG-Status, EGFR-Mutation, Alter). Durch den multiplikativen Zusammenhang all dieser Ausprägungen ergibt sich eine große (oft sogar dreistellige) Anzahl an Subgruppenanalysen und zwangsläufig auch eine Reduktion der jeweiligen Teilmengen auf wenige Patienten.

5. Kriterien der wissenschaftlichen medizinischen Fachgesellschaften

Diese Vorschläge beschränken sich auf die Situation der Nutzenbewertung und auf die Positionierung der wissenschaftlichen medizinischen Fachgesellschaften im aktuellen Prozess der frühen Nutzenbewertung. Zum Zeitpunkt der Stellungnahmen liegen bereits die Festlegungen des G-BA zur Bildung von Subgruppen und zur zweckmäßigen Vergleichstherapie, das Dossier des pharmazeutischen Unternehmers und der Bericht von IQWiG bzw. G-BA vor. Gleichzeitig sind die Experten vertraut mit den Ergebnissen der Zulassungs- und weiterer Studien sowie mit Text und Diskussion der Arzneimittelbehörden.

Im Folgenden werden die zentralen Inhalte der Bewertung von Subgruppen dargestellt.

5.1. Klinische Relevanz

Bei der Subgruppenbildung im Rahmen der Nutzenbewertung sind vor allem zwei verschiedene Situationen zu unterscheiden:

- Subgruppen mit unterschiedlicher Standard- und Vergleichstherapie (prädiktive Merkmale)
- Subgruppen mit unterschiedlichem Verlauf (prognostische Merkmale)
 - Krankheitsverlauf
 - Ansprechen auf die Therapie

5.1.1. Unterschiedliche Standard- und Vergleichstherapie (prädiktive Merkmale)

Die Subgruppenbildung ist von hoher Bedeutung im Rahmen der Nutzenbewertung, da sie die Festlegung der zweckmäßigen Vergleichstherapie betrifft und damit maßgeblichen Einfluss auf die endgültige Festlegung des Zusatznutzens hat. Wird eine Vergleichstherapie festgelegt, für die es (bekanntermaßen) keine randomisierte Studie gibt, ist das Ergebnis „Zusatznutzen nicht belegt“ vorhersehbar. Im bisherigen AMNOG-Prozess haben die wissenschaftlichen medizinischen Fachgesellschaften die Festlegung der jeweiligen Vergleichstherapie(n) in mehr als 30% der Verfahren kritisiert.

Entscheidendes Kriterium ist, dass die jeweilige Subgruppenbildung in aktuellen Leitlinien oder aktuellen Empfehlungen von Fachgesellschaften angewandt wird. Das stellt sicher, dass klinische Experten sich unabhängig von der Bewertung eines einzelnen Arzneimittels mit der klinischen Relevanz dieses Merkmals beschäftigt haben.

Positives Beispiel (Beispiel 1)

Ein Beispiel für berechtigte Subgruppenbildung in dieser Kategorie ist die Nutzenbewertung von Apixaban zur Prophylaxe thrombembolischer Komplikationen bei Risikopatienten mit nichtvalvulärem Vorhofflimmern [11]. Angemessene Vergleichstherapie ist die Behandlung mit Vitamin-K-Antagonisten (VKA). Eine klinisch relevante Subgruppe sind Patienten mit Kontraindikationen gegen Vitamin-K-Antagonisten. Für diese Patienten wurde zum Zeitpunkt der Planung der Zulassungsstudie eine Behandlung mit Acetylsalicylsäure (ASS) empfohlen. Für Apixaban wurde deshalb neben der Zulassungsstudie zum Vergleich Apixaban vs VKA (ARISTOTLE) eine zweite randomisierte Studie zum Vergleich Apixaban vs ASS bei Patienten mit Kontraindikationen gegen VKA (AVERROES) gestartet. Im Verfahren der frühen Nutzenbewertung nach dem AMNOG wurden diese zwei Subgruppen mit unterschiedlichen Vergleichstherapien festgelegt:

- Patienten, die für eine Therapie mit Vitamin-K-Antagonisten (VKA) geeignet sind: Vergleichstherapie VKA
- Patienten, die für eine Therapie mit Vitamin-K-Antagonisten nicht geeignet sind: Vergleichstherapie Acetylsalicylsäure (ASS)

Dieser Subgruppenbildung haben die wissenschaftlichen medizinischen Fachgesellschaften in ihren Stellungnahmen und in der Anhörung zugestimmt. In der Festlegung des Zusatznutzens durch den G-BA wurden die Patienten mit Kontraindikationen gegen VKA nicht aufgeführt. Dies beeinträchtigt aber die inhaltliche Berechtigung dieser Subgruppenbildung nicht.

Negatives Beispiel (Beispiel 2)

Beispiele für Subgruppenbildung, die die Kriterien dieser Kategorie nicht erfüllt, waren die Verfahren zu Abirateron [20] und Cabazitaxel [9], jeweils bei Patienten mit kastrationsresistentem Prostatakarzinom nach Vortherapie mit Docetaxel. Zu diesen beiden neuen Arzneimitteln gab es jeweils nur eine Zulassungsstudie. Abirateron wurde in der COU-AA-301-Studie gegenüber Placebo, Cabazitaxel in der TROPIC-Studie gegen Mitoxantron getestet. Der G-BA hat zwei Subgruppen mit unterschiedlichen Vergleichstherapien gebildet:

- Patienten, für die eine Docetaxel Retherapie infrage kommt
- Patienten, für die eine Docetaxel-Retherapie nicht infrage kommt.

Eine erneute Chemotherapie mit derselben Substanz kommt bei Patienten infrage, die bei der Vortherapie gut auf dieses Arzneimittel angesprochen haben und bei denen keine begrenzende Toxizität auftrat. Die Docetaxel-Retherapie ist jedoch kein Therapiestandard, sondern eine individuelle Entscheidung. Es liegen keine Daten randomisierter Studien zum Vergleich einer Docetaxel-Retherapie versus Best Supportive Care oder versus einer anderen Therapieform vor. Entsprechend ist sie auch in keiner Leitlinie, und explizit auch nicht in deutschen Leitlinien als Therapiestandard definiert.

Die stellungnehmenden Fachgesellschaften haben diese Subgruppenbildung kritisiert. In zwei späteren Verfahren zu den neuen Arzneimitteln Enzalutamid [13] und Radium-223 [15] in derselben Indikation wurde diese Subgruppe nicht mehr festgelegt, siehe auch [Anhang](#). Das spricht für das lernende System der frühen Nutzenbewertung, aber auch für eine verbindliche Festlegung von Subgruppenkriterien.

5.1.2. Unterschiedlicher Verlauf

5.1.2.1. Krankheitsverlauf

In jeder Erkrankung gibt es ein Kontinuum von Patienten mit sehr ungünstigen bis zu sehr günstigen Verläufen. Für viele Entitäten gibt es eine Vielzahl, manchmal fast eine Unzahl, prognostischer Marker. Eine Voraussetzung für die Akzeptanz als Subgruppenmerkmal ist ihre Validierung. Das betrifft auch die Zusammenfassung reproduzierbarer Merkmale in Form von Risiko-Scores.

Innerhalb der Nutzenbewertung ist die Subgruppenbildung aufgrund prognostischer Marker kritisch zu sehen. Prognostisch relevante, validierte und allgemein anerkannte Marker oder Scores sollen präspezifiziert und als Stratifikationsparameter genutzt werden. Die Zulassungsstudien hatten in keinem der bisher abgeschlossenen Verfahren ausreichende Fallzahlen, um signifikante Unterschiede in prognostisch relevanten Subgruppen nachzuweisen.

5.1.2.2. Ansprechen auf die Therapie

Beim Ansprechen kann es eine klare Entscheidung „ja“ oder „nein“ geben, aber auch ein Kontinuum von Patienten mit sehr ungünstigen bis zu sehr günstigen Verläufen und Heilungen. Das unterschiedliche Ansprechen auf eine neue Behandlung ist relevant für die Nutzenbewertung und entspricht den Vorgaben des Gesetzgebers. Auch hier ist zu berücksichtigen, dass die Zulassungsstudien in der Regel nicht die statistische Power haben, um signifikante Unterschiede in prognostisch relevanten Subgruppen nachzuweisen. Entsprechend müssen weitere Kriterien erfüllt sein, siehe [Kapitel 4.2](#), [Kapitel 4.3](#). und [Kapitel 4.5](#).

Positives Beispiel (Beispiel 3)

Ein Beispiel für berechtigte Subgruppenbildung in dieser Kategorie ist die Nutzenbewertung von Afatinib zur Therapie von Patienten mit fortgeschrittenem oder metastasiertem, nichtkleinzelligem Lungenkarzinom (NSCLC) mit Nachweis aktivierender Mutationen des Epidermal Growth Factor Receptor (EGFR) [14, 18]. Diese genetischen Aberrationen sind heterogen. In den relevanten Exonen des EGFR-Gens können Deletionen, Punktmutationen, Insertionen und Duplikationen vorkommen. Bisher sind drei Tyrosinkinase-Inhibitoren für die Therapie zugelassen. In einer Metaanalyse war plausibel gezeigt worden, dass Patienten mit der Aberration del19 besonders gut auf die gezielte Therapie ansprechen [26]. In der Zulassungsstudie zu Afatinib war die Art der genetischen Aberration präspezifiziert.

Tabelle 2: Subgruppenbildung bei Afatinib

Studienname / Erstautor / Jahr	Kontrolle	Neue Therapie	N ¹	ÜLZ ⁶ (HR ⁵)
LUX-Lung 3 [34] alle	Cisplatin Pemetrexed	Afatinib	345	23,5 vs 23,1 0,93 n. s. ⁹
LUX-Lung 3 [34] Del19	Cisplatin Pemetrexed	Afatinib	170	21,6 vs 32,6 0,55 p = 0,0059

Während Afatinib in der Studie LUX-Lung 3 insgesamt nicht zu einer statistisch signifikanten Verlängerung der medianen Überlebenszeit führte, war der Unterschied in der Subgruppe von Patienten mit del19 signifikant mit einem auch klinisch relevanten Unterschied. Dieser Effekt ist bemerkenswert, weil insgesamt 70% der Patienten im Chemotherapie-Arm einen EGFR-TKI als Folgetherapie erhielten. Die Daten zeigen, dass nicht nur die Therapie mit dem TKI an sich hochwirksam ist, sondern dass auch der Einsatz in der Erstlinientherapie indiziert ist.

Negatives Beispiel (Beispiel 4)

Die Verfahren zu neuen Arzneimitteln bei Patienten mit chronischer Hepatitis C sind Spitzenreiter bei der Zahl von Subgruppen im Rahmen der Nutzenbewertung. Patienten mit kompensierter Leberzirrhose aufgrund der chronischen Hepatitis C und Patienten mit HIV-Koinfektion sind klinisch relevante Subgruppen zum einen in Bezug auf das Krankheitsstadium, zum anderen in Bezug auf Komorbidität. In der Nutzenbewertung wurden Subgruppen für Patienten ohne Zirrhose und mit kompensierter Zirrhose u. a. in den Verfahren zu Daclatasvir, Ledipasvir/Sofosbuvir, Ombitasvir/Paritaprevir/Ritonavir/Dasabuvir und Simeprevir gebildet.

Allerdings wurde bisher nicht nachgewiesen, dass Patienten mit kompensierter Zirrhose schlechter auf die Therapie mit direkt antiviralen Substanzen ansprechen als als Patienten mit früheren Fibrosestadien. Die bisher publizierten Daten zeigen gleich hohe Raten eines anhaltenden virologischen Ansprechens (Sustained Virologic Response, SVR). Aufgrund dieser studienbasierten Erkenntnisse empfehlen die Fachgesellschaften deshalb z. B. Ledipasvir plus Sofosbuvir für Therapie-naive und -erfahrene Patienten mit HCV Typ 1- und 4-6-Infektion mit und ohne kompensierte Zirrhose [32]. Damit ist diese Subgruppenbildung nicht erforderlich.

5.2. Präspezifikation

Alle Subgruppenanalysen im Rahmen klinischer Studien sollen präspezifiziert sein [36]. Dies betrifft die exakte Beschreibung der jeweiligen Patientenpopulation, z. B. bei Analysen nach Geschlecht, und die Festlegung der Grenzwerte bei kontinuierlichen Variablen, z. B. bei Analysen nach dem Alter. Die Präspezifikation nimmt Faktoren auf, die einen nachgewiesenen oder vermuteten Einfluss auf den individuellen Therapieeffekt haben.

5.2.1. Primärer Endpunkt

Die Präspezifikation von Subgruppen orientiert sich in den meisten Studien am primären Endpunkt, kann aber auch – mit unterschiedlichen Merkmalen – auf andere Endpunkte, z. B. Nebenwirkungen gerichtet sein. In der frühen Nutzenbewertung wird diese Differenzierung nicht gemacht. Die Subgruppenbildung und die Festlegung der zweckmäßigen Vergleichstherapie betreffen alle vom G-BA als relevant akzeptierten Endpunkte.

Neben der Charakterisierung der Merkmale muss das Studienprotokoll auch Methoden der statistischen Auswertung enthalten.

Post-Hoc-Analysen haben grundsätzlich ein erhöhtes Verzerrungspotenzial [20]. Dies betrifft sowohl die nachträgliche Wahl von Variablen als auch die nachträgliche Festlegung von Grenz- oder Trennwerten. Grundsätzlich hat die extensive Post-Hoc-Analyse eine lange Tradition in der Medizin und einen hohen Stellenwert in der Identifikation relevanter Faktoren für Krankheits- und Therapieverläufe. Die aus solchen Analysen generierten Daten sind jedoch vorwiegend nur hypothesenbildend und nicht dazu angelegt, unterschiedliche Wirksamkeit oder Zusatznutzen zu belegen.

Eine neue Form der letzten Jahre sind sogenannte prospektiv-retrospektive Studien. Hintergrund sind die Fortschritte in der Identifikation von Biomarkern. Wenn diese parallel zum Verlauf der Studie oder zu einem späteren Zeitpunkt entwickelt werden, kann sie zu einer Subgruppenbildung post hoc (*a posteriori*) führen. Ein Beispiel ist die Entdeckung der Relevanz des Mutationsstatus der *KRAS*- (Kirsten rat sarcoma 2 viral oncogene) und der *NRAS*- (Neuroblastoma Rat Sarcoma oncogene) Gene für die Wirksamkeit der Anti-EGFR-Antikörper Cetuximab und Panitumumab beim metastasierten kolorektalen Karzinom. Die retrospektive Analyse des asservierten Tumormaterials in zwei großen randomisierten Cetuximab-Studien zeigte, dass nur die Patienten mit den unmutierten *RAS*-Genen (Wildtyp) auf die Behandlung mit den Antikörpern ansprechen, während Patienten mit Mutationen nicht ansprechen und möglicherweise sogar eine kürzere Überlebenszeit haben [39]. Diese Daten führten zur Änderung der Zulassung. Da das post hoc analysierte Untersuchungsmaterial vor Therapie und ergebnisoffen asserviert wurde, werten wir diese Art der Subgruppen auf der Basis von prospektiv-retrospektiven Studien als präspezifiziert. Das zusätzliche Kriterium als Stratifikationsmerkmal (siehe [Kapitel 4.2.3](#)) wird nicht erfüllt, ebenfalls meistens nicht das Kriterium der Festlegung von Auswertemethoden im Studienprotokoll. Eine zusätzliche, zukünftige Herausforderung ist die Integration von Methoden zur Nutzung kontinuierlicher Variablen aus prospektiv-retrospektiven Genom-Analysen zur Bildung sinnvoller Subgruppen [27].

5.2.2. Andere Endpunkte

In einigen Zulassungsstudien enthalten die Studienprotokolle unterschiedliche Präspezifikationen von Subgruppen in Bezug auf die unterschiedlichen Endpunkte, z. B. in Bezug auf Nebenwirkungen. Das ist bei der Bewertung von Subgruppen zu berücksichtigen. Zulassungsstudien haben nicht die notwendige Stichprobengröße, um signifikante Unterschiede für sekundäre Endpunkte zu zeigen. Wenn dazu noch Analysen von Subgruppen in Bezug auf diese Endpunkte kommen, haben die Analysen eher exploratorischen Charakter.

5.2.3. Stratifizierung

Die Präspezifikation relevanter Faktoren führt in vielen Studienprotokollen zu einer Stratifikation der Randomisierung nach den gewählten Faktoren, um eine ausgeglichene Verteilung in den Studienarmen zu gewährleisten. Eine exemplarisch transparente Darstellung der präspezifizierten Subgruppen findet sich im Dossier des pharmazeutischen Unternehmers zu Siltuximab, einem Antikörper zur Therapie des multizentrischen Morbus Castleman [16].

Tabelle 3: Subgruppenbildung bei Siltuximab [16]

Parameter	Kategorien	Stratifizierung
Geschlecht	weiblich männlich	nein
Ethnie	Kaukasier Nicht-Kaukasier	nein
Geographische Region	Nordamerika Europa Asien-Pazifik-Raum Latein-Amerika	nein
Alter	< 65 Jahre ≥ 65 Jahre	nein
Corticosteroid-Gebrauch zu Studienbeginn	ja nein	ja
MCD Histologie durch Unabhängige Bewertung	hyalin-vaskulärer Typ plasmazytärer Typ gemischter Typ	nein
Systemische Therapie vor Studienbeginn erhalten	ja nein	nein

Hier wird nicht nur transparent, welche Subgruppenmerkmale a priori festgelegt wurden, sondern auch ihr Einsatz als Stratifikationsmerkmal. Eine solche Transparenz fehlt in vielen Dossiers. Da zum Zeitpunkt der Nutzenbewertung nicht immer alle Informationen der Zulassungsstudien publiziert sind – und auch die bereits erschienenen Publikationen nicht alle Informationen vollständig enthalten – wäre eine solche Tabelle in jedem Dossier hilfreich.

5.3. Plausible Begründung

Die meisten Präspezifikationen von Subgruppen beruhen auf biologischen, pathophysiologischen, pharmakologischen, klinischen oder sozialen Gründen. In den bisherigen Fragen der frühen Nutzenbewertung wurde die Frage nach der Rationale kritisch diskutiert, wenn die Berechnungen aus Homogenitäts- bzw. Interaktionstests einen Unterschied im Ansprechen auf ein neues Arzneimittel anzeigten, ohne dass eine nachvollziehbare Rationale präsentiert wurde. In einer früheren Auswertung hatten wir 19 Verfahren identifiziert, in denen der Bericht des IQWiG die Bildung zusätzlicher Zusatzgruppen aufgrund positiver Interaktionstests und dem Hinweis auf Effektmodifikationen vorschlug [39]. Inhalte waren Krankheitsstadium, Komorbidität, Vortherapien und wiederholt Altersgrenzen. Mit Ausnahme des Verfahrens zu Afatinib (siehe [Kapitel 4.1.2.2](#)) wurden diese Vorschläge nicht in die Festlegungen des G-BA aufgenommen. Eine zentrale Rolle in Stellungnahmen, die die Vorschläge des IQWiG-Berichtes ablehnten, spielte das Fehlen von plausiblen Begründungen. Eine plausible Begründung allein reicht nicht zur Definition einer Subgruppe, wenn sie z. B. aus Subgruppen-Analysen von Phase-II-Studien stammen, nicht validiert oder und nicht in Leitlinien aufgenommen sind.

5.4. Inkonsistenz und Heterogenität

Zur Untersuchung der Homo- bzw. Heterogenität von Daten sollen Interaktions- bzw. Heterogenitätstests eingesetzt werden. In den wegweisenden Publikationen wird daraufhin gewiesen, dass zur Erklärung von Unterschieden biologische, soziale, klinische, geographische oder andere Gründe heranzuziehen sind, siehe [Kapitel 4.3.](#) Hintergrund sind die methodischen Einwände gegen Subgruppenanalysen:

- Subgruppenbildende Merkmale können arbiträr sein
- Subgruppenanalysen erreichen keine ausreichende Stichprobengröße.
- Mit zunehmender Zahl von Subgruppenanalysen steigt die Wahrscheinlichkeit, dass irgendein signifikantes Ergebnis erreicht wird, obwohl es sich in Wahrheit um ein zufälliges Ergebnis handelt [4, 38].

In Publikationen wird aber auch darauf hingewiesen, dass die Ergebnisse einer validen Interaktionsanalyse auch ohne vernünftige biologische Rationale kommuniziert werden sollen.

Als Signifikanzschwelle schlagen einige Autoren und die GRADE Working Group – vorsichtig – für Heterogenitätstest einen p-Wert bei 0,1 oder I^2 bei 30% als Schwelle fest [3, 7]. Studien zur Validierung eines bestimmten Grenzwertes mit Endpunkten einer Nutzenbewertung sind uns nicht bekannt. Aus Sicht wissenschaftlicher medizinischer Fachgesellschaft ist eine Kongruenz zwischen Prozessen der Leitlinienerstellung und anderer Verfahren wie Zulassung und Nutzenbewertung sinnvoll. Deshalb erscheint eine Signifikanzschwelle mit einem p-Wert bei 0,1 als vernünftig.

5.5. Patientenzahl

Jede klinische Studie nimmt eine gewisse Heterogenität in Kauf, um ausreichende Patientenzahlen für ein statistisch aussagekräftiges Ergebnis zu erlangen. Regelhaft haben Zulassungsstudien nicht die Stichprobengröße, um Unterschiede bzw. ggf. Nicht-Unterlegenheit in Subgruppen zu belegen.

Kritisch ist die Bildung sehr kleiner Subgruppen. Beispiel ist die zweite Subgruppe bei Axitinib, definiert durch Vorbehandlung mit einem Zytokin. Die Patientenzahl wurde auf 1% der Zielpopulation geschätzt, absolut 6 Patienten pro Jahr. Ein weiteres Beispiel ist die zweite Subgruppe bei Ledipasvir/Sofosbuvir, definiert als therapienaive Patienten mit Leberzirrhose und chronischer Infektion durch Hepatitis C, Genotyp 1 [16]. Für diese Subgruppe wurde ein relativer Anteil von 0,6% im Beschluss des G-BA angekommen, auch wenn diese Zahl fragwürdig ist. Ein anderes Beispiel für Schwierigkeiten beim Umgang mit Subgruppen war der Vorschlag eines unterschiedlichen Nutzens für Patienten <65 oder ≥65 Jahre im IQWiG Bericht bei der erneuten Nutzenbewertung von Vandetanib.

Die Feststellung einer Heterogenität zwischen den Altersgruppen beruhte im Vergleichsarm auf einer Subgruppe von 12 Patienten [11].

In der konkreten Behandlungssituation ist die Berücksichtigung dieser individuellen Faktoren relevant. In Leitlinien ist für solch kleine Patientengruppen oft kein Therapiestandard definiert.

Allein auf der Basis der Patientenzahl besteht darüber hinaus das Risiko einer Ungleichbehandlung von Patienten mit seltenen Erkrankungen. Zahlenmäßig liegen diese Subgruppen oft in der Häufigkeit von Erkrankungen mit Orphan-Drug-Status [39]. Dann gälte der Zusatznutzen als belegt und würde bei wenigen, nicht sicher aussagekräftigen Daten als nicht quantifizierbar festgelegt.

Ein grundlegendes Problem ist die Unsicherheit der Angaben zu den Patientenzahlen. Aufgrund des Fehlens flächendeckender und unabhängiger Register in vielen Indikationen sind die Schätzungen unsicher. Für eine zuverlässige Berechnung der Kosten sind Regeln für die Vorlage unabhängiger und sicherer Daten erforderlich.

5.6. Perspektiven

Die Subgruppenbildung im Rahmen der Nutzenbewertung kann auch strategische Ziele haben. So könnte die Strategie eines pharmazeutischen Unternehmens darauf abzielen, wenigstens mit einem „Anhaltspunkt für einen geringen Zusatznutzen“ für eine Subpopulation aus dem G-BA Verfahren zu kommen. Dies ermöglicht die Forderung nach einem Preis oberhalb der zweckmäßigen Vergleichstherapie in den sich anschließenden, vertraulichen Preisverhandlungen mit dem GKV-Spitzenverband. Der Umgang mit multiplen Endpunkten in multiplen Subgruppen fördert das Erreichen eines solchen Ziels. Das Abschwächen der statistischen Power durch Subgruppenbildung wirkt diesem Ziel entgegen. Letzteres wiederum kann ein strategisches Ziel von Krankenkassen sein. Die Forderung nach multiplen Subgruppen kann die Chance der Bildung von Patientenkollektiven ohne ausreichende Aussagekraft zur Festlegung eines Zusatznutzens eröffnen. Das Abschwächen der statistischen Power durch Subgruppenbildung wirkt diesem Ziel entgegen. Ob und unter welchen Umständen beide antagonistische Mechanismen zu einem realistischen Ergebnis gelangen ist im Moment völlig unklar.

Eine neue methodische Entwicklung umgeht klassische Subgruppenanalysen und orientiert sich an der Verwendung von Prädiktionsmodellen für das individuelle Therapieansprechen. Wie die Entwicklung eines prädiktiven Modells aufgrund von Daten eines RCTs erfolgen kann, wurde 2012 beschrieben [27]. Vickers et al. beschrieben eine Strategie zur Validierung des Effektes einer individuellen Therapiezuweisung unter Verwendung eines prädiktiven Modells im Vergleich zur Gabe der innovativen Substanz an alle [37]. Mit diesen Ansätzen werden viele Probleme, die im Rahmen komplexer Subgruppenanalysen aufgetreten sind, umgangen.

Statistische Strategien zur Entwicklung komplexer SGAs werden ebenfalls vermehrt vorgeschlagen [33]. Sie beinhalten die modellbasierte rekursive Aufteilung einer Menge von Biomarkern, um damit Algorithmen-basiert prädiktive Subgruppen von Patienten zu entdecken. Diese Verfahren kommen den von der EMA in ihrer Guideline gestellten Anforderungen sehr nahe, da sie eine unverzerrte Auswahl von prädiktiven Markern und eine statistische Fehlerkontrolle erlauben. Solche Verfahren reduzieren die Suche nach interessanten Patientengruppen und vermeiden komplexe Analysen, wie sich diese aus den momentanen Empfehlungen des IQWiG ergeben.

6. Fragenkatalog

6.1. Modell

Auf der Basis der obigen Überlegungen schlagen wir pragmatisch den folgenden Fragenkatalog vor, siehe [Tabelle 4](#).

Tabelle 4: Vorschlag für einen Fragenkatalog wissenschaftlicher medizinischer Fachgesellschaften zur Subgruppenbildung im Rahmen der frühen Nutzenbewertung

Kriterien	G-BA					pU		IQWiG	
	1	2	3	4	5	1	2	1	2
klinisch relevant · eigener Therapiestandard · unterschiedliche Wirksamkeit der Therapie									
Präspezifikation · für den primären Endpunkt · für andere Endpunkte · Stratifizierung oder separate Studie									
Rationale (plausible Begründung)									
positiver Homogenitäts- bzw. Interaktionstest									
relative Häufigkeit									

6.2. Beispiel

Tabelle 5: Fragenkatalog wissenschaftlicher medizinischer Fachgesellschaft zur Subgruppenbildung im Rahmen der frühen Nutzenbewertung von Dabrafenib (neues Anwendungsgebiet)

Kriterien für Subgruppen	G-BA			IQWiG	
	nicht vorbehandelt		vorbe- handelt	Frauen	Männer
	BRAF V600 positiv	BRAF V600 negativ			
klinisch relevant · eigener Therapiestandard in aktuellen Leitlinien · unterschiedliche Wirksamkeit der Therapie	■	■	■	■	■
Präspezifikation · für den primären Endpunkt · für andere Endpunkte · Stratifizierung oder separate Studie	■	■	■	■	■
Rationale (plausible Begründung)	■	■	■	■	■
positiver Homogenitäts- bzw. Interaktionstest	■	■	■	■	■
relative Häufigkeit	27	31	42	45	55

■ ja, zutreffend ■ nein, nicht zutreffend ■ keine Angaben

7. Literatur

1. Arzneimittel-Nutzenbewertungsverordnung vom 28. Dezember 2010 (BGBl. I S. 2324). <http://www.gesetze-im-internet.de/am-nutzenv/BJNR232400010.html>
2. Assmann SF, Pocock SJ, Enos LE, Kasten LE: Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 355: 1064-1069, 2000. PMID: 10744093
3. Brookes ST, Whitely E, Egger M et al.: Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol* 57:229-236, 2004. PMID: 15066682
4. Counsell CE, Clarke MJ, Slattery J, Sandercock PA: The miracle of DICE therapy for acute stroke: fact or fictional product of subgroup analysis? *BMJ* 309:1677-1681, 1994. PMID: 7819982
5. Cui L, Hung HMJ, Wang SJ, Tsong Y: Issues related to subgroup analysis in clinical trials. *J Biopharm Stat* 12: 347-358, 2002. PMID: 12448576
6. European Medicines Agency: Guideline on the investigation of subgroups in confirmatory clinical trials. EMA/CHMP/539146/2013. Draft 2014.
7. Feinstein AR: The problem of cogent subgroups: A clinicostatistical tragedy. *J Clin Epidemiol* 1998; 51: 297–299, 1998. PMID: 9539884
8. G-2 Gutachten: Aussagekraft von Subgruppenanalysen. Medizinischer Dienst der Spitzenverbände der Krankenkassen e.V.; 2004.
9. <http://www.g-ba.de/informationen/nutzenbewertung/10/>
10. <http://www.g-ba.de/informationen/nutzenbewertung/19/>
11. <http://www.g-ba.de/informationen/nutzenbewertung/31/>
12. <http://www.g-ba.de/informationen/nutzenbewertung/57/>
13. <https://www.g-ba.de/informationen/nutzenbewertung/76/>
14. <https://www.g-ba.de/informationen/nutzenbewertung/87/>
15. <https://www.g-ba.de/informationen/nutzenbewertung/96/>
16. <https://www.g-ba.de/informationen/nutzenbewertung/122/>
17. <https://www.g-ba.de/informationen/nutzenbewertung/150/>
18. <https://www.g-ba.de/informationen/nutzenbewertung/170/>
19. Gesetz zur Neuordnung des Arzneimittelmarktes in der gesetzlichen Krankenversicherung (Arzneimittelmarkneuordnungsgesetz – AMNOG), Bundesgesetzblatt Teil I vom 27. 12. 2010. [http://www.bgbl.de/Xaver/start.xav?startbk=Bundesanzeiger_BGBI&bk=Bundesanzeiger_BGBI&start=//*\[@attr_id=%27bgbl110s2262.pdf%27\]#__Bundesanzeiger_BGBI__%2F%2F%5B%40attr_id%3D%27bgbl110s2262.pdf%5D__1374326855765](http://www.bgbl.de/Xaver/start.xav?startbk=Bundesanzeiger_BGBI&bk=Bundesanzeiger_BGBI&start=//*[@attr_id=%27bgbl110s2262.pdf%27]#__Bundesanzeiger_BGBI__%2F%2F%5B%40attr_id%3D%27bgbl110s2262.pdf%5D__1374326855765)
20. Grouin JM, Coste M, Lewis J. Subgroup analyses in randomized clinical trials: statistical and regulatory issues. *J Biopharm Stat* 15: 869-882, 2005. PMID: 16078390
21. Guyatt G, Wyer P, Ioannidis J. When to believe a subgroup analysis. In: Guyatt G, et al, editors. *The users' guides to the medical literature: a manual for evidence-based clinical practice*. New York, NY: McGraw-Hill; 2008.
22. Guyatt G, Oxman AD, Akl E et al.: GRADE guidelines 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 64:383-394, 2013. DOI: 10.1016/j.jclinepi.2010.04.026
23. Guyatt GH, Oxman AD, Kunz R et al.: GRADE guidelines: 7. Rating the quality of evidence-inconsistence. *J Clin Epidemiol* 64: 1294-1302, 2013. DOI: 10.1016/j.jclinepi.2011.03.017
24. IQWiG: Allgemeine Methoden 4.0, 2011. https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf
25. IQWiG: Aktualisierung einiger Abschnitte der Allgemeinen Methoden 4.0 sowie neue Abschnitte zur Erstellung der Allgemeinen Methoden Version 4.1, Entwurf, 2013. https://www.iqwig.de/download/13-04-18_Entwurf_Aktualisierung_Allgemeine_Methoden_4-0.pdf
26. Lee CK, Wu YL, Ding PN et al.: Impact of specific Epidermal Growth Factor Receptor (EGFR) mutations and clinical characteristics on outcomes after treatment with EGFR tyrosine kinase inhibitors versus chemotherapy in EGFR-mutant lung cancer: a meta-analysis *J Clin Oncol* 17: 1958-1965, 2015. DOI: 10.1200/JCO.2014.58.1736
27. Matsui S, Simon R, Qu P et al.: Developing and validating continuous genomic signatures in randomized clinical trials for predictive medicine. *Clin Cancer Res* 18:6065-6073, 2012. DOI: 10.1158/1078-0432.CCR-12-1206

28. Oxman AD, Guyatt GH: A consumer's guide to subgroup analyses. *Ann Intern Med* 116:78-84, 1992. PMID: 1530753
29. Rasch A, Dintsios CM: Subgruppen in der frühen Nutzenbewertung von Arzneimitteln: eine methodische Bestandsaufnahme, *Z. Evid. Fortbild. Qual. Gesundheitswesen (ZEFQ)* 109:69-78, 2015. DOI: 10.1016/j.zefq.2015.01.001
30. Rothwell PM: Treating individuals 2: subgroup analysis in randomised controlled trials; importance, indications, and interpretation. *Lancet* 365: 176-186, 2005. PMID: 15639301
31. Rothwell PM, Mehta Z, Howard SC et al.: Treating individuals 3: from subgroups to individuals: general principles and the example of carotid endarterectomy. *Lancet* 365:256-265, 2005. PMID: 15652609
32. Sarrazin C, Berg T, Buggisch P et al.: Aktuelle Empfehlung zur Therapie der chronischen Hepatitis C, S3 guideline hepatitis C addendum. http://www.dgvs.de/fileadmin/user_upload/Leitlinien/Therapie_der_chronischen_Hepatitis_C/09.07.2015_Aktuelle_Empfehlung_zur_Therapie_der_chronischen_Hepatitis_C.pdf
33. Seibold H, Zeileis A, Hothorn T: Model-based recursive partitioning for subgroup analyses. *International Journal of Biostatistics*, 2015. Accepted 2015-10-02.
34. Sequist SV et al.: Phase III Study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR Mutations. *J Clin Oncol* 31:3327-3334, 2013. DOI: 10.1200/JCO.2012.44.2806
35. Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010; 340: c117. DOI: 10.1136/bmj.c117
36. Sun X, Ioannidis PA, Agoitsas T, Alba AC, Guyatt G: How to use a subgroup analysis. *Users' guides to the medical Literature. JAMA* 311:405-411, 2014. DOI:10.1001/jama.2013.285063
37. Vickers AJ, Kattan MW, Daniel S: Method for evaluating prediction models that apply the results of randomized trials to individual patients. *Trials* 8:14, 2007. PMID: 17550609
38. Wing LM, Reid CM, Ryan P, et al.: Second Australian National Blood Pressure Study Group. A comparison of outcomes with angiotensin-converting-enzyme inhibitors and diuretics for hypertension in the elderly. *N Engl J Med* 348:583-592, 2003. PMID: 12584366
39. Wörmann B et al.: Frühe Nutzenbewertung neuer Arzneimittel in Deutschland. Gesundheitspolitische Schriftenreihe der DGHO Band 6, 2014. <https://www.dgho.de/informationen/gesundheitspolitische-schriftenreihe/band-6>

Anhang: Subgruppen in Verfahren der frühen Nutzenbewertung 2011 – 2015

Arzneimittel ¹	Präparat ²	Indikation ^{3, 9}	Subgruppen		
			präspezi- fiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Abirateron	Zytiga®	Prostatakarzinom, CRPC	10	2	Docetaxel nicht indiziert
					Docetaxel Retherapie indiziert
Abirateron (n. A. ⁶)	Zytiga®	Prostatakarzinom, CRPC	8	1 ¹⁰	
Aclidiniumbromid	Bretaris®	COPD ₉	1 ¹⁰	2	ab Therapiestufe II
					ab Therapiestufe III/IV mit mehr als 2 Exazerbationen
Aclidiniumbromid/ Formoterol	Duaklir®/Brimi- ca®, Genuair®	COPD	4	4	Stufe II
					Stufe III
					Stufe IV
					Stufe III/IV mit mehr als 2 Exazerbationen/Jahr
Afatinib	Giotrif®	Lungenkarzinom, EGFR+, metastasiert	10	5	Erstlinie, guter AZ, Deletion 19
					Erstlinie, guter AZ, EGFR Mutation L858R
					Erstlinie, guter AZ, EGFR andere Aberrationen
					Erstlinie, reduzierter AZ (ECOG >2)
					Zweitlinientherapie nach Chemo- therapie
Afatinib	Giotrif®	Lungenkarzinom, EGFR+, metastasiert	10	5	Erstlinie, guter AZ, Deletion 19
					Erstlinie, guter AZ, EGFR Mutation L858R
					Erstlinie, guter AZ, EGFR andere Aberrationen
					Erstlinie, reduzierter AZ (ECOG >2)
					Zweitlinientherapie nach Chemo- therapie
Aflibercept	Eylea®	diabetisches Makulaödem	12	1	
Aflibercept	Zaltrap®	kolorektales Karzinom, metastasiert	10	1	
Aflibercept	Eylea®	Makuladegeneration	8	1	
Aflibercept	Eylea®	Makulaödem, retinaler Venen- verschluss	16	1	
Aflibercept	Eylea®	Makulaödem, Zentralvenenverschluss	14	1	
Albiglutid	Eperzan®	Diabetes mellitus Typ 2	8	5	Monotherapie, Metformin nicht indiziert
					Kombi mit Metformin
					Kombi mit einem anderem Antidiabetikum
					Kombi mit >2 anderen Antidiabetika
					Kombi mit Insulin
Aliskiren/ Amlodipin	Rasiamlo®	essenzielle Hypertonie	n.e.	1	
Alipogentipar- vovec	Glybera®	Hyperlipoproteinämie Typ I	n.e.	1	
Apixaban	Eliquis®	VTE ₇ nach Gelenkoperation	9	2	Initialbehandlung, bis zu 6 Monate
					Langzeitprophylaxe
Apixaban (n. A.)	Eliquis®	Vorhofflimmern	21	2	für Vitamin-K-Antagonisten geeignet
					nicht für Vitamin-K-Antagonisten geeignet

Arzneimittel ¹	Präparat ²	Indikation ^{3,9}	Subgruppen		
			präspezi- fiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Apremilast	Otezla®	Psoriasis	25	2	Arthritis, Vorbehandlung mit DMARD
					Plaque, nach anderer systemischer Therapie
Ataluren	Translarna™	Muskeldystrophie Duchenne	8	1	
Axitinib	Inlyta®	Nierenzellkarzinom, metastasiert	1	2	Vorbehandlung mit Sunitinib
					Vorbehandlung mit Zytokinen
Belatacept	Nulojix®	Nierentransplantation	2	2	Transplantat nach Standardkriterien
					Transplantat nach erweiterten Kriterien
Belimumab	Benlysta®	systemischer Lupus erythematoses (SLE)	1	1	
Boceprevir	Victrelis®	chronische Hepatitis C	18	2	therapienaiv, Genotyp 1
					therapieerfahren, Genotyp 1
Bosutinib	Bosulif®	Chronische myeloische Leukämie (CML)	1	1	
Brentuximab Vedotin	Adcetris®	Lymphom	13	3	CD30 positive Hodgkin Lymphome, nach ASZT
					CD30 positive Hodgkin Lymphome, keine ASZT
					anaplastische großzellige Lymphome
Bromfenac	Yellox®	Entzündung nach Augen-Operation	1	1	
Cabazitaxel	Jevtana®	Prostatakarzinom, CRPC	7	2	Docetaxel nicht indiziert
					Docetaxel Retherapie indiziert
Cabozantinib	Cometriq®	Schilddrüsenkarzinom, medullär	14	1	
Canagliflozin	Invokana®	Diabetes mellitus Typ 2	2	5	Monotherapie, Metformin nicht geeignet
					Kombi mit Metformin
					Kombi mit Sulfonylharnstoff
					Kombi mit > 2 anderen Antidiabetika
					Kombi mit Insulin
Canagliflozin/ Metformin	Vokanamet®	Diabetes mellitus Typ 2	1	3	Kombi mit Metformin
					Kombi mit anderen Antidiabetika
					Kombi mit Insulin
Ceritinib	Zykadia®	Lungenkarzinom, ALK+, metastasiert	11	2	Docetaxel oder Pemetrexed indiziert
					Docetaxel oder Pemetrexed nicht indiziert
Cholsäure	Orphacol®	Gallensäuresynthesestörung	1	1	
Cobicistat	Tybost®	HIV Infektion	2	1	
Colestilan	Bindren®	Hyperphosphatämie	4	2	kalziumhaltige Phosphatbinder geeignet
					kalziumhaltige Phosphatbinder nicht geeignet
Crizotinib	Xalkori®	Lungenkarzinom	3	2	Chemotherapie indiziert
					Chemotherapie nicht indiziert
Dabrafenib	Tafinlar®	Melanom, metastasiert	9	1	
Daclatasvir	Daklinza®	chronische Hepatitis C	5	7	therapienaiv, ohne Zirrhose, Genotyp 1
					therapienaiv, kompensierte Zirrhose, Genotyp 1
					therapieerfahren, Genotyp 2
					therapienaiv- und erfahren, ohne Zirrhose, Genotyp 3
					therapienaiv- und erfahren, ohne Zirrhose, Genotyp 4
					therapienaiv, Genotyp 4, Kombi mit IFN
					therapieerfahren, Genotyp 4, Kombi mit IFN/Ribavirin

Arzneimittel ¹	Präparat ²	Indikation ^{3, 9}	Subgruppen		
			präspezi- fiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Dapagliflozin	Forxiga®	Diabetes mellitus	2	4	Monotherapie, Metformin nicht indiziert
					Kombi mit Metformin
					Kombi mit anderen Antidiabetika
					Kombi mit Insulin
Dapagliflozin/ Metformin	Xigduo®	Diabetes mellitus	4	3	Kombi mit Metformin
					Kombi mit anderen Antidiabetika
					Kombi mit Insulin
Dasabuvir	Exviera®	chronische Hepatitis C			therapienaiv, ohne Zirrhose, Genotyp 1a/b
					therapienaiv, kompensierte Zirrhose, Genotyp 1a/b
					therapieerfahren, ohne Zirrhose, Genotyp 1 a/b
					therapieerfahren, kompensierte Zirrhose, Genotyp 1a/b
					therapienaiv- oder erfahren, HIV+, Genotyp 1 a/b
Decitabin	Dacogen®	Akute Myeloische Leukämie (AML)	6	1	
Dimethylfumarat	Tecfidera®	Multiple Sklerose	11	1	
Dolutegravir	Tivicay®	HIV Infektion ⁸	4	5	therapienaiv, Erwachsene
					therapienaiv, Jugendliche
					therapieerfahren, Erwachsene, Integrase-Inh. indiziert
					therapieerfahren, Erwachsene, Integrase-Inh. Nachrangig
					therapieerfahren, Jugendliche
Dolutegravir/Aba- cavir/Lamivudin	Triumeq®	HIV Infektion ⁸	4	5	therapienaiv, Erwachsene
					therapienaiv, Jugendliche
					therapieerfahren, Erwachsene, Integrase-Inh. indiziert
					therapieerfahren, Erwachsene, Integrase-Inh. Nachrangig
					therapieerfahren, Jugendliche
Dulaglutid	Trulicity®	Diabetes mellitus Typ 2	9	5	Monotherapie, Metformin nicht indiziert
					Zweifachkombi mit Metformin
					Zweifachkombi mit anderem oralen Antidiabetikum
					Dreifachkombi mit 2 oralen Antidiabetika
					Kombi mit Insulin
Eliglustat	Cerdelga®	Morbus Gaucher	2	1	
Elosulfase alpha	Vimizim®	Mukopolysaccharidose Typ IVA	5	1	
Elvitegravir/ Cobicistat/...	Stribild®	HIV Infektion	6	2	therapienaiv
					therapieerfahren, keine HIV Mutation, Resistenz
Empagliflozin	Jardiance®	Diabetes mellitus Typ 2	9	5	Monotherapie, Metformin nicht indiziert
					Zweifachkombi mit Metformin
					Zweifachkombi mit anderem oralen Antidiabetikum
					Kombi mit mindestens 2 oralen Antidiabetika
					Kombi mit Insulin

Arzneimittel ¹	Präparat ²	Indikation ^{3, 9}	Subgruppen		
			präspezi- fiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Emtricitabin/ Ralpirivir/T...	Eviplera®	HIV Infektion	n.a.	1	
Emtricitabin/ Ralpirivir/T... (n. A.)	Eviplera®	HIV Infektion	n.a.	1	
Enzalutamid	Xtandi®	Prostatakarzinom	13	1	
Enzalutamid (n. A.)	Xtandi®	Prostatakarzinom	13	1	
Eribulin	Halaven®	Mammakarzinom	16	2	Taxane oder Anthrazykline indiziert
					Taxane oder Anthrazykline nicht indiziert
Eribulin (nach Fristablauf, n. A.)	Halaven®	Mammakarzinom	16	3	Taxane oder Anthrazykline indiziert
					Taxane oder Anthrazykline nicht indiziert
					HER2 positiv, Anti-HER2-Therapie indiziert
Extrakt aus Cannabis sativa	Sativex®	Multiple Sklerose	1	1	
Fampridin	Fampyra®	Multiple Sklerose	1	1	
Fidaxomicin	Dificlir®	Clostridium difficile	11	2	milder Krankheitsverlauf
					schwerer und/oder rezurrenter Krankheitsverlauf
Fingolimod	Gilenya®	Multiple Sklerose	1	3	hochaktiv, Nicht-Ansprechen
					hochaktiv, nicht ausreichend mit IFN behandelt
					rasch fortschreitender, schwerer Verlauf
Fingolimod (nach Fristablauf)	Gilenya®	Multiple Sklerose	1	3	hochaktiv, Nicht-Ansprechen auf IFN
					hochaktiv, nicht ausreichend mit IFN behandelt
					rasch fortschreitender, schwerer Verlauf
Fingolimod (n. A.)	Gilenya®	Multiple Sklerose	1	2	hochaktiv, Nicht-Ansprechen auf Therapie (nicht IFN)
					hochaktiv, nicht ausreichend vorbehandelt (nicht IFN)
Fluticason/ Vilanterol	Relvar®, Ellipta®	Asthma, COPD	n.a.	1	
Ibrutinib	Imbruvica®	CLL, MZL	27	3	Mantelzell-Lymphom, rezidiviert/ refraktär
					CLL, mind. 1 Vortherapie
					CLL, Erstlinientherapie, del17p oder mutTP53
Idelalisib	Zydelig®	CLL, FL	21	6	CLL, Rezidiv, Chemo indiziert
					CLL, Rezidiv, Chemo nicht indiziert
					CLL, refraktär, Chemo oder Ofatumumab indiziert
					CLL, refraktär, Chemo oder Ofatumumab nicht indiziert
					CLL, Erstlinientherapie, del17p oder mutTP53
					Folikuläres Lymphom, refraktär nach 2 Vortherapien

Arzneimittel ¹	Präparat ²	Indikation ^{3,9}	Subgruppen		
			präspezi- fiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Indacaterol/ Glycopyrronium	Ultibro®, Breezhaler®	COPD			Stufe II
					Stufe III, höchstens 1 Exazerbation pro Jahr
					Stufe IV, höchstens 1 Exazerbation pro Jahr
					Stufe III/IV, >2 Exazerbationen pro Jahr
Ingenolmebutat	Picato®	aktinische Keratose	5	1	
Insulin deludec (n. A.)	Tresiba®	Diabetes mellitus	1	4	Typ 2, Monotherapie
					Typ 2, Kombi mit >1 oralem Antidiabetikum
					Typ 2, Kombi mit Bolusinsulin
					Typ 1
Insulin degludec (n. A.)	Tresiba®	Diabetes mellitus	n.a.	1	
Insulin degludec (n. A.)	Tresiba®	Diabetes mellitus, Kinder/Erwachsene	1	3	Typ 1
					Typ 2, Monotherapie
					Typ 2, Kombi mit anderen Antidiabetika
Insulin degludec/ Liraglutid	Xultophy®	Diabetes mellitus	1	3	Typ 2, Kombi, wenn orale Antidiabetika nicht ausreichen
					Typ 2, Kombi mit Metformin, wenn orale Antidiabetika und Basalinsulin nicht ausreichen
					Typ 2, Kombi mit anderen oralen Antidiabetika, wenn orale Antidiabetika und Basalinsulin nicht ausreichen
Ipilimumab	Yervoy®	Melanom, metastasiert	5	1	
Ipilimumab (n. A.)	Yervoy®	Melanom, metastasiert	3	2	BRAF-V600-Mutation negativ
					BRAF-V600-Mutation positiv
Ivacaftor	Kalydeco®	Zystische Fibrose	4	2	Kinder
					Jugendliche und Erwachsene
Ivacaftor (n. A.)	Kalydeco®	Zystische Fibrose	1	1	
Ivermectin	Soolantra®	Rosacea	n.a.	1	
lebende Larven von L. serrata	BioBag®, BioMonde®	Wundheilungsstörungen	n.a.	1	
Ledipasvir/ Sofosbuvir	Harvoni®	chronische Hepatitis C	13	7	therapienaiv, ohne Zirrhose, Genotyp 1
					therapienaiv, kompensierte Zirrhose, Genotyp 1
					therapieerfahren, Genotyp 1
					therapienaiv- und erfahren, Genotyp 3
					therapienaiv- und erfahren, Genotyp 4
					therapienaiv- und erfahren, HIV+, Genotyp 1
					dekompensierte Zirrhose, Genotyp 1
Lenvatinib	Lenvima®	Schilddrüsenkarzinom, papillär	n.a.	1	
Linaclotid	Constella®	Reizdarmsyndrom	2	1	
Linagliptin	Trajenta®	Diabetes mellitus Typ 2		3	Monotherapie
					Kombi mit Metformin
					Kombi mit Metformin und Sulfonylharnstoff
Linagliptin (5b ⁸)	Trajenta®	Diabetes mellitus Typ 2	10	2	Monotherapie
					Kombi mit Metformin

Arzneimittel ¹	Präparat ²	Indikation ^{3, 9}	Subgruppen		
			präspezi- fiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Linagliptin (n. A.)	Trajenta®	Diabetes mellitus, Typ 2	n.a.	1	
Lisdexamfetamin	Elvanse®	ADHS	3	1	
Lixisenatid	Lyxumia®	Diabetes mellitus Typ 2	1	4	Zweifachkombi mit Metformin
					Zweifachkombi mit anderem oralen Antidiabetikum
					Dreifachkombi mit oralen Antidiabetika
					Kombi mit Insulin
Lomitapid	Lojuxta®	Hypercholesterinämie	n.a.	2	medikamentöse und diätetische Optionen ausgeschöpft
					medikamentöse und diätetische Optionen nicht ausgeschöpft
Lomitapid (nach Fristablauf)	Lojuxta®	Hypercholesterinämie	n.a.	3	medikamentöse und diätetische Optionen ausgeschöpft, keine LDL-Apherese
					medikamentöse und diätetische Optionen ausgeschöpft, LDL-Apherese durchgeführt
					medikamentöse und diätetische Optionen nicht ausgeschöpft
Lurasidon	Latuda®	Schizophrenie	10	2	Akuttherapie
					Rückfallprophylaxe
Macitentan	Opsumit®	pulmonal arterielle Hypertonie	1	1	
Mikrobielle Collagenase	Xiapex®	Dupuytren'sche Kontraktur	n.a.	4	Stadium N
					Stadien N/1, I und II
					Stadien III und IV
					Stadien III und IV, KI für partielle Fasziektomie
Mirabegron	Betmiga™	überaktive Blase	2	1	
Nalmefen	Selincro®	Alkoholbedingte Erkrankungen	3	1	
Nepafenac	Nevanac®	Schmerz nach Augenoperation	n.a.	1	
Nintedanib	Vargatef®	Lungenkarzinom	10	1	
Nintedanib	Ofev®	idiopathische Lungenfibrose	8	1	
Obinutuzumab	Gazyvaro®	CLL	1	1	
Ocriplasmin	Jetrea®	vitreomakuläre Adhäsion	3	2	leichte Symptomatik
					schwere Symptomatik
Olaparib	Lynparza®	Ovarialkarzinom	9	1	
Olodaterol	Striverdi®, Respimat®	COPD	n.a.	1	
Ombitasvir/Paritaprevir/Ritonavir	Viekirax®	chronische Hepatitis C	5	7	therapienaiv, ohne Zirrhose, Genotyp 1a/b
					therapienaiv, kompensierte Zirrhose, Genotyp 1a/b
					therapieerfahren, ohne Zirrhose, Genotyp 1 a/b
					therapieerfahren, kompensierte Zirrhose, Genotyp 1a/b
					therapienaiv- oder erfahren, ohne Zirrhose, Genotyp 4
					therapienaiv- oder erfahren, kompensierte Zirrhose, Genotyp 4
					therapienaiv- oder erfahren, HIV+, Genotyp 1 a/b
Pasireotid	Signifor®	Hypophysendysfunktion	n.a.	1	
Pasireotid (n. A.)	Signifor®	Hypophysendysfunktion	n.a.	1	

Arzneimittel ¹	Präparat ²	Indikation ^{3, 9}	Subgruppen		
			präspezi- fiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Perampanel	Fycompa®	Epilepsie	3	1	
Perampanel (n. A.)	Fycompa®	Epilepsie	3	1	
Pertuzumab	Perjeta®	Mammakarzinom	10	3	viszerale Metastasierung
					nicht-viszerale Metastasierung
					lokal-rezidiert, inoperabel
Pirfenidon	Esbriet®	idiopathische pulmonale Fibrose	n.a.	1	
Pixantron	Pixuvri®	aggressive Lymphome	1	1	
Pomalidomid	Imnovid®	Multiples Myelom	5	1	
Ponatinib	Iclusig®	CML, Ph+ ALL	9	2	CML
					Ph+ ALL
Propranolol	Hemangirol®	infantiles Hämangiom	2	3	lebens- oder funktionsbedrohend
					ulzeriert, schmerzhaft
					Gefahr von bleibenden Narben oder Entstellung
Radium-223	Xofigo®	Prostatakarzinom	8	2	Docetaxel indiziert
					Docetaxel nicht indiziert
Ramucirumab	Cyramza®	Magenkarzinom	19	2	in Kombination mit Paclitaxel
					Monotherapie
Regadenoson	Rapiscan®	Myokardperfusion (Bildgebung)	n.a.	1	
Regorafenib	Stivarga®	kolorektales Karzinom	12	1	
Regorafenib (n. A.)	Stivarga®	Gastrointestinale Stromatumore	n.a.	1	
Retigabin	Trobalt®	Epilepsie	1	1	
Retigabin (erneute Nb. ⁷)	Trobalt®	Epilepsie	n.a.	1	
Rilpivirin	Edurant®	HIV Infektion ⁷	2	1	
Riociguat	Adempas®	pulmonale Hypertonie	n.a.	2	CTEPH
					PAH
Ruxolitinib	Jakavi®	primäre Myelofibrose	n.a.	1	
Ruxolitinib (erneute Nb. ⁷)	Jakavi®	primäre Myelofibrose	9	1	
Ruxolitinib (n. A.)	Jakavi®	Polycythaemia vera	8	1	
Safinamid	Xadago®	Parkinson	1	1	
Saxagliptin	Onglyza®	Diabetes mellitus	1	4	Zweifachkombi mit Metformin
					Zweifachkombi mit Sulfonylharnstoff
					Dreifachkombi mit Metformin und Sulfonylharnstoff
					Kombi mit Insulin
Saxagliptin/ Metformin	Komboglyze®	Diabetes mellitus	1	2	Zweifachkombi mit Metformin
					Dreifachkombi mit Insulin
Saxagliptin/ Metformin (n. A. ⁶)	Komboglyze®	Diabetes mellitus	1	1	
Secukinumab	Consentyx®	Plaque-Psoriasis	12	3	für systemische und/oder Phototherapie geeignet
					für systemische und/oder Phototherapie nicht geeignet, mit Biologika-Vorbehandlung
					für systemische und/oder Phototherapie nicht geeignet, ohne Biologika-Vorbehandlung
Siltuximab	Sylvant®	Multizentrischer Morbus Castleman	7	1	

Arzneimittel ¹	Präparat ²	Indikation ^{3, 9}	Subgruppen		
			präspezi- fiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Simeprevir	Olysio®	chronische Hepatitis C	7	7	therapienaiv, Genotyp 1
					therapieerfahren, Genotyp 1
					therapieerfahren, Non-Responder, Genotyp 1 a/b
					therapienaiv- oder erfahren, Genotyp 4
					therapieerfahren, Non-Responder, Genotyp 4
					therapienaiv- oder erfahren, HIV+, ohne Zirrhose, Genotyp 1, 4
					therapienaiv- oder erfahren, HIV+, mit Zirrhose, Genotyp 1, 4
Simoctocog alfa	Nuwiq®	Hämophilie A	1	1	
Sipuleucel-T	Provenge®	Prostatakarzinom, CRPC	27	1	
Sitagliptin	Xnuvia®, Xelevia®	Diabetes mellitus Typ 2	1	5	Monotherapie
					Zweifachkombi mit Metformin
					Zweifachkombi mit Sulfonylharnstoff
					Dreifachkombi mit Metformin und Sulfonylharnstoff
					Kombi mit Insulin
Sitagliptin/ Metformin	Xnumet®, Velmetia®	Diabetes mellitus Typ 2	1	3	Zweifachkombi mit Metformin
					Dreifachkombi mit Metformin und Sulfonylharnstoff
					Dreifachkombi mit Insulin
Sofosbuvir	Sovaldi®	chronische Hepatitis C	14	1	
Sucroferic	Velphoro®	Hypophosphatämie	7	2	kalziumhaltige Phosphatbinder indiziert
					kalziumhaltige Phosphatbinder nicht indiziert
Tafluprost/Timolol	Taptiqom®	Glaukom	5	1	
Tafamidis Meglumin	Vyndaqel®	Amyloidose	2	1	
Teduglutid	Revestive®	Malabsorptionssyndrom	n.a.	1	
Tegafur	Teysuno®	Magenkarzinom	n.a.	1	
Telaprevir	Incivo®	chronische Hepatitis C	7	2	therapienaiv, Genotyp 1
					therapieerfahren, Genotyp 1
Teriflunomid	Aubagio®	Multiple Sklerose	10	1	
Ticagrelor	Brilique®	Akutes Koronarsyndrom	19	4	NSTEMI
					STEMI, medikamentöse Therapie
					STEMI, perkutane Koronarintervention
					Myokardinfarkt
Trastuzumab Emtansin	Kadcyla®	Mammakarzinom	9	3	lokal fortgeschritten, inoperabel
					nach Trastuzumab, Taxanen und Anthrazyklinen
					nach Trastuzumab und Taxanen
Turoctocog alfa	NovoEight®	Hämophilie A	n.a.	1	
Umeclidinium/ Vilanterol	Anoro®, Laventair®	COPD	1	2	mittlere Schweregrade
					darüber hinaus gehende Schweregrade
Vandetanib	Caprelsa®	Schilddrüsenkarzinom, medullär	9	1	
Vandetanib (5b ⁸)	Caprelsa®	Schilddrüsenkarzinom, medullär	9	1	

Arzneimittel ¹	Präparat ²	Indikation ^{3,9}	Subgruppen		
			präspezifiziert ⁴	Gemeinsamer Bundesausschuss ⁵	
				n	n
Vedolizumab	Entyvio®	Colitis ulcerosa, M. Crohn	n.a.	4	Colitis ulcerosa, nach konventioneller Therapie
					Colitis ulcerosa, nach TNF-alpha-Antagonist
					Morbus Crohn, nach konventioneller Therapie
					Morbus Crohn, nach TNF-alpha-Antagonist
Vemurafenib	Zelboraf®	Melanom, metastasiert	9	1	
Vemurafenib (nach Fristablauf)	Zelboraf®	Melanom, metastasiert	9	1	
Vildagliptin	Galvus®, Xlra®, Xiliarx®	Diabetes mellitus Typ 2	4	5	Monotherapie
					Zweifachkombi mit Metformin
					Zweifachkombi mit Sulfonylharnstoff
					Dreifachkombi mit Metformin und Sulfonylharnstoff
					Kombi mit Insulin
Vildagliptin (erneute Nb.)	Galvus®, Xlra®, Xiliarx®	Diabetes mellitus Typ 2	5	1	
Vildagliptin/Metformin	Eucreas®, Ican-dra®, Zomarist®	Diabetes mellitus Typ 2	4	3	Zweifachkombi mit Metformin
					Dreifachkombi mit Metformin und Sulfonylharnstoff
					Dreifachkombi mit Insulin
Vismodegib	Erivedge®	Basalzellkarzinom	8	2	symptomatisch metastasiert
					lokal fortgeschritten
Vortioxetin	Brintellix®	Depression	4	3	leichte Episode
					mittelschwere Episode
					schwere Episode

Legende:

- 1 Arzneimittel – alphabetische Auflistung, Arzneimittel mit mehr als einem abgeschlossenen Verfahren sind getrennt aufgeführt
- 2 Präparatename – Name(n) der zugelassenen Präparate
- 3 Indikation – Kurzbeschreibung der Zulassungsindikation
- 4 präspezifiziert – Anzahl der im Dossier des pharmazeutischen Unternehmers (Modul 4, Abschnitt 4.2.5.5) aufgeführten Subgruppen; in diesem Abschnitt sind a priori geplante und im Studienprotokoll festgelegte Subgruppenanalysen für patientenrelevante Endpunkte darzustellen. Wenn nicht ausdrücklich anders erwähnt, sind wir davon ausgegangen, dass die genannten Subgruppenmerkmale präspezifiziert waren.
- 5 Gemeinsamer Bundesausschuss – Anzahl der Subgruppen und Kurzbeschreibung der Merkmale
- 6 n. A. – neues Anwendungsgebiet
- 7 erneute Nb. – erneute Nutzenbewertung
- 8 erneute Nutzenbewertung nach § 35a Abs. 5b SGB V
- 9 Diagnosen: ADHS – Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung, Chemo – Chemotherapie, CML – chronische myeloische Leukämie, COPD – chronisch obstruktive Lungenerkrankung, CTEPH – Chronisch thromboembolische pulmonale Hypertonie, DMARD – Disease Modifying Anti-Rheumatic Drugs, HIV – Humaner Immundefizienzvirus, Integrase-Inh. – Integrase-Inhibitor, Kombi – Kombinationstherapie, NSTEMI – Instabile Angina pectoris (IA)/Myokardinfarkt ohne ST-Strecken-Hebung, PAH – pulmonale arterielle Hypertonie, Ph+ ALL – Philadelphia Chromosom positive Akute Lymphatische Leukämie, SH – Sulfonylharnstoff, STEMI – Myokardinfarkt mit ST-Strecken-Hebung
- 10 Anzahl der Subgruppen – 1 bedeutet, dass die Indikation als Ganzes ohne weitere Subgruppenbildung bewertet wurde

ISBN 978-3-9816354-9-2