

AMNOG-Subgruppen versus wirtschaftliche Verordnungsweise aus Sicht der Industrie

Von Dr. Markus Frick | Verband forschender Arzneimittelhersteller

Das AMNOG läuft in der Praxis nicht rund. Dies gilt einerseits für zahlreiche methodische Fragen. Insbesondere die Aufteilung einer Zielpopulation in viele Subgruppen gehört zu einem großen Problemfeld. Oft wird zudem die bestverfügbare Evidenz aus der Zulassung wegen abweichender methodischer Vorgaben des G-BA überhaupt nicht ausgewertet. Andererseits zeigen sich grundsätzliche Probleme, die vor allem aus der zentralen Machtposition des GKV-Spitzenverbandes im gesamten Verfahren resultieren. Der G-BA splittet seine Bewertung nach Subgruppen auf, auf die der GKV-SV bei der Preisbildung durch „Mischpreise“ aufsetzen kann. Trotz dieser zentralen Preisregulierung wird auf regionaler Ebene anschließend die Wirtschaftlichkeit von Erstattungsbeträgen bezweifelt und Ärzten mit Regress gedroht.

In practice, the AMNOG does not run smoothly, in particular with regard to numerous methodological questions. Slicing of target populations into various subgroups constitutes a major issue in this respect. Moreover, the best available evidence is frequently not even evaluated because study methodology deviates from the specifications of the G-BA. Additionally, fundamental problems arise from the dominant position of the GKV-Spitzenverband (GKV-SV) in the overall process. The G-BA appraisal is divided into subgroups, which may be used by the GKV-SV to justify mixed prices during price negotiations. Moreover, despite this central mechanism of price regulation, the economic efficiency of reimbursement amounts is subsequently questioned on a regional level and physicians are faced with the risk of financial penalties.

Perzeption des AMNOG
Mit dem Arzneimittelmarktneuordnungsgesetz (AMNOG) hat der Gesetzgeber im Jahr 2010 eine Neuregelung der Erstattungsbedingungen für innovative Arzneimittel vorgenommen. Damit wurde ein zweistufiges Bewertungs- und Erstattungsverfahren eingeführt. Die eigentliche Bewertung der von pharmazeutischen Unternehmen eingereichten Unterlagen (Dossiers) übernimmt in den meisten Fällen das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Der G-BA beschließt dann über den jeweiligen Zusatznutzen. Anschließend vereinbart der einzelne pharmazeutische Unternehmer innerhalb von weiteren sechs Monaten mit dem GKV-Spitzenverband (GKV-SV) für sein Produkt einen Erstattungsbetrag. Finden die Verhandlungsparteien zu keiner Einigung, legt eine Schiedsstelle den Erstattungsbetrag zwölf Monate nach der Markteinführung innerhalb von drei Monaten rückwirkend fest.

Inzwischen befinden sich die am AMNOG beteiligten Akteure im fünften Jahr seit Inkrafttreten des Gesetzes. Seit Januar 2011 wurden ca. 150 Verfahren der frühen Nutzenbewertung gestartet, darunter die ersten Verfahren für neue Anwendungsgebiete bereits bewerteter Produkte. Etwa 120 Nutzenbewertungsverfahren sind bis heute abgeschlossen worden. In der zweiten AMNOG-Phase sind in über 75 Fällen Erstattungsbeträge bestimmt worden. Weitere Produkte wurden in das Festbetragssystem eingruppiert oder aber vom Hersteller vom deutschen Markt zurückgezogen.

Mediale Rezeption des AMNOG hat sich geändert

Folgt man den offiziellen Verlautbarungen des G-BA und des GKV-SV, so erscheint das AMNOG weitgehend friktionsfrei. Die zentrale Deutung sowie die mediale Wahrnehmung der AMNOG-Ergebnisse änderten sich jedoch im zeitlichen Ablauf. Während zunächst die Botschaft im Mittelpunkt stand,

dass das AMNOG funktioniere und die Hürde des Zusatznutzens nicht zu hoch geschraubt sei („AMNOG funktioniert“, „Bei 64 Prozent der Arzneimittel konnte ein Zusatznutzen festgestellt werden“), veränderte sich diese Botschaft nach erfolgter Etablierung des AMNOG und es wurde relativierend festgestellt, dass ein Großteil des Zusatznutzens oft nur Subgruppen betreffe oder vom Ausmaß nur „gering“ sei.

Dies spiegelt sich in der medialen Rezeption, die bei nahezu unveränderten Nutzenbewertungsergebnissen ganz unterschiedliche Deutungen reflektiert. Noch im Jahr 2012 hieß es bei SPIEGEL Online beispielsweise: „Neu-Medikamente nützlicher als erwartet“ (Spiegel-online 2012). In 2014 lautete die Deutung derselben Autorin: „Nur jedes fünfte Arzneimittel bringt mehr Nutzen“ (Spiegel-online 2014). Dabei lag der Anteil von Produkten mit vom G-BA zugesprochenem Zusatznutzen bei 36 Prozent (Sep-

tember 2012) und 41 Prozent (Mai 2014) weitgehend stabil. Der Grund lag darin, dass im neueren Artikel allein auf den „beträchtlichen“ Zusatznutzen abgestellt wurde und „geringer“ oder „nicht quantifizierbarer“ Zusatznutzen nicht als Zusatznutzen wahrgenommen wurde.

Dass sich hinter der Zuweisung eines „geringen“ Zusatznutzens durchaus relevanter Zusatznutzen für Patienten verbergen kann, zeigen beispielhaft zwei Bewertungen in der Abbildung 1. Im erstgenannten Beispiel (Abbildung links) geht es um Lebensverlängerung bzw. Senkung der Mortalität. Mit Verweis auf Nebenwirkungen wurde dieser Zusatznutzen als lediglich „gering“ eingestuft (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen 2012). Im zweiten Beispiel wird diese Tendenz noch deutlicher, weil eine Senkung der Sterblichkeit hier nicht durch ein mögliches Mehr an Nebenwirkungen erkaufte wurde (Abbildung rechts) Hier führen positive Effekte in Bezug auf Gesamtmortalität, Schlaganfälle, Blutungen, unerwünschte Ereignisse und schwere unerwünschte Ereignisse trotz fehlender negativer Effekte ebenfalls aus Sicht des G-BA zu einem lediglich „geringen“ Zusatznutzen (Gemeinsamer Bundesausschuss 2013).



Dr. med. Markus Frick MPH ist seit September 2012 als Geschäftsführer Markt und Erstattung im vfa tätig. Der Mediziner arbeitete mehrere Jahre im Institut für Transplantationsdiagnostik und Zelltherapeutika der Medizinischen Einrichtungen der Heinrich-Heine Universität und wechselte 2001 als Leiter der Abteilung Medizin zur Westdeutschen Brust-Centrum GmbH. 2003 trat er bei der Aventis Pharma Deutschland GmbH ein und bekleidete dort verschiedene Funktionen; zuletzt leitete er die Abteilung Market Access.

Subpopulationen und Subgruppen im AMNOG

Eine Auswertung aller vorliegenden Beschlüsse des G-BA ergibt insgesamt ein differenziertes Bild. So attestierte der G-BA in etwa 53 Prozent der Verfahren einen Zusatznutzen. Dieser Zusatznutzen wurde allerdings regelmäßig nur in Teilen eines Anwendungsgebietes beschieden. Auf der Ebene aller dabei beschlossenen Subpopulationen sah der G-BA nur in etwa 36 Prozent der Fälle einen Zusatznutzen. Für die anschließenden Erstattungsbetragsverhandlungen im Sinne eines „Mischpreises“ ist zudem auch die Prävalenzebene relevant, in der die jeweiligen Subpopulationen nach Anzahl der behandelbaren Patienten gewichtet werden. Hier erweist sich die Bewertungspraxis des G-BA als noch konservativer:

IQWiG-Dossierbewertung A11-24

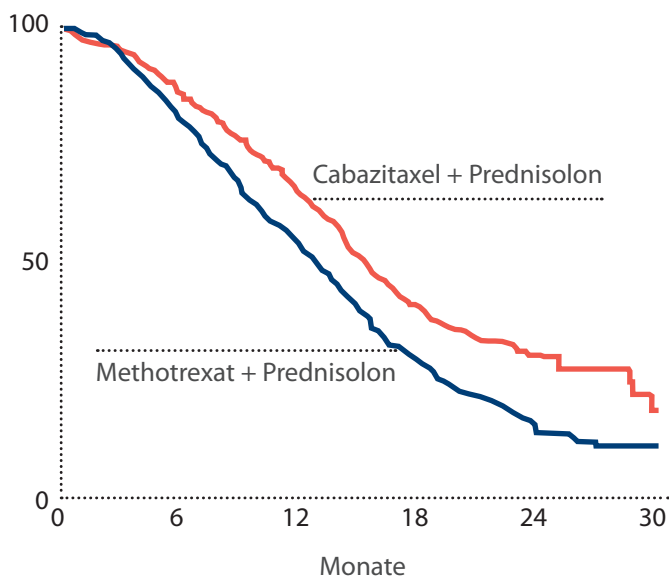


Abb. 1: Beispiele für einen vom G-BA zugewiesenen geringen Zusatznutzen: Dahinter kann sich ein für die Patienten durchaus relevanter Zusatznutzen verbergen.

In den bisher abgeschlossenen Verfahren sprach der G-BA nur für unter 20 Prozent der populationsgewichteten Teilpopulationen einen Zusatznutzen aus. In abgeschlossenen Verfahren wurde das zugelassene Anwendungsgebiet in den Beschlüssen des G-BA in etwa 47 Prozent der Fälle aufgeteilt. Beim IQWiG zeigt sich eine deutlichere Tendenz zur Bewertung nach Teilmengen. Hier wurden in rund 65 Prozent mindestens zwei Teilpopulationen berichtet. Das Maximum erreicht hierbei die Bewertung zu Sitagliptin mit insgesamt zehn Einzelempfehlungen zum Zusatznutzen. Der einzelnen Betrachtung solcher Teile der Anwendungsgebiete kommt damit eine besondere Bedeutung zu (vgl. Abbildung 2).

Es ist zunächst auf die vom Gesetzgeber verwendete Beschreibung hinzuweisen, die zugleich auch als rechtliche Begründung für die gebildeten Teilmengen sowie durchzu-

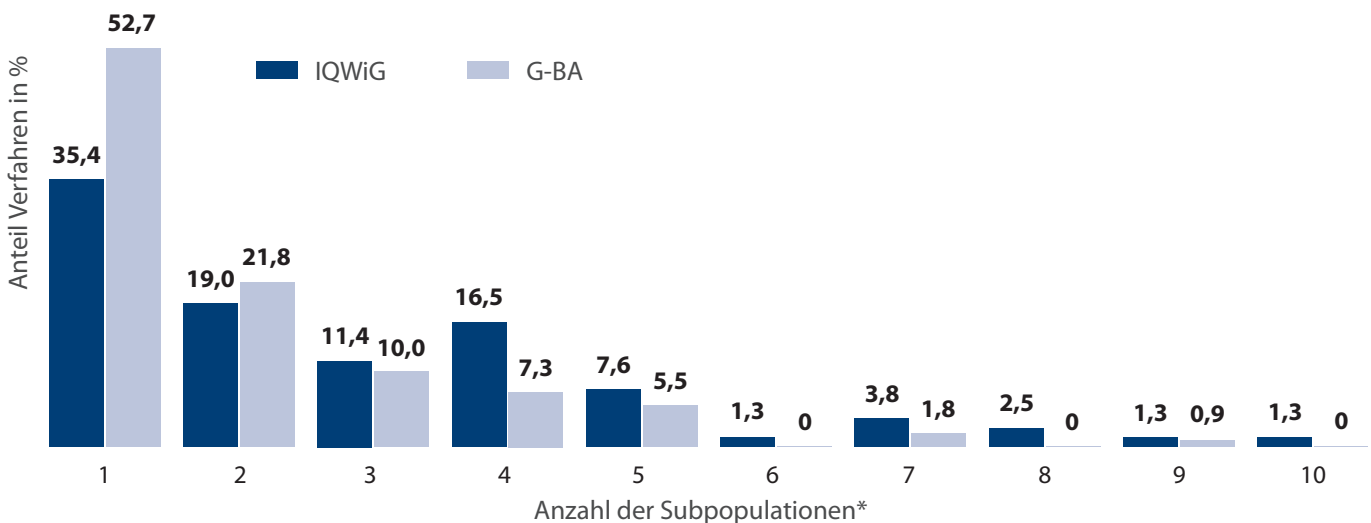
Darstellung nach G-BA Apixaban in VHF

Positive Effekte	Negative Effekte
Gesamtmortalität	keine
Schlaganfall (ischämisch, hämorrhagisch oder unbekannte Ursache)	keine
Schlaganfall (hämorrhagisch)	keine
größere Blutung, klinisch relevant nicht größere Blutung	keine
Gesamte Unerwünschte Ereignisse	keine
Therapieabbrüche wegen SUE	keine
Schlaganfall, SE, größere Blutungen oder Mortalität	keine

führende Analysen dient und im Paragraph 35a Abs. 1 Satz 3 Nr. 4 SGB V enthalten ist. Gesetzliche Grundlage für die Aufteilung der Bewertung in Teilmengen ist die Feststellung des AMNOG, dass die Unterlagen der Hersteller unter anderem Angaben über „Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht“ enthalten sollen. Tatsächlich gibt es eine potenziell infinite Anzahl möglicher Subgruppen und die Definition und Auswahl der tatsächlich verwendeten Subgruppen erfolgt nicht im Dossier, sondern im Bewertungsprozess des G-BA. Nach Paragraph 7 Abs. 2 der AM-NutzenV (Arzneimittel-Nutzenbewertungsverordnung) wird zudem geprüft, „welcher Zusatznutzen für welche Patientengruppen in welchem Ausmaß belegt ist“.

Insgesamt kann zwischen Subpopulationen und Subgrup-

Tendenz zur Bewertung von Teilmengen ist beim IQWiG besonders ausgeprägt



* Datenbasis: 110 abgeschlossene Verfahren durch G-BA, 79 Nutzenbewertungen durch IQWiG

Abb. 2: Verteilung der Subpopulationen und Subgruppen im AMNOG durch IQWiG und Gemeinsamen Bundesausschuss (Stand: 4. 3. 2015).

pen unterschieden werden. Als Subpopulationen werden im Folgenden zunächst solche Patientengruppen bezeichnet, die sich aus dem Wortlaut des zugelassenen Anwendungsgebietes ergeben (z.B. Schweregrade einer Erkrankung oder sequenzielle Therapiestufen) (Rasch und Dintsios 2015). Auch die vom G-BA je Subpopulation abgeleiteten unterschiedlichen zweckmäßigen Vergleichstherapien können eine Aufteilung begründen. Als Subgruppen werden im Weiteren solche Teilpopulationen einer Studie bezeichnet, die sich aufgrund einer potenziellen Effektmodifikation ergeben. Hierzu hat der G-BA durch seine Verfahrensordnung eine Vielzahl unterschiedlicher Analysen vorgesehen und dies in jüngster Zeit noch um Subgruppenanalysen nach Länder- und Zentrumseffekten erweitert.

Insbesondere die Zerlegung einer Zielpopulation im Rah-

men von Subgruppenanalysen gehört zu einem regelmäßig diskutierten methodischen Problemfeld (Rothwell 2005). Das dabei praktizierte formal-abstrakte Vorgehen ist auch aufgrund eines hohen Fehlerpotenzials überaus kritisch zu sehen. Es lassen sich dabei drei wesentliche Problemfelder identifizieren:

Erstens stellen a priori gebildete Subgruppen mit einer stratifizierten Randomisierung und anschließenden Interaktionstests den Goldstandard dar.

Zweitens ergibt sich ein wesentliches methodisches Problem durch das multiple Testen von mehreren Subgruppenmerkmalen. Bei genügend vielen Tests kann dies sowohl zu einer falsch-positiven Annahme (Fehler 1. Art) als auch zu einer falsch-negativen Ablehnung (Fehler 2. Art) eines therapeutischen Effekts in einer Subgruppe führen.

Drittens ergibt sich insbesondere auf der Studienebene das Problem eines Verlustes an statistischer „Power“, so dass das Risiko steigt, tatsächlich vorhandenen Nutzen nicht mehr nachzuweisen.

Zusammenfassend sind solche Analysen nur dann ausreichend valide, wenn die Anzahl präspezifizierter Subgruppen gering bleibt, die Effekte groß, konsistent und auch aus klinischer Sicht plausibel. Als große Problemfelder der Subgruppenanalytik benennt das Methodenpapier den fehlenden Beweischarakter von Subgruppenanalysen, das Risiko falscher positiver Resultate durch Multiplizität und das Risiko falscher negativer Ergebnisse durch Verringerung der Studienpower.

Auswirkungen der geänderten IQWiG-Methodik

Zentral bei der Subgruppenanalytik ist ferner, dass die ausgewählten Subgruppen sich statistisch überhaupt unterscheiden, was durch sogenannte Interaktionstests zu prüfen ist. An dieser Stelle ist anzumerken, dass die aktuelle IQWiG-Methodik das Signifikanzniveau der Interaktionstests gesenkt hat, so dass schon bei einem p-Wert ab 0,2 von einem Hinweis auf Signifikanz ausgegangen wird und entsprechende Subgruppen getrennt analysiert werden. Dadurch besteht ein höheres Risiko, dass tatsächlich nicht unterschiedliche Gruppen getrennt bewertet werden. Grundsätzlich ist die erkenntnistheoretische Aussagekraft einer Subgruppenanalyse geringer als die der Primäranalyse, so dass sogar das IQWiG in seinem Methodenpapier fordert, dass sie „nicht das Ergebnis der primären Analyse dominieren [dürften]“. Tatsächlich findet aber genau dies regelmäßig statt, wenn eine Bewertung nur für Subgruppen erfolgt, ohne die Gesamtbewertung zum primären Bewertungsgegenstand zu machen.

Eine Ausnahme von den oben genannten Regeln wird zudem dann gesehen, „wenn sozialrechtliche Implikationen solche Analysen erforderlich machen“, insbesondere weil im Paragraf 139 a Abs. 2 SGB V die Berücksichtigung von alters-,

geschlechts- und lebenslagenspezifischen Besonderheiten erwähnt ist, sowie im Kontext der frühen Nutzenbewertung auch von Patientengruppen die Rede sei (IQWiG 2013). Hinsichtlich der bisherigen Praxis einer exzessiven Subgruppenaufteilung der Zulassungspopulationen stellt sich jedoch die Frage, ob die erkenntnistheoretischen Gesetzmäßigkeiten und die wissenschaftliche Validität tatsächlich durch sozialrechtliche Gesetzestexte außer Kraft gesetzt werden können. Vielmehr ist davon auszugehen, dass die Intention des Gesetzgebers trotz der beispielhaft genannten Merkmale auf die Einhaltung der Standards der Evidenzbasierten Medizin in der Handhabung von Subpopulationen abzielt. Die formal-abstrakte Gesamtkonstruktion zur Ableitung des Zusatznutzenausmaßes anhand des Konfidenzintervalls erweist sich ebenfalls als umso problematischer und konservativer, je extensiver eine Studienpopulation in kleinere Subgruppen aufgeteilt wird. Zudem suggeriert eine Differenzierung in mehrere kleinere Teilpopulationen bei einer unreflektierten Betrachtung eine Pseudogenauigkeit.

Die zahlreichen methodischen Limitationen bleiben auch bei schwachen Hinweisen auf Effektmodifikation oder bei sonstigen Aufteilungen systemseitig kaum diskutiert und fungieren axiomatisch als selbstverständliche und scheinbar unangefochtene Eckpfeiler. Dabei zeigt sich eine gewisse Ambivalenz, da einerseits im Methodenpapier des IQWiG zahlreiche Limitationen von Subgruppenanalysen sowie Standards zur Interpretation publizierter Ergebnisse beschrieben werden. Andererseits scheinen diese Standards weniger zu greifen, wenn es sich um obligatorisch vorzulegende oder vom IQWiG oder vom G-BA selbst interpretierte Subgruppenanalysen handelt (Rasch und Dintsios 2015).

Ein der dargestellten Subgruppenproblematik verwandtes Problem ist, dass die „negativen“ Beschlüsse des G-BA sich weit überwiegend nicht aus einer detaillierten Studienauswertung ergeben, sondern aufgrund von Inkongruenzen

zwischen den Bewertungsmaßstäben des G-BA und der bestverfügbaren Evidenz nach der Zulassung (wissenschaftliche Erkenntnisse aus Studien). Eingereichte Zulassungsstudien werden in vielen Fällen von vornherein als nicht bewertungsrelevant eingestuft, so dass wegen dieser Inkongruenz zwischen Anforderungen der Zulassung und des IQWiG bzw. G-BA kein Zusatznutzen festgestellt wird, ohne dass die Evidenz im Einzelnen betrachtet wird.

Verfügbare Evidenz nicht detailliert ausgewertet

Analysiert man die Beschlüsse ohne Zusatznutzen, so zeigt sich, dass in mehr als der Hälfte der Fälle die verfügbare Evidenz nicht detailliert ausgewertet wurde. Bei einer Betrachtung von negativ bewerteten Subpopulationen steigt dieser Anteil nochmals deutlich (siehe Abbildung 3). Trotz regelhaft vorliegender randomisiert-kontrollierter Studien, sah der G-BA relevante Abweichungen im Regime der Vergleichstherapie, im Therapieregime oder in der untersuchten Population, die zum „formalen Scheitern“ des Produkts in der frühen Nutzenbewertung führten.

Die idealtypischen Anforderungen der Nutzenbewertung können damit in der Praxis oft kaum erfüllt werden, zumal eine Diskussion der Übertragbarkeit der Studienergebnisse im Sinne eines Analogieschlusses im Rahmen der Nutzenbewertung nur in Einzelfällen vorgenommen wird. Hier gilt es insbesondere zu hinterfragen, inwiefern eine Harmonisierung mit den Anforderungen der Zulassung bestärkt werden sollte, um die beste verfügbare Evidenz nicht verloren zu geben. Ebenso gilt dies für eine rechtzeitige Einbindung der Zulassungsbehörden sowie der Fachgesellschaften in den gesamten Prozess. Zum gegenwärtigen Zeitpunkt existiert hier offensichtlicher Verbesserungsbedarf.

Grundsätzlich führt eine Aufteilung des Anwendungsbereiches in einzelne Teilpopulationen insbesondere bei unterschiedlichen Ausprägungen des Zusatznutzens auch zu einer

Gründe für „formales Scheitern“ im AMNOG

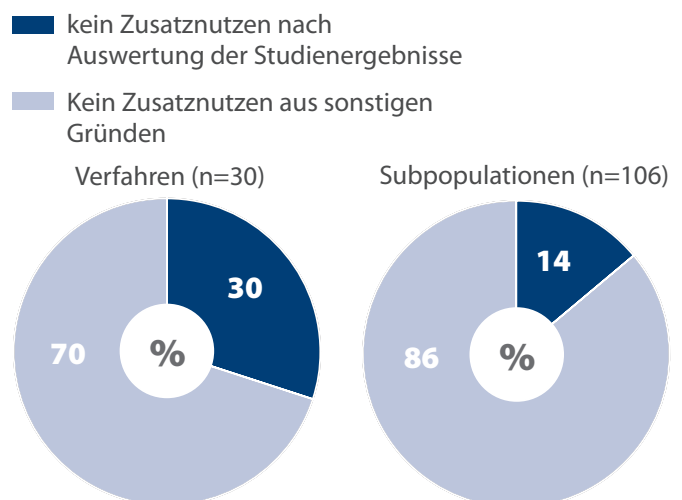


Abb. 3: Idealtypische Anforderungen der Nutzenbewertung und ihre Folgen.

paradigmatischen Mischkalkulation im Rahmen von Erstattungsbetragsverhandlungen. Dies lässt sich an einem fiktiven Beispiel verdeutlichen. Bei einem Produkt mit zwei Subpopulationen mit jeweils gleichen Patientenzahlen und einer durch die Verhandlungspartner durchgeführten monetären Übersetzung des Zusatznutzens in einerseits 10 und andererseits 30 Euro, ergäbe sich ein produktbezogener Erstattungsbetrag von 20 Euro. Damit subventionieren Verordnungen in Teilpopulationen mit einem höheren Zusatznutzen solche mit einem relativ niedrigen Zusatznutzen, da das AMNOG keine indikationenbasierten oder subgruppenspezifischen Erstattungsbeträge kennt. Für die Versorgungspraxis ergibt sich daraus dann ein Dilemma, wenn trotz einer Mischkalkulation insbesondere in Subpopulationen ohne festgestellten Zusatznutzen eine Verordnung durch Krankenkassen oder Kassenärztliche Vereinigungen als „unwirtschaftlich“ angeprangert und mit Regress bedroht wird.

Anteil am zu erwartenden Verbrauch in %

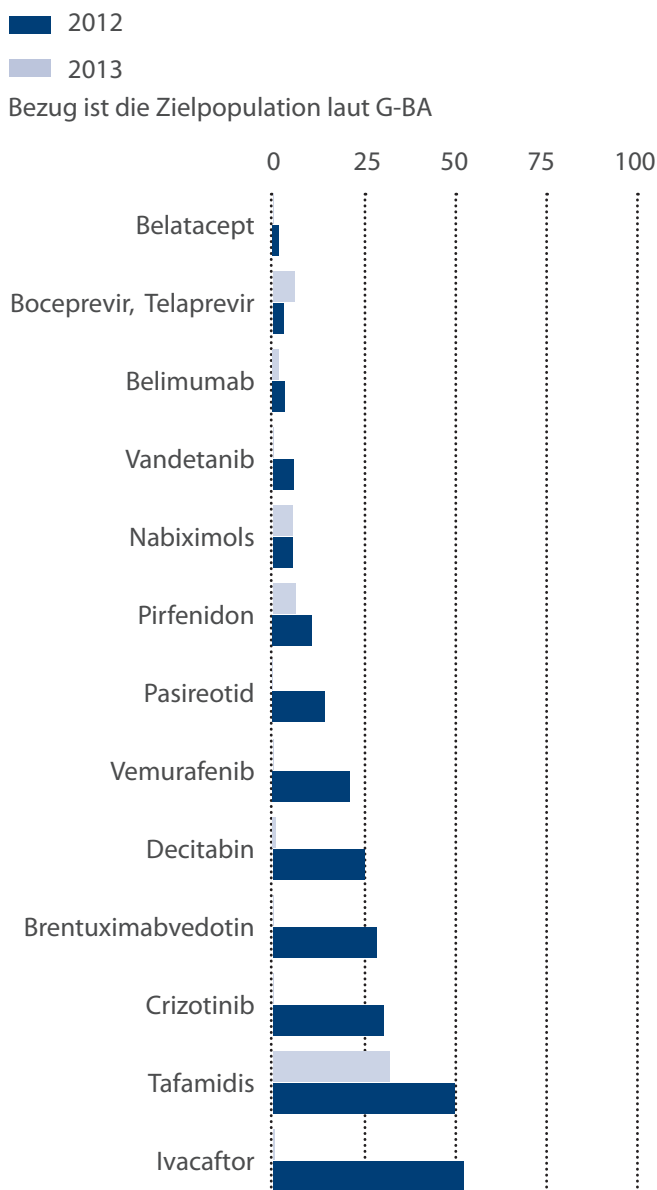


Abb. 4: Trotz attestiertem Zusatznutzen werden neue Arzneimittel nur unzureichend im Verordnungsalltag eingesetzt. Quelle: eigene Darstellung nach Häussler et. al 2014

Gegenläufige Wirkung durch Quotenregelungen

Hinzu kommt, dass die gesetzliche „Soll-Vorgabe“, bundesweite Praxisbesonderheiten für AMNOG-Produkte zu vereinbaren, tatsächlich keine Anwendung findet. Evident wird dieser regulatorische Antagonismus zwischen AMNOG und dezentraler Wirtschaftlichkeitsprüfung, wenn einerseits vom G-BA ein Zusatznutzen attestiert wird, andererseits für dasselbe Arzneimittel begrenzende Quotenregelungen durch Arzneimittelvereinbarungen nach Paragraph 84 Abs. 1 SGB V getroffen werden und gefordert wird, die Mehrzahl der Patienten mit der unterlegenen Vergleichstherapie zu behandeln.

Hinsichtlich der Verordnungszahlen ist insgesamt zu erkennen, dass innovative Arzneimittel trotz attestiertem Zusatznutzen nur unzureichend im Versorgungsalltag eingesetzt werden. Verdeutlicht wird dies durch Abbildung 4. Um eine Aussageverzerrung durch mögliche Alternativtherapien zu vermeiden, werden hier nur die AMNOG-Produkte dargestellt, bei denen es keine Überschneidung in der Zielpopulation mit anderen Wirkstoffen gibt und für die 2013 keine Zulassung eines neuen Anwendungsgebiets entstanden ist. Alle Produkte haben gemeinsam, dass sie deutlich weniger Patienten erreichen, als die vom G-BA definierten Zielpopulationen mit Zusatznutzen erwarten lassen.

Eine enge Verzahnung der „Slicing“-Problematik zeigt sich auch bei der Festlegung der zweckmäßigen Vergleichstherapie, die eine zentrale Stellschraube im gesamten AMNOG-Verfahren darstellt. Hier gilt es zu unterscheiden, ob die vom G-BA vorgegebenen Komparatoren im Alternativ- oder Ergänzungsverhältnis zueinander stehen. Bei einem Alternativverhältnis stellen alle Optionen nach Auffassung des G-BA eine gleichermaßen zweckmäßige Therapiealternative dar. Diese Konstellation ist insbesondere dann problematisch, wenn ein Zusatznutzen als nicht belegt angesehen wird und in der Konsequenz die wirtschaftlichste Alternative und damit häufig das billigste Generikum als preisliche Obergrenze fun-

Auswahl der zweckmäßigen Vergleichstherapie

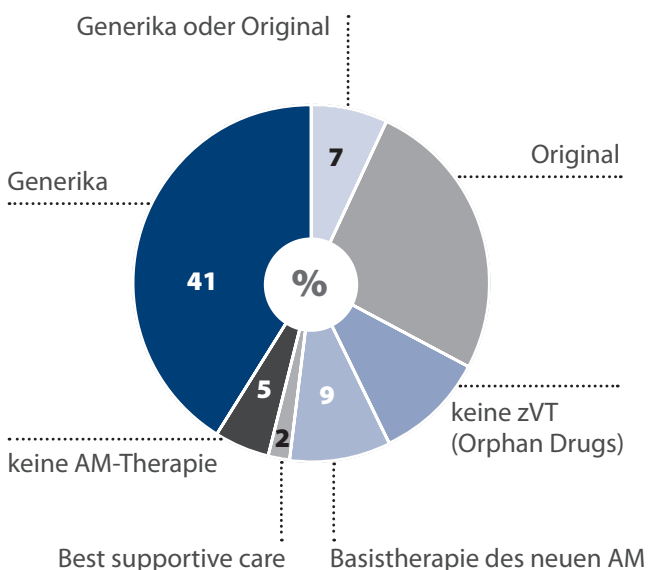


Abb. 5: Bei einem Großteil der Subpopulationen werden kostengünstige Vergleichstherapien wie etwa Generika benannt, mit Folgen für die spätere Preisverhandlung.

giert. Ein Ergänzungsverhältnis führt dagegen ggf. zu einer Aufteilung des Anwendungsgebietes in Subpopulationen mit der jeweils gültigen zweckmäßigen Vergleichstherapie. In solchen Fällen dürften die Subpopulationen auch im Rahmen der anschließenden Erstattungsbetragsverhandlungen separat betrachtet werden.

Keine spürbare Änderung durch AMG-Novelle

Das Fehlsteuerungspotenzial sollte im Rahmen des Dritten AMG-Änderungsgesetzes im Sommer 2013 vermieden werden. Es wurde dabei festgelegt, dass die Auswahl der Komparatoren allein nach medizinischen Kriterien erfolgen muss und nicht auch nach wirtschaftlichen Erwägungen.

Zudem wurde dem G-BA die Möglichkeit eingeräumt, sofern medizinisch sinnvoll, mehrere Komparatoren vorzugeben, gegenüber denen der Hersteller den Zusatznutzen seines Arzneimittels nachweisen kann.

Leider hat diese Adjustierung in der Praxis des G-BA nicht zu den spürbaren Veränderungen geführt, die der Gesetzgeber damit erreichen wollte. Weiterhin werden die meisten negativen Zusatznutzenbeschlüsse aus quasi-formalen Gründen gefasst und weiterhin sind die Nutzenbewertung und die Preisverhandlung nicht entkoppelt. Bei einem großen Teil der Subpopulationen werden kostengünstige Vergleichstherapien (z. B. Generika oder „Best supportive care“) benannt, die sich vor allem mit Blick auf die anschließenden Verhandlungen als problematisch erweisen (Abbildung 5).

Das Preisniveau der Generika lässt nämlich die Notwendigkeit einer Refinanzierung der Aufwendungen für die Forschung und Entwicklung systematisch außer Acht. Insbesondere werden dabei solche Investitionen in Indikationen benachteiligt, in denen länger kein Fortschritt stattgefunden hat. Ausgerechnet die therapeutischen Solisten erhalten nach der bestehenden Logik der Verhandlungen die billigste Preisbasis. Wenn die zweckmäßige Vergleichstherapie nicht nur als Nutzenreferenz, sondern gleichzeitig als Preisreferenz genutzt wird, wie es der GKV-SV und offensichtlich auch die Schiedsstelle praktizieren, mutiert die Nutzenbewertung zu einer Schwundform einer Kostennutzenbewertung, die zudem völlig falsche Anreize setzt, da gerade Entwicklungen in Bereichen, in denen längere Zeit kein Fortschritt stattgefunden haben, durch eine generische Preisreferenz „bestraft“ werden.

Erstattungsbeträge und Festbetragssystem

Bei den bis Dezember 2014 hierdurch bestimmten Erstattungsbeträgen finden sich 82 Prozent der Erstattungsbeträge, die seit dem 14. SGB V-Änderungsgesetz gleichzeitig zu

„Listenpreisen“ aufgewertet wurden, auf einem Erstattungsniveau wieder, das unterhalb der europäischen Listenpreise liegt. Rund ein Drittel der Arzneimittel markieren sogar das untere europäische Preisniveau. Als Folge dieser niedrigen Listenpreise, die über Parallelexporte und internationale Preisreferenzierung sofort internationale Wirkung auslösen, wurden bis dato 13 Produkte vom deutschen Markt zurück gezogen.

Seitdem dem Erstattungsbetrag durch das 14. SGBV-Änderungsgesetz Listenpreischarakter zugewiesen worden ist, hat der GKV-SV angekündigt, das Erstattungsbetragsystem und das Festbetragsystem vermengen zu wollen (Bode und Tebinka-Olbrich 2014). Dann würden für mehrere Arzneimittel zunächst Nutzenbewertungen und Erstattungsbetragsverhandlungen durchgeführt und anschließend diese Arzneimittel gemeinsam in eine Festbetragsgruppe kommen und denselben Festbetrag erhalten. Dies ignoriert, dass beide Systeme von sich widersprechenden Grundannahmen ausgehen (Festbetragsystem von der Annahme austauschbarer Arzneimittel, AMNOG von Arzneimitteln mit Zusatznutzen und deshalb fehlender Austauschbarkeit) und dass dieser Gegensatz sich in völlig unterschiedlicher Systematik und Regelsetzung widerspiegelt.

Eine Vermischung würde dazu führen, dass verschiedene Arzneimittel mit unterschiedlichem Zusatznutzen denselben Festbetrag erhielten. Die zentrale AMNOG-Idee der nutzenbasierten Erstattung würde damit ad absurdum geführt. Um genau dies zu vermeiden, hat der Gesetzgeber das AMNOG speziell für „nicht festbetragsfähige Arzneimittel“ eingeführt. Und genau deshalb findet nach AMNOG gleich zu Verfahrensbeginn eine dichotome Unterteilung statt in solche Arzneimittel ohne Zuweisung eines Zusatznutzens, die in die Festbetragschiene kommen, und solche mit Zusatznutzen, für die ein nutzenbasierter Erstattungsbetrag verhandelt wird, der den Zusatznutzen widerspiegeln soll.

Schlussfolgerung und Ausblick

Es zeigt sich, dass die frühe Nutzenbewertung in der Praxis keineswegs reibungslos läuft. Dies gilt einerseits für zahlreiche Kritikpunkte, die wegen der (zu) engen Verzahnung von Nutzenbewertung und Erstattungsbetragsverhandlung nicht ausschließlich methodische Qualität haben. Beispielfähig zu nennen sind hier die häufige Aufteilung der Anwendungsgebiete in Subpopulationen und die Verwendung der medizinischen Vergleichstherapie als Kostenreferenz sowie die fehlende tatsächliche Beteiligung von Zulassungsbehörden und Fachgesellschaften. Das grundsätzliche Problem des AMNOG ist seine unzureichende Governance, die zu einer dysfunktionalen Vermischung von Nutzenbewertung und Preisverhandlung führt.

Irritierend ist ferner, dass sich inzwischen zeigt, dass auch Arzneimittel mit vom G-BA bestätigtem Zusatznutzen nur unzureichend die Patienten erreichen. Dies ist dem Faktum geschuldet, dass die nachgelagerten Wirtschaftlichkeitsprüfungen das AMNOG konterkarieren und die Wirtschaftlichkeit erstattungsbetrags geregelter Arzneimittel regelmäßig in Zweifel ziehen und die Ärzte sich von Regress bedroht sehen. Dabei ist völlig unverständlich, warum ein erstattungsbetrags geregeltes Arzneimittel unwirtschaftlich sein sollte: Wenn kein Zusatznutzen zugewiesen wurde, verursacht es maximal die Kosten der billigsten Vergleichstherapie; bei Zusatznutzen wurde ein entsprechender Erstattungsbetrag mit der GKV vereinbart, der auch mögliche Subgruppen ohne Zusatznutzen einpreist. Hier wird eine fehlende Abstimmung von zentralem AMNOG-Prozess und regionaler Arzneimittelsteuerung deutlich, die nachweisbar zu medizinischer Unterversorgung von Patienten führt. Es ist offensichtlich die gesetzliche Klarstellung erforderlich, dass erstattungsbetrags geregelte Arzneimittel bei zulassungsgemäßer Verordnung selbstverständlich wirtschaftlich sind.

Literatur

Spiegel-online (2012): Zwischenbilanz: Neu-Medikamente nützlicher als erwartet. Internet: <http://www.spiegel.de/gesundheit/diagnose/amnog-jedes-sechste-neue-medikament-bringt-mehr-nutzen-fuer-patienten-a-853520.html>.

Spiegel-online (2014): G-BA-Bilanz: Nur jedes fünfte neue Medikament bringt mehr Nutzen. Internet: <http://www.spiegel.de/gesundheit/diagnose/amnog-bilanz-g-ba-bescheinigt-nur-wenigen-medikamenten-zusatznutzen-a-970976.html>.

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2012): Cabazitaxel – Nutzenbewertung gemäß § 35a SGB V. Auftrag: A11-24.

Gemeinsamer Bundesausschuss (2013): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Apixaban (neues Anwendungsgebiet).

Rasch A, Dintsios CM (2015): Subgruppen in der frühen Nutzenbewertung von Arzneimitteln: eine methodische Bestandsaufnahme. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen. 109(1): 69-78.

Rothwell PM (2005): Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications and interpretation. Lancet. 365 (9454): 176-86.

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2013): Allgemeine Methoden. Version 4.1 vom 28.11.2013.

Häussler B, Höfer A, Hempel, E (2014): Arzneimittel-Atlas 2014. Der Arzneimittelverbrauch in der GKV. Dordrecht: Springer.

Bode CH A, Tebinka-Olbrich A (2014): Ein ideales Paar: Erstattungs- und Festbeträge zur Regulierung von Arzneimittelpreisen. Gesundheits- und Sozialpolitik (3): Band 7-14.