
**Von der Beratung über die Nutzenbewertung bis zur
Erstattungsbetragsfindung von neuen Arzneimitteln:
Eine Analyse des AMNOG auf Basis einer
Methodentriangulation unter Verwendung
biometrischer, akteurszentrierter und
gesundheitsökonomischer Ansätze**

Dr. rer. biol. hum. Charalabos-Markos Dintsios

**Von der Beratung über die Nutzenbewertung bis zur
Erstattungsbetragsfindung von neuen Arzneimitteln:
Eine Analyse des AMNOG auf Basis einer
Methodentriangulation unter Verwendung
biometrischer, akteurszentrierter und
gesundheitsökonomischer Ansätze**

Dr. rer. biol. hum. Charalabos-Markos Dintsios

Inauguraldissertation
zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaft (Dr. rer. oec.)
an der Fakultät für Wirtschaftswissenschaft
— Schumpeter School of Business and Economics —
der Bergischen Universität Wuppertal

vorgelegt von
Dr. Charalabos-Markos Dintsios
aus Nürnberg

Wuppertal, den 3. September 2021

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20220630-094932-9

[<http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20220630-094932-9>]

DOI: 10.25926/qr7b-mp03

[<https://doi.org/10.25926/qr7b-mp03>]

Erstgutachterin: Prof. Dr. Juliane Köberlein-Neu

Zweitgutachter: Prof. Dr. Volker Ulrich

Tag der mündlichen Prüfung: 17. März 2022

für meine Kinder

Inhaltsverzeichnis

1	Einführung	1
1.1	Motivation	2
1.2	Zielsetzung und Beitrag	11
1.3	Überblick.....	19
1.4	Liste der Publikationen der kumulativen Dissertation.....	53
2	Literatur	55

Abbildungsverzeichnis

Abbildung 1.1: Verfahrensablauf der frühen Nutzenbewertung von Arzneimitteln.....	4
Abbildung 1.2: Konzeptioneller Rahmen der Dissertation	12

Einführung¹

¹ Zur Vermeidung von nicht etablierten Übersetzungen unterschiedlicher Akteure im deutschen Gesundheitswesen und komplizierter sozialrechtlicher Termini ins Englische und der damit einhergehenden begrifflichen Unschärfen wird die Einführung auf Deutsch verfasst. Gleichwohl werden die einzelnen Veröffentlichungen im Rahmen der kumulativen Dissertation in der jeweiligen Publikationssprache (Deutsch oder Englisch) präsentiert. Ferner wird aus Gründen der Sprachökonomie und wegen des relativ großen Umfangs der Arbeit die männliche Form sowohl im Singular als auch im Plural ohne jegliche Intention einer geschlechtlichen Diskriminierung verwendet.

1.1 Motivation

Mit dem am 1.1.2011 in Kraft getretenen „Gesetz zur Neuordnung des Arzneimittelmarktes in der gesetzlichen Krankenversicherung“ (AMNOG) ist es zu einem Paradigmenwechsel im Umgang mit Arzneimittelinnovationen in Deutschland gekommen. Neue Arzneimittel müssen sich ab dem 1.1.2011 nach § 35a des fünften Sozialgesetzbuches (SGB V) einer auf klinischen Studien fußenden Nutzenbewertung im Vergleich zu einem vom Gemeinsamen Bundesausschuss (G-BA)², dem Selbstverwaltungsorgan des deutschen Gesundheitswesens – aufgrund seiner ausgestatteten regulativen Machtfülle als untergesetzlicher Normgeber agierend, festgelegten Komparator (zweckmäßige Vergleichstherapie) zur Bestimmung ihres Zusatznutzens unterziehen, der bei den anschließenden Erstattungsbetragsverhandlungen zwischen Arzneimittelhersteller und dem GKV Spitzenverband (GKV-SV) nach § 130b SGB V u.a. als Basis für den Erstattungsbetrag 12 Monate nach Markteintritt fungiert.

Im Gesetzentwurf zum AMNOG (Deutscher Bundestag 2010) werden dezidiert drei Ziele genannt: 1. Den Menschen müssen im Krankheitsfall die besten und wirksamsten Arzneimittel zur Verfügung stehen; 2. Die Preise und Verordnungen von Arzneimitteln müssen wirtschaftlich und kosteneffizient sein; 3. Es müssen verlässliche Rahmenbedingungen für Innovationen, die Versorgung der Versicherten und die Sicherung von Arbeitsplätzen geschaffen werden. Einfacher und prägnanter formuliert lassen sich die Ziele in der Trias Versorgungsqualität, Kostendämpfung und Innovationsförderung zusammenfassen. Die Umsetzung des AMNOG wird auf Ebene der Exekutiven (Bundesministerium für Gesundheit) durch die Arzneimittel-Nutzenbewertungsverordnung - AM-NutzenV und beim untergesetzlichen Normengeber G-BA durch seine eigene Verfahrensordnung (G-BA VerfO) begleitet.³

² Bestehend aus den zwei Bänken der Leistungserbringer (ambulant niedergelassene Ärzte mit kassenärztlicher Zulassung vertreten durch die Kassenärztliche Bundesvereinigung KBV und der Deutschen Krankenhausgesellschaft DKG) und der Kostenträger (vertreten durch den GKV-SV) sowie drei unparteiischen Mitgliedern mit Stimmrecht und den Patientenvertretern mit rein beratender Funktion ohne Stimmrecht.

³ Das AMNOG hat seit seiner Einführung vor 10 Jahren mehrere Ergänzungen und Änderungen erfahren, die sich oft auf technische Details beschränkten. Hervorzuheben darunter sind: (i) die durch das am 25.10.2012 in Kraft getretene Zweite Gesetz zur Änderung arzneimittelrechtlicher und anderer Vorschriften (2. AMGÄndG) eingeführte Beteiligung der Zulassungsbehörden bei Beratungen des G-BA zur Planung klinischer Studien sowie Berücksichtigung europäischer Preise in den Erstattungsbetragsverhandlungen, (ii) die durch das am 13.8.2013 in Kraft getretene Dritte Gesetz zur Änderung arzneimittelrechtlicher und anderer Vorschriften (3. AMGÄndG) eingeführte Streichung des Wirtschaftlichkeitskriteriums bei der Bestimmung der zweckmäßigen Vergleichstherapie, (iii) die durch das am 1.4.2014 in Kraft getretene Vierzehnte Gesetz zur Änderung des Fünften Buches Sozialgesetzbuch (14. SGB V ÄndG) eingeführte Aufhebung des Bestandsmarktaufrufs, (iv) die durch das am 13.5.2017 in Kraft getretene Gesetz zur Stärkung der Arzneimittelversorgung (GKV-AMVSG) eingeführte Geltung des Erstattungsbetrags auch für Krankenhäuser und Sollregelung zum Erstattungsbetrag hinsichtlich seiner Nichtüberschreitung der Kosten der zweckmäßigen Vergleichstherapie bei fehlendem Zusatznutzen und (v) die durch das am 16.8.2019 in Kraft getretene Gesetz für mehr Sicherheit in der Arzneimittelversorgung (GSAV) eingeführte Beteiligung von Fachgesellschaften sowie der Arzneimittelkommission der deutschen Ärzteschaft in den Beratungen mit den Herstellern und die Anwendungsbegleitende Datenerhebung.

Dem AMNOG vorausgegangen waren mehrere gesetzliche Reformen mit Kostendämpfungscharakter bzw. -zielen – auch im Arzneimittelbereich, ohne allerdings eine Erstattung von Arzneimittelinnovationen auf Basis einer (Zusatz-)Nutzenbewertung bis dahin vorzusehen. Vor der Verabschiedung des AMNOG lagen Defizitschätzungen des Schätzerkreises beim damaligen BVA⁴ in Höhe von 7,4 Mrd. € für 2010 und 11 Mrd. € für 2011 vor (GKV-Schätzerkreis), die den Gesetzgeber u.a. dazu veranlassten, im Rahmen des GKV-ÄndG⁵ 2010 ein Preismoratorium für verschreibungspflichtige Arzneimittel einzuführen, das in der Realität die Arzneimittelpreise auf dem Stand von 2009 bzw. dem ersten Preis bei späterer Markteinführung einfrieren ließ sowie Herstellerabschläge in unterschiedlicher Höhe vorsah⁶.

Der vorherrschende Ansatz der in den vergangenen drei Dekaden eingeleiteten Gesundheitsreformen lag in der Dämpfung von Kostensteigerungen mit relativ einfachen Mitteln der Kostenkontrolle wie Moratorien und Budgetierungen. Als Resultat dessen lässt sich hervorheben, dass die von der GKV getragenen Gesundheitsausgaben etwa 6% des BIP in den letzten zwanzig Jahren ausmachten.⁷ Im Arzneimittelbereich beinhalteten weitere Kostendämpfungsmaßnahmen Festbeträge und die aut idem Regelung, welche es den Apothekern ermöglichte, Arzneimittel durch billigere Präparate mit demselben Wirkstoff zu substituieren.⁸ Des Weiteren waren und sind Rabattverträge mit einzelnen Krankenversicherungen nach § 130a SGB V möglich. Mit dem AMNOG wurde zusätzlich § 130c SGB V eingeführt, welcher aktuell abweichend von § 130b SGB V Vereinbarungen über die Erstattung von sowie zur Versorgung ihrer Versicherten mit Arzneimitteln mit den Herstellern vorsieht, insbesondere mengenbezogene Staffelung des Preisnachlasses, jährliches Umsatzvolumen mit Ausgleich von Mehrerlösen oder eine Erstattung in Abhängigkeit von messbaren Therapieerfolgen.⁹ Vor dem AMNOG wurden nur in einzelnen Fällen patentgeschützte Arzneimittel auf Grundlage einer nach § 139b (1) SGB V durchgeführten Nutzenbewertung des Instituts für Qualität und Wirtschaftlichkeit (IQWiG) durch den G-BA wegen Unwirtschaftlichkeit gemäß § 12 SGB V von der Erstattung ausgeschlossen.¹⁰ Gesundheitsökonomische Evaluationen als Basis für Höchstbetragssetzungen¹¹ von Arzneimitteln waren nach dem GKV-WSG¹² 2007 vorgesehen, kamen aber aufgrund der lang andauernden Methodenentwicklung des IQWiG

⁴ Bundesversicherungsamt, ab 1.1.2020 in Bundesamt für Soziale Sicherung (BAS) umbenannt.

⁵ zu den Inhalten des GKV-Änderungsgesetzes siehe BGBl 2010 I Nr. 39, S. 983 ff.

⁶ oszillierend ab 2003 zwischen 6% und 16%, aktuell 7%. Für Generika seit 2006 unverändert 10%.

⁷ Berechnet auf Basis der GKV-Ausgaben in Relation zu den gesamten Gesundheitsausgaben in toto als BIP Anteil aus RKI Heft 45 „Ausgaben und Finanzierung des Gesundheitswesens“, 2009 für den Zeitraum 2000-2009 und aus den jährlichen Daten des Statistischen Bundesamtes ab 2009.

⁸ vgl. § 35 & § 129 SGB V.

⁹ Auch unter dem englischen Terminus ‚Pay for Performance‘ (P4P) in der Literatur bekannt.

¹⁰ bspw. Beschluss des G-BA über eine Änderung der Arzneimittel-Richtlinie in Anlage 10: Kurzwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2 aus dem Jahre 2008.

¹¹ Höchstbeträge nach § 31 (2a) SGB V alt (abgelöst durch das AMNOG am 01.01.2011).

¹² GKV-Wettbewerbsstärkungsgesetz.

zum Effizienzgrenzansatz (Caro et al. 2010a) und den diesen Prozess begleitenden Methodenstreit primär mit den deutschen Gesundheitsökonomern (Greiner et al. 2010; Schwarzbach et al. 2009) bzw. deren damaligen fachwissenschaftlichen Vertretung¹³ aber auch international (Brouwer & Rutten 2010; Caro et al. 2010b; Dintsios & Gerber 2010; Gerber & Dintsios 2010; Sculpher & Claxton 2010) nicht für die Bestimmung von Erstattungspreisen zur Anwendung.¹⁴

Die folgende Abbildung dient zur Veranschaulichung des komplexen Verfahrensablaufs der frühen Nutzenbewertung mit allen involvierten Akteuren und Verfahrensschritten.

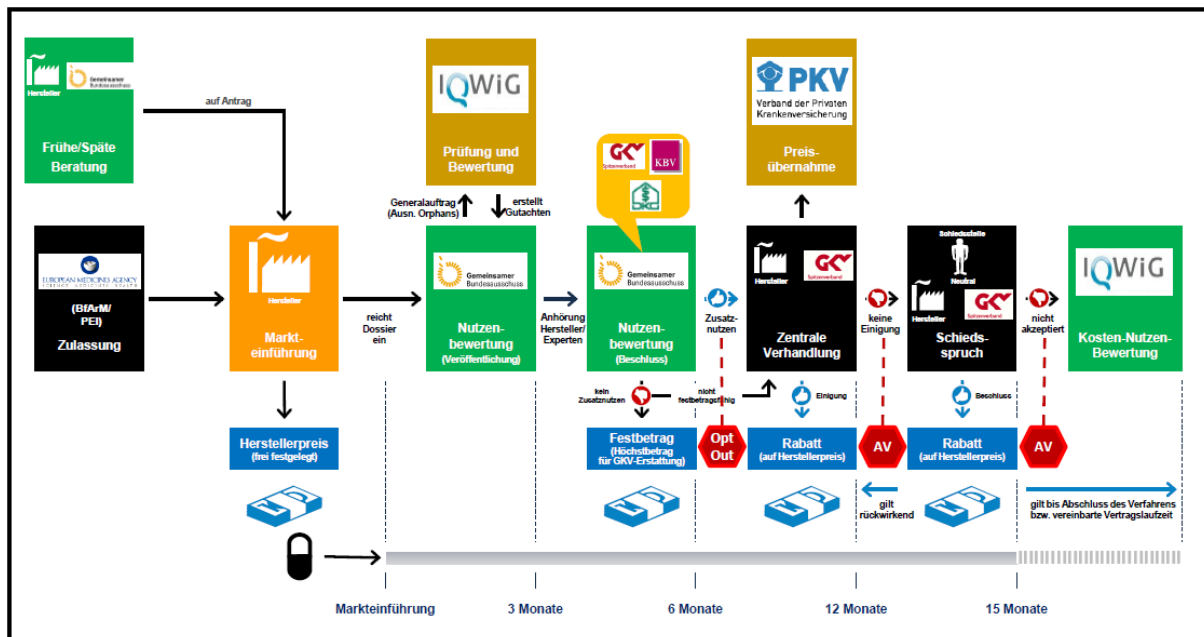


Abbildung 1.1: Verfahrensablauf der frühen Nutzenbewertung von Arzneimitteln nach § 35a SGB V (in Anlehnung an die Darstellung des Bundesgesundheitsministeriums)

Mit dem Tag des Inverkehrbringens eines von den zuständigen Zulassungsbehörden auf europäischer¹⁵ bzw. nationaler¹⁶ Ebene zugelassenen Arzneimittels mit neuem Wirkstoff, definiert durch Eintrag in der Großen Deutschen Spezialitätentaxe (sog. Lauer-Taxe), beginnt offiziell das Verfahren der frühen Nutzenbewertung. Vorher ist eine (frühe) Beratung¹⁷ mit dem G-BA für die Hersteller hinsichtlich spezieller Fragestellungen zum Studiendesign und insbesondere zur zweckmäßigen Vergleichstherapie möglich.

¹³ Sektion Gesundheitsökonomie des Vereins für Socialpolitik

¹⁴ Einige Ausnahmen zur modelhaften Simulation der Effizienzgrenze ohne Konsequenzen allerdings für die Erstattung stellt der IQWiG Pilotbericht G09-01 zur Kosten-Nutzenbewertung von Antidepressiva (IQWiG 2013) und zu Behandlungsoptionen von Hepatitis C (Mühlbacher & Sadler 2017). Der Versuch den Effizienzgrenzansatz mit externen Schwellenwerten zu kombinieren (Sandmann et al. 2018) bzw. ähnliche Effekte auf die Langzeitkostenentwicklung zwischen Effizienzgrenzansatz in Deutschland und externen Schwellenwertansatz in England zu postulieren (Gandjour 2020) scheint hinsichtlich seiner weiteren Anwendung nicht zu fruchten.

¹⁵ durch die EMA (European Medicines Agency)

¹⁶ durch das BfArM (Bundesinstitut für Arzneimittel und Medizinprodukte) oder für spezifische Produkte wie bspw. monoklonale Antikörper durch das PEI (Paul-Ehrlich-Institut)

¹⁷ frühe Beratung meint in der Regel noch vor Durchführung von Zulassungsstudien (Phase II/III-Studien).

Grundlage für die Nutzenbewertung ist ein Dossier, das vom Hersteller¹⁸ beim G-BA eingereicht wird. Das Dossier muss laut AM-NutzenV Informationen über das zugelassene Anwendungsgebiet, den medizinischen Nutzen, den medizinischen Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie, die Anzahl der Patienten, die Kosten der Therapie und die Anforderung an eine qualitätsgesicherte Anwendung enthalten. Das Dossier wird in einem Zeitabschnitt von drei Monaten nach Einreichung durch den G-BA selber (Orphan Drugs) oder im Auftrag durch das IQWiG bewertet¹⁹. Im Anschluss an die Veröffentlichung der Bewertung folgt ein Stellungnahmeverfahren inklusive einer Anhörung. Spätestens drei Monate nach Abschluss der Bewertung beschließt der G-BA über den Zusatznutzen.²⁰ Insofern ein Zusatznutzen als belegt anerkannt wird, kann dieser folgende Ausmaße annehmen („erheblich“ – „beträchtlich“ – „gering“ oder für den Fall, dass das Ausmaß nicht bestimmt werden kann: „nicht quantifizierbar“). Weiter kann ein Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie nicht belegt oder der Nutzen des zu bewertenden Arzneimittels sogar geringer als derjenige der zweckmäßigen Vergleichstherapie ausfallen. Die den Zusatznutzen umgebende Ergebnissicherheit (ausgedrückt als Wahrscheinlichkeit²¹) wird trichotomisiert dargestellt: „Beleg“ – „Hinweis“ – „Anhaltspunkt“. Insgesamt beträgt somit der maximale Zeitraum von der Einreichung des Dossiers bis zur Beschlussfassung sechs Monate. In § 35a (5) SGBV ist ferner geregelt, dass der Hersteller frühestens ein Jahr nach Veröffentlichung des Beschlusses eine erneute Nutzenbewertung beantragen kann.

Stellt der G-BA einen Zusatznutzen fest, folgt nach der entsprechenden Beschlussfassung die Erstattungsbeitragsverhandlung zwischen Hersteller und dem GKV-Spitzenverband innerhalb von sechs Monaten nach § 130b (1) SGB V. Die Kriterien für diese Verhandlung sind in einer zwischen dem GKV-SV und den Verbänden der pharmazeutischen Unternehmer getroffenen Rahmenvereinbarung (GKV-SV 2012) relativ detailliert festgelegt. So werden die Jahrestherapiekosten in der jeweiligen Indikation vergleichbarer Arzneimittel und die tatsächlichen Abgabepreise in anderen, nach einem in der Rahmenvereinbarung festgelegten Kriterienkatalog zur entsprechenden Bevölkerungsgröße und zur wirtschaftlichen Leistung ausgewählten, europäischen Ländern, gewichtet nach den jeweiligen Arzneimittelmarktumsätzen und Kaufkraftparitäten neben den Jahrestherapiekosten der zweckmäßigen Vergleichstherapie und dem Zusatznutzen berücksichtigt.

¹⁸ in den entsprechenden Rechtstexten wird vom pharmazeutischen Unternehmen (pU) gesprochen. Dies hat arznei- und patentrechtliche Implikationen, da Hersteller und Patenhalter sich unterscheiden können. Um Begriffsverwirrungen aufgrund synonyme Verwendung der Termini Hersteller und pU (bspw. beim Begriff Herstellerrabatte oder Herstellerabgabepreis) zu vermeiden, wird in der Arbeit primär der geläufigere Begriff Hersteller verwendet.

¹⁹ im internationalen Sprachgebrauch hat sich hierfür der Begriff „Assessment“ etabliert.

²⁰ die Beschlussfassung von Entscheidungsträgern wird international als „Appraisal“ bezeichnet.

²¹ vgl. § 5 (5) AM-NutzenV, allerdings nicht im statistischen Sinne unter prozentualer Angabe.

Vergibt der G-BA für ein Arzneimittel keinen Zusatznutzen, so wird es nach § 35a (4) SGB V einer Festbetragsgruppe zugeordnet. Wenn keine entsprechende Festbetragsgruppe hierfür existiert, verhandeln die Hersteller mit dem GKV-SV einen Erstattungsbetrag, der allerdings nicht zu höheren Jahrestherapiekosten der zweckmäßigen Vergleichstherapie bzw. der wirtschaftlichsten darunter, für den Fall, dass mehrere bestimmt wurden (§ 130b (3) SGB V), führen darf. Alternativ zum Eintritt in die Verhandlungen hat der Hersteller unabhängig vom Vorliegen eines Zusatznutzens die Möglichkeit sein Produkt innerhalb der ersten vier Wochen nach Bekanntgabe des Beschlusses vom deutschen Markt zu nehmen²² und auf eine Verhandlung zu verzichten. Es wird dann kein Erstattungsbetrag gelistet. Nach diesen vier Wochen ist zwar jederzeit die Möglichkeit einer „Außer Vertriebs-Stellung“²³ gegeben, diese führt aber nicht dazu, dass das Verhandlungsverfahren ausgesetzt wird und birgt für den Hersteller das Risiko, dass der GKV-SV bei einem Fernbleiben des Herstellers mit Re- bzw. Parallelimporteuren verhandelt oder auch vor die Schiedsstelle nach § 130b (4) SGB V in Abhängigkeit vom Verhandlungsergebnis tritt und dort einen Erstattungsbetrag „erwirkt“, der anschließend auch gelistet wird. Wenn der Hersteller und der GKV-SV sich nicht innerhalb von sechs Monaten einigen können, wird der strittige Vertragsinhalt – darunter nicht nur der Erstattungsbetrag – innerhalb von drei Monaten von der oben genannten Schiedsstelle festgelegt. Dieser festgelegte Erstattungsbetrag gilt dann ab dem 13.ten Monat rückwirkend.

Im Anschluss an einen Schiedsstellenentscheid kann gemäß § 35b SGB V eine Kosten-Nutzen-Bewertung²⁴ beim G-BA durch jede Vertragspartei beantragt werden, die vom IQWiG durchgeführt wird. Außerdem kann gegen einen Schiedsstellenentscheid eine gerichtliche Klage erhoben werden, allerdings haben weder die Kosten-Nutzen-Bewertung noch eine Klage eine aufschiebende Wirkung, sodass der festgesetzte Erstattungsbetrag weiterhin gilt.

Das gesamte Verfahren kann synoptisch betrachtet als ein über die vorangestellte Nutzenbewertung des G-BA unter Aufbereitung der vorhandenen Studienevidenz durch sein Auftragsinstitut IQWiG dem ‚Value-based pricing‘ Ansatz nahekommendes Vorgehen zur Erzielung eines zeitversetzten Erstattungsbetrags nach vorausgegangener Markteinführung neuer Arzneimittel mit freier Preisbildung auf Grundlage einer Verhandlungslösung mit einem Monopsonist (GKV-SV)²⁵, welcher konstituierender Akteur des G-BA ist, bezeichnet werden. Daraus lassen sich für die Untersuchung des primären Erkenntnisobjekts Kategorien

²² siehe Option „Opt-Out“ in Abb. 1.1.

²³ siehe Option „AV“ in Abb. 1.1.

²⁴ Beim § 35b SGB V handelt es sich um einen „toten“ Paragraphen, da diese Option seit Einführung des AMNOG von keiner Vertragspartei gezogen wurde. Eine Kosten-Nutzen-Bewertung kann auch nach §35a (5a) SGB V auf Verlangen des Herstellers bei nicht attestiertem Zusatznutzen vom G-BA nach § 35b oder §139a (3) Nr. 5 an das IQWiG in Auftrag gegeben werden, wenn der Hersteller hierfür die Kosten trägt. Auch dieser Paragraph kam bisher nicht zur Anwendung.

²⁵ nach § 130b (1) SGB V verhandelt der GKV-SV im Benehmen mit dem PKV-Verband die Erstattungsbeträge.

wie Evidenz-basierte Medizin bzw. Gesundheitstechnologiebewertung²⁶, Value-based Pricing, Selbstverwaltung, Marktmacht und weitere als analytisches Raster zur gesetzlich intendierten Zielerreichung ableiten, deren Erfassung allerdings aufgrund ihrer unterschiedlichen erkenntnistheoretischen Provenienz eine Methodentriangulation mittels qualitativer sowie quantitativer Methoden unter Verwendung sowohl biometrischer und akteurszentrierter als auch (gesundheits-)ökonomischer Ansätze erfordert.

Mit Einführung des AMNOG stellten Gerber et al. (2011) die Frage, wieweit das neue Vorgehen vom Ansatz des Value-based Pricing entfernt ist, ohne eine klare Antwort darauf zu geben. In der Preispolitik wird der 3-K-Ansatz für Kosten-basierte, Konkurrenz-orientierte und Kunden-basierte Preisbildung genannt, wobei sich ersterer als Gewinnzuschlag auf die Kosten, zweiter als Positionierung in Relation zu den Wettbewerbern und letzterer als Wert spezifischer Nutzenattribute aus Kundensicht (anders ‚Value-based Pricing‘) definieren lässt (Hinterhuber 2004). Es besteht ein breiter Konsens, dass ein sich an den Wert aus Kundensicht und deren Zahlungsbereitschaft orientierender Preis die Preismacht eines Unternehmens im Sinne höherer Preisfindung im Vergleich zu den beiden anderen K-Ansätzen positiv beeinflussen kann (Liozu 2017). Allerdings sei darauf hingewiesen, dass im Kontext der Preisbildung im Gesundheitswesen die GKV als zahlender Nachfrager für Gesundheitsprodukte konsumierende und Dienstleistungen in Anspruch nehmende Patienten auftritt, wobei sie die Interessen ihrer Versicherten und nicht explizit der Patienten darunter vertritt. Im Sinne des AMNOG handelt es sich mit Bezug auf Value-based Pricing um eine ex-post Preisfestsetzung²⁷ durch Bestimmung der Wertschöpfung eines Angebots für den GKV-SV oder durch empirische Bestimmung der angebotsbezogenen Zahlungsbereitschaft des GKV-SV und Ermittlung des resultierenden Gewinnbeitrags für den Hersteller.

Value-based Pricing für Arzneimittel ist kein deutsches Novum²⁸ und wird in einigen europäischen Gesundheitssystemen bereits verfolgt (Jommi et al. 2020; Prieto-Pinto et al. 2020) und in Japan, dem zweitgrößten Arzneimittelmarkt, aktuell eingeführt (Shiroiwa 2020), wobei sich sowohl Kriterien als auch Messung des Wertes von Gesundheitsleistungen je nach Gesundheitssystem oder Akteursperspektive relativ stark unterscheiden können (Towse & Barnsley 2013; Vermeulen & Krabbe 2018). Sein theoretischer Ursprung für Gesundheitsleistungen findet sich beim bekannten Harvard Wettbewerbsstrategen Porter (Porter &

²⁶ im internationalen Sprachgebrauch unter ‚Health Technology Assessment‘ (HTA) bekannt.

²⁷ Ex ante wird der Hersteller den Kosten- oder Konkurrenzbasierten Ansatz verfolgen, wobei er durch den zeitlich limitierten Patentschutz bei echten Arzneimittelinnovationen (keine „Me-too“ Präparate) im Rahmen der freien Preisfestsetzung eine passagere Monopolstellung erlangen kann. Letzteres gilt im AMNOG allerdings nur für die ersten 12 Monate nach Markteintritt.

²⁸ Alternative Vorschläge zur Bepreisung von Arzneimittelinnovationen aus dem deutschsprachigen Raum mit Ableitung der Zahlungsbereitschaft für eine Nutzeneinheit aus inkrementellen Kosten-Effektivitäts-Verhältnissen in der palliativen Behandlung bzw. Pflege (Gandjour & Chernyak 2011) oder auf Grundlage eines sozialen Gesundheitsindex (Kifmann 2010) haben keinen Widerhall im deutschen Gesundheitssystem gefunden.

Teisberg 2006; Porter 2010) und in der Arzneimittelversorgung kann man in Schweden ab 2002 den Vorläufer bei der Umsetzung, zu einer Zeit als der Begriff noch nicht kursierte, sehen (Moïse & Docteur 2007), während das englische NHS trotz etablierter Kosten-Effektivitäts-Analyse von Arzneimitteln erst später nachzog (Claxton et al. 2008; Raftery 2013). Folgt man der Taxonomie von Sussex et al. (2013), so ist das AMNOG Vorgehen im breiteren Sinn dem Ansatz zuzuordnen und nach Jommi et al. (2020) erfüllt es das Kriterium der Berücksichtigung des Wertes eines neuen Arzneimittels (Zusatznutzen) sowie die Bedingung zweiter Ordnung der Berücksichtigung des Einflusses auf das Ausgabengeschehen.²⁹ Allerdings ist aus der internationalen Literatur außer deskriptiven Informationen zu den vorgesehenen Regularien hinsichtlich Arzneimittelpreisverhandlungen wenig über deren reale Umsetzung bekannt (Jommi et al. 2020).

Aus der mikroökonomischen Standardliteratur (bspw. (Pindyck & Rubinfeld 2009)) ist hinlänglich bekannt, dass für Monopsonisten die Grenzausgaben höher als Durchschnittsausgaben liegen, womit bei Gleichsetzung des Grenzwertes mit den Grenzausgaben in Abhängigkeit von der Angebotselastizität sich geringere Mengen bei niedrigeren Preisen im Vergleich zum Wettbewerbsmarkt ergeben. Die Monopsonmacht führt zu Nettowohlfahrtsverlusten, da sich das Produktionsniveau unter dem Wettbewerbsniveau bewegt oder anders ausgedrückt, Konsumenten mit höherer Zahlungsbereitschaft als die Grenzkosten der Produktion nicht zum Zuge kommen (Bungenstock 2011; Pauly 1988). Folgt man dem ‚Alles oder Nichts‘ Modell (Herndon 2002), ergibt sich sogar die Möglichkeit für den Monopsonist auf dem Arzneimittelmarkt, indem er die Hersteller vor dem Dilemma einer garantierten Gleichgewichtsmenge des Wettbewerbsmarkts zu einem niedrigeren Preis oder keiner Abnahme von neuen Arzneimitteln setzt, zumindest kurzfristig die Gleichgewichtsmenge nicht zu reduzieren (Bungenstock 2011), weil die Hersteller um auf dem Markt zu bestehen, ihre Preise soweit reduzieren werden, bis ihre Produzentenrente vollkommen aufgebraucht ist. Dies kann sogar zu niedrigeren Preisen als im Monopson ohne ‚Alles oder Nichts‘ Angebot führen, wobei langfristig innovative Hersteller mit höheren Durchschnittskosten aus dem Markt ausscheiden werden. Aus diesem Verhalten des (absoluten) Monopsonisten werden reduzierte und verzögerte Einführungen von Arzneimittelinnovationen trotz passagerer Monopolposition³⁰ durch zeitlich begrenzten Patentschutz postuliert (Bungenstock 2011).

²⁹ Aus den Kriterien der RahmenV lassen sich neben der primär mit dem AMNOG intendierten wertbasierten Preisfindung auf Basis der Nutzenbewertung auch Elemente der internen Preisreferenzierung (IPR) durch die Berücksichtigung der Preise vergleichbarer Arzneimittel in der Indikation von Interesse sowie der externen Preisreferenzierung (EPR) (Rémuzat et al. 2015) über die europäischen Vergleichspreise identifizieren. Zu den positiven und negativen Effekten von ERP auf die Arzneimittelausgaben und -versorgung siehe Kanavos et al. (2020). Eine Gegenüberstellung der Kurz- und Langzeiteffekte von Value-based Pricing und External Price Referencing liefern Kanavos et al. (2010).

³⁰ Zwar kann die Schaffung gegengewichtiger Marktmacht durch ein bilaterales Monopol bzw. einen Monopson-Monopol Markt zu einem Pareto-Optimum wie unter vollständiger Konkurrenz führen (Fritsch 2014), allerdings ist ein solcher Arzneimittelmarkt aufgrund der Substitutionskonkurrenz (relative Monopole) nicht organisierbar.

Empirisch konnte Danzon (1997) für die Europäische Union nachweisen, dass sich die Arzneimittelpreise von optimalen Ramsey-Preisen, d.h. einer zweitbesten Lösung zur Regulierung natürlicher Monopole über umgekehrt proportional zur Preiselastizität kalkulierte Aufschläge auf die Grenzkostenpreise zur bestmöglichen Allokationseffizienz (Pindyck & Rubinfeld 2009), aufgrund von Parallelhandel und monopsonistischer Marktstrukturen weit entfernt liegen.

Hinsichtlich des AMNOG Zieles der Gewährleistung einer hohen Versorgungsqualität mit Arzneimitteln gab es in der Vergangenheit schon Zweifel, ob aufgrund eines in der Praxis nicht hinreichenden Informationsstands und wirtschaftlich bedingter Fehlanreize auf Seiten der Ärzte sich die Ergebnisse der Nutzenbewertung im ärztlichen Verschreibungsverhalten niederschlagen (Bauer et al. 2016). Zwischen Hersteller und GKV-SV im Rahmen der Erstattungsbetragsverhandlungen verhandelte Praxisbesonderheiten, welche die Verordnung von bewerteten Arzneimitteln aus der Wirtschaftlichkeitsprüfung ausschließen, scheinen auf Ebene regionaler Kassenärztlicher Vereinigungen (KV) nicht umgesetzt oder beachtet zu werden, womit sie nach Bauer et al. (2016) jede Relevanz verloren haben. Mit dem GKV-AMVSG (2017) ist es zu einer weiteren Regionalisierung der ärztlichen Wirtschaftlichkeitsprüfungen gekommen³¹ und das Arzteinformationssystem eingeführt worden, dass die Aufnahme der Nutzenbewertungsergebnisse in die Praxissoftware vorsieht.

Gandjour (2018) sagt ausgehend von einer deutlich größeren Unter- als Überversorgung mit innovativen Arzneimitteln in Deutschland vorher, dass die Aufnahme der Nutzenbewertungsergebnisse in die Praxissoftware zu einer Umsatzsteigerung der pharmazeutischen Industrie führen wird und argumentiert, dass der richtige Ansatz in Einklang mit der ökonomischen Theorie darin läge, entweder die Informationsverfügbarkeit für Patienten zu verbessern oder im Falle eines Staatsversagens die bestehenden Vorschriften zu ändern. Letzteres kann die Aufhebung von Sanktionen für ärztliche Verschreibungen und die Verhängung von Verschreibungsausschlüssen beinhalten. In ihrer empirischen Analyse gelangen Blankart und Stargardt (2020) zum Ergebnis, dass die Veröffentlichung der Bewertungsergebnisse sich auf die Verbreitung der Arzneimittel auswirkte, da vor Ergebnisveröffentlichung kein signifikanter Unterschied zwischen positiv und negativ bewerteten Arzneimitteln in der Geschwindigkeit ihres Eingangs in die Versorgung bestand, während nach Veröffentlichung positiv bewertete Arzneimittel signifikant schneller angenommen wurden und schlussfolgerten, dass HTA über die Verwendung in Preisverhandlungen hinaus Mehrwert stiftet.

Wenn die Verfügbarkeit positiv nutzenbewerteter Arzneimittel und deren steigender Marktanteil als Indikator für eine hohe Versorgungsqualität herhalten, so stellt sich die Frage

³¹ Auf regionaler Ebene handelt es sich um reine Mengen-regulierende Steuerungsinstrumente (bspw. Wirkstoffvereinbarungen in Bayern, Richtwertvolumina in Baden-Württemberg, Referenzfallwerte in Thüringen, Zielquoten in Mecklenburg-Vorpommern usw.), die auch patentgeschützte Arzneimittel betreffen.

nach der kontinuierlichen Weiterentwicklung von Innovationen, die mit dem dritten Ziel des AMNOG, der Sicherung des Innovationsstandorts Deutschland einhergehen, womit in Konsequenz beide Ziele (Arzneimittelversorgungsqualität und Innovationssicherung) direkt interagieren. Zu den potenziell negativen Auswirkungen von Monopsonen auf Forschung und Entwicklung und die dadurch bedingten Wohlfahrtsverluste durch Marktaustritt von Herstellern wurde weiter oben eingegangen. Eine weitere Facette, die HTA betrifft, wird von Cook und Golec (2017) aufgeworfen. Sie untersuchen in ihrem realen Options-Modell, wie sich die Nichtberücksichtigung von realen Optionen (bspw. die Lebensverlängerung von Patienten mit Arzneimitteln bis zur Entwicklung kurativer Therapien oder nachfolgender ebenfalls lebensverlängernder Arzneimittel) durch HTA auf die Forschungs- und Entwicklungsinvestitionen auswirkt und gelangen zum Ergebnis, dass die Risikobereitschaft der Investoren sinkt und damit die Innovationsbereitschaft betroffen ist.

Das mit dem AMNOG verfolgte Kostendämpfungsziel wurde anhand der in der Gesetzesbegründung genannten Höhe von 2 Mrd. Euro (Deutscher Bundestag 2010) erst 2018, sieben Jahre nach Gesetzesverabschiedung, mit Einsparungen von 2,65 Mrd. Euro erreicht (Schwabe & Ludwig 2020). Das Gesetz schätzte in seiner Begründung (Deutscher Bundestag 2010) die Bürokratiekosten für ein Dossier nach § 35a (1) sehr naiv auf 1.250 Euro und veranschlagte 2,5 Arbeitstage je Dossier. Es scheinen Transaktions- und Vertragsabschlusskosten (Voigt 2009) durch die Interaktion der Hersteller mit allen involvierten internen und externen Akteuren im Rahmen der Nutzenbewertung und die Erstattungsbetragsverhandlungen mit dem GKV-SV bzw. zusätzlich mit der Schiedsstelle vollkommen ignoriert worden zu sein.^{32,33}

Betrachtet man in Abbildung 1.1 die in der Nutzenbewertung involvierten Akteure, kommt man kaum umhin, Anleihen aus der Soziologie und Institutionenökonomik wie bspw. den akteurszentrierten Institutionalismus zur Untersuchung der Steuerung und Organisation von komplizierten Politikfeldern (hier die Selbstverwaltung im Gesundheitswesen) zu bemühen, und zwar im Sinne einer Forschungsheuristik. Diese soll ausgehend von einem bestimmten Sachverhalt, Rückschlüsse auf dessen Ursachen ermöglichen (Maynzt & Scharpf 1995). Dies umso mehr als Problematiken der Governance beim G-BA, insbesondere hinsichtlich der Mitwirkung des GKV-Spitzenverbandes bei der Nutzenbewertung (Huier et al. 2021) mit Auswirkungen auf die Festlegung der Vergleichstherapien und Bestimmung des Zusatznutzens aus verschiedenen Perspektiven moniert werden (Vorderwülbecke 2013). Cassel und

³² In seiner Pressemitteilung 9/2018 vom 22.6.2018 gibt der Verband der forschenden Arzneimittelhersteller (vfa e.V.) an, dass eine Dossiererstellung zur vorgeschriebenen Nutzenbewertung eines Arzneimittels im Schnitt etwa 870.000 Euro kostet (<https://www.vfa.de/de/presse/pressemitteilungen/pm-009-2018-buerokratiekosten-draengen-arzneimitteltherapien-aus-der-versorgung.html>). Berücksichtigt man dazu den auf mehrere Tausend Seiten angeschwollenen Umfang und die im Vergleich zur Zulassung aus HTA Perspektive zusätzlich anzufertigende Datenauswertung, scheinen die angegebenen Kosten durchaus realistisch. Anekdotisch kann dies der Verfasser vorliegender Arbeit aus seiner eigenen Erfahrung in der pharmazeutischen Industrie bestätigen.

³³ Weitere Kosten können im Falle der Anrufung der für das SGB V zuständigen Sozialgerichtsbarkeit entstehen.

Ulrich (2015) charakterisieren diesen Sachverhalt als „institutionelle Omnipräsenz“ des GKV-Spitzenverbands. In der Studie der Stiftung Münch zu Interessenkonflikten im G-BA (Coenen et al. 2016) stellen die Autoren fest, dass derzeit Partikularinteressen beim G-BA dominieren und fordern eine stärkere Orientierung am Gemeinwohl, insbesondere bei der Bewertung von Innovationen. Die Analyse untersuchte aus wettbewerbstheoretischer und regulierungsökonomischer Sicht, inwieweit die Struktur und Arbeitsweise des G-BA grundsätzlich geeignet sei, eine möglichst optimale und effiziente Gesundheitsversorgung zu gewährleisten. Auf der regulatorischen Governance Ebene werden die Eigenheiten des deutschen Systems insbesondere im Vergleich mit dem englischen Vorgehen ersichtlich (Wright et al. 2017), das als Vorbild für viele andere nationale HTA Systeme gilt. Letztlich weisen Gerhardus und Dintsios (2005) schon in ihrer länger zurückliegenden systematischen Übersichtsarbeit darauf hin, dass es nur wenig Erfahrung mit Methoden gibt, die eine valide Bewertung des Einflusses von HTA-Berichten auf die gesundheitspolitische Entscheidungsfindung ermöglichen und dass nur vorsichtige Aussagen hierüber getroffen werden. Sie warnen davor, von der inhaltlichen Übereinstimmung zwischen Empfehlungen aus HTA-Berichten und nachfolgenden Entscheidungen regelmäßig auf einen kausalen Zusammenhang zu schließen. Um evidenzbasierte Schlussfolgerungen über den Einfluss von HTA-Berichten ziehen zu können, wäre es notwendig, hinsichtlich ihrer Validität überprüfte Indikatoren einzusetzen und die Untersuchungen so zu gestalten, dass die Einflüsse anderer Faktoren kontrolliert werden können (Gerhardus & Dintsios 2005).

Vorliegende Dissertation analysiert die mit dem AMNOG eingeführte Nutzenbewertung von neuen Arzneimitteln ihrem prozeduralen Ablauf entlang unter Berücksichtigung der gesetzlich intendierten Zielsetzung und unter Einsatz aus ihren strukturellen Merkmalen abgeleiteter analytischer Raster.

1.2 Zielsetzung und Beitrag

Ziel vorliegender Dissertation ist zum Verständnis der komplexen frühen Nutzenbewertung mit all ihren involvierten Akteuren, Interessenkonstellationen bzw. Motivationslagen, Verfahrensschritten sowie dem dahinterstehenden methodischen Instrumentarium, dort wo es zur Anwendung kommt, beizutragen und Facetten aufzudecken, die sonst in der eher deskriptiv gehaltenen Literatur bzw. dem periodischen Berichtswesen³⁴ nicht identifiziert bzw. kaum tiefergehend analysiert werden. Des Weiteren soll etwas Licht in die Abläufe der Selbstverwaltung im Gesundheitswesen auch hinsichtlich ihrer Interaktion mit Akteuren wie

³⁴ Eine Reihe von jährlichen Berichten zum AMNOG wie bspw. der AMNOG-Report der DAK (Angestellten Krankenkasse) oder der Arzneiverordnungs-Report des WiDO (Wissenschaftliches Institut der AOK) bzw. auch Jahresberichte der pharmazeutischen Industrieverbände wie der AMNOG-Report des BPI (Bundesverbandes des Pharmazeutischen Industrie) oder von medizinischen Fachgesellschaften wie der DGHO.

der pharmazeutischen Industrie geworfen werden, da der G-BA selbst für die akademische Szene, geschweige denn für die Versicherten und Patienten, zumindest perzeptiv oft als eine Black-Box erscheint, was auch eine entsprechende Governance-Problematik nach sich zieht.³⁵ Letztlich bilden die drei Oberziele der Gesetzgebung zum AMNOG den Optionsraum für das gesamte Vorgehen und ermöglichen somit auch die Identifizierung potenzieller Bruchstellen des Systems der Selbstverwaltung im Gesundheitswesen bei der Zielerreichung.

Die Arbeit bedient sich aufgrund des komplexen Geschehens mehrerer methodischer Erklärungsansätze. Der konzeptionelle Rahmen der Arbeit wird in Abbildung 1.2 dargestellt.

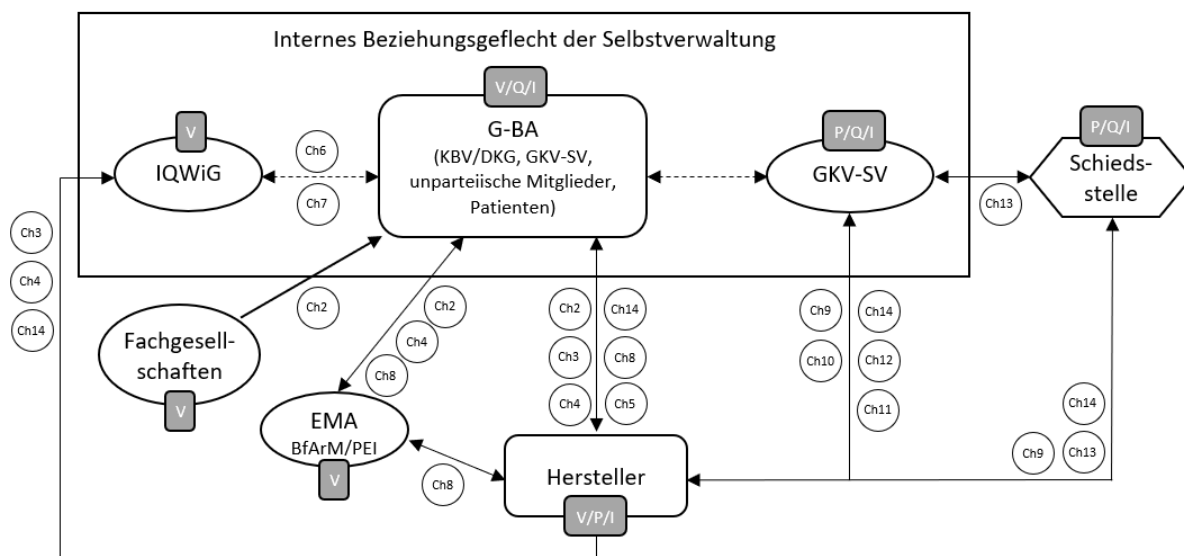


Abbildung 1.2: Konzeptioneller Rahmen der Dissertation

Notiz: gestrichelte Linien bilden das interne Beziehungsgeflecht der Selbstverwaltung ab, durchgezogene Linien Interaktionen mit externen Akteuren. Linien mit zwei Pfeilen implizieren bilaterale Interaktionen, mit einem Pfeil dagegen einseitigen Input. V: Value bzw. (Zusatz-)Nutzen, P: Preis, Q: Qualität der Versorgung und I: Innovation. Ch(N): der Beitrag jedes einzelnen Kapitels (Chapter), hinter welchem jeweils die entsprechende Publikation steht.

Im konzeptionellen Rahmen werden die Handlungsfolgen der jeweiligen Akteure auf die gesetzlichen Zielbereiche abgebildet. Da das Value-based Pricing zwei Dimensionen bereits impliziert, den (Zusatz-)Nutzen und den hiervon abgeleiteten Preis (hier Erstattungsbetrag), wird es in zwei eigene Betrachtungskomponenten aufgeteilt (Value und Preis). Die beiden anderen Zielbereiche, Arzneimittelversorgungsqualität (Q) und innovationsfördernde Rahmenbedingungen (I) sind interdependent, werden aber aufgrund der dezidierten Nennung in der Zielsetzung des AMNOG getrennt betrachtet.

Chapter 2 ist die einzige publizierte Studie zu den Erfahrungen der pharmazeutischen Industrie mit den Beratungen des G-BA im Rahmen der frühen Nutzenbewertung. Sie bedient

³⁵ So wurde der G-BA zu mehr Transparenz durch das Oberverwaltungsgericht NRW auf Klage eines Herstellers verpflichtet (Az. 8 A 467/11 vom 15.1.2014) und musste die Identität sowohl der Mitglieder des Unterausschusses Arzneimittel sowie sämtlicher Personen, die eine Stellungnahme im Rahmen des Verfahrens abgegeben hatten als auch zu den Sitzungsprotokollen und zu den Voten der Patientenvertreter offenlegen.

sich zweier qualitativer und einer quantitativen longitudinalen Erhebung. Sie zeigt, dass der G-BA über seine Erfahrungswerte die Qualität der Beratungen aus Herstellerperspektive erhöhen konnte, womit auch die These eines „lernenden“ Systems bestätigt wird. Allerdings werden auch die Bruchstellen hinsichtlich zentraler Fragestellungen der Nutzenbewertung wie die Definition der zweckmäßigen Vergleichstherapie, die Verwertung der Studienendpunkte hinsichtlich ihrer Patientenrelevanz und die zumindest in den Augen der pharmazeutischen Industrie exzessive Subgruppenaufteilung ersichtlich. Ferner wird eine niedrige Beratungskompetenz des G-BA zum Studiendesign moniert. Da die Beratung die Leitplanken für die nachfolgende Nutzenbewertung festsetzt, insbesondere bezogen auf die zweckmäßige Vergleichstherapie, ist ihr Einfluss auf das Ergebnis der Nutzenbewertung kardinal und implizit auch auf die anschließenden Erstattungsbetragsverhandlungen, bei denen die zweckmäßige Vergleichstherapie einen Preisanker darstellt. In Anbetracht der Tatsache, dass sich in Wettbewerb befindende pharmazeutische Unternehmen nicht ohne Weiteres zu eigenen Erfahrungen öffentlich äußern, konnte die Studie nur anonym und ohne direkten Produktbezug durchgeführt werden.

In Chapter 3 wird der Umgang mit dem Endpunkt gesundheitsbezogene Lebensqualität näher betrachtet. Gesundheitsbezogene Lebensqualität ist in Zeiten patientenzentrierter Versorgung von großer Wichtigkeit und erlangt vor allem in onkologischen Indikationsgebieten neben dem Überleben als klinischen Endpunkt einen enormen Stellenwert. Sie wird auch deswegen vom G-BA systematisch erfragt. Die Publikation zeigt die zwiespältige Situation, in der sich sowohl IQWiG und G-BA als auch die Hersteller befinden: einerseits wird der Endpunkt, der in Zulassungsstudien meist nur explorativ mitgeführt wird, berichtet, andererseits genügt er den an ihn gestellten sehr hohen Anforderungen des IQWiG und G-BA nicht, was zu einer unbefriedigenden Gesamtkonstellation führt. Die Publikation trägt zum Erkenntnisgewinn der Verwertung dieses Endpunktes bei, indem sie alle Bewertungen im Betrachtungszeitraum systematisch analysiert und mit Bezug auf die jeweils eingesetzten Erhebungsinstrumente formale und inhaltliche Ablehnungsgründe identifiziert und systematisch aufbereitet.

Chapter 4 geht der vergleichenden Frage nach der confirmatorischen Evidenz im Zulassungsverfahren und in der frühen Nutzenbewertung nach. Ausschlaggebend hierfür ist der direkte Bezug auf die Zulassungsstudien für die Nutzenbewertung in § 5 (3) Satz 3 der AM-NutzenV (2010), wonach für die erstmalige Bewertung nach § 35a SGB V zum Zeitpunkt der Markteinführung für die Bewertung des Arzneimittels mit neuen Wirkstoffen grundsätzlich die Zulassungsstudien zugrunde zu legen sind. Es wird durch die Gleichsetzung von confirmatorischen und explorativen Endpunkten in der Nutzenbewertung evident, dass die Nutzen-Schaden-Abwägung des IQWiG bzw. G-BA sich hinsichtlich ihrer confirmatorischen Basis vom Nutzen-Risiko-Verhältnis der Zulassungsbehörden gravierend unterscheidet. Dies

führt zu dem Paradoxon, dass dadurch die konfirmatorische Basis in der Nutzenbewertung geschmälert wird, weil Endpunkte unabhängig von ihrer Powerung in den Zulassungsstudien in ihrer Wertigkeit, insofern als patientenrelevant erachtet, gleichgestellt werden, was der klassischen Hypothesenlogik von Zulassungsstudien widerspricht. Der Beitrag der Studie liegt darin, dass durch die akribische Auswertung aller verfügbaren Zulassungsdaten und Gegenüberstellung mit den entsprechenden Nutzenbewertungen diese Sachlage bestätigt und unter Verwendung von Übereinstimmungsstatistiken überhaupt quantifiziert werden konnte.

In der nächsten Publikation, Chapter 5, werden auf Indikations- und Subgruppenebene primär formale Gründe für das Scheitern von Bewertungen im Rahmen der frühen Nutzenbewertung systematisch untersucht und klassifiziert sowie indikationsgruppenspezifisch (Miss-)Erfolgswahrscheinlichkeiten samt ihrer Präzision berechnet. Die Publikation hebt das Dilemma der beiden Welten HTA und Zulassung aus der deutschen Perspektive hervor und zeigt die Konsequenzen der unzureichenden Abstimmung dieser beiden Welten für die Nutzenbewertung auf. Sie trägt damit dazu bei, den Abstimmungsbedarf zu identifizieren und beide Seiten, HTA und Zulassung weiterhin zu sensibilisieren. Während in Chapter 4 auf die spezifische Problematik der Endpunktverwertung aus den Zulassungsstudien eingegangen wurde, bewegt sich die Analyse hier auf einer höheren hierarchischen Ebene und zeigt weitere Implikationen für die Nutzenbewertung aus den Zulassungsdaten auf. Besonders die Nichtakzeptanz von Analogieschlüssen der Zulassungsbehörden durch IQWiG bzw. G-BA, die sich auf die Labelformulierung auswirken, führen zu unterschiedlichen Konstellationen von Evidenzlücken in Bezug auf das umgesetzte Studiendesign der Zulassungsstudien aus HTA Sicht. Die Identifizierung und genaue Darstellung dieser Konstellationen ist ein weiterer Beitrag der Publikation.

Chapter 6 ist biometrisch ausgerichtet und geht den Kritikpunkten zur IQWiG Methodik bei der endpunktspezifischen Zusatznutzenquantifizierung nach. Ergänzend zum vorherigen Chapter 5 wird hier nicht das rein formale, sondern das methodisch inhaltliche Vorgehen des Auftragsinstituts der Selbstverwaltung in den Vordergrund gestellt. Empirisch lassen sich einige der von unterschiedlichen Akteuren, darunter auch Biostatistikern und Gesundheitsökonomen, geäußerten Kritikpunkte nicht bestätigen (bspw. die Bevorzugung bei der endpunktbezogenen Zusatznutzenausprägung von Indikationen, die in ihren Zulassungsstudien den Mortalitätsendpunkt mitführen). Andererseits wird aus der Analyse ersichtlich, dass das methodische Vorgehen aufgrund seiner expliziten Annahmen so ausgerichtet ist, dass es zwangsläufig zu einem Powerverlust führt. Dies lässt sich sowohl für die Annahme zweier vorhandener pivotalen Studien als auch für die Dichotomisierung von stetigen Endpunkten bestätigen, letzteres mit Hilfe einer Simulation. Die Studie unterbreitet einen Vorschlag, wie durch Modifikationen der IQWiG Methodik einige der stichhaltigen Kritikpunkte ausgemerzt werden könnten. Neben ihrem Beitrag zur Methodendiskussion und -weiterentwicklung

leistet sie auch einen Beitrag dazu, das in den vorausgegangenen Chapters herausgearbeitete Dilemma zwischen Zulassungs- und HTA Welt aufzulösen und somit Teile der verworfenen Evidenz aus der Zulassung für HTA Zwecke zugänglicher zu machen.

Chapter 7 geht in seiner Betrachtung explizit auf das interne Beziehungsgeflecht der Selbstverwaltung ein (hier IQWiG versus G-BA bzw. ‚Assessment‘ versus ‚Appraisal‘). Aus dem Ergebnisvergleich ergeben sich relativ starke Unterschiede bei der Zusatznutzenausprägung und ihrer Wahrscheinlichkeit (die adäquate methodische Terminologie müsste hier Ergebnissicherheit lauten), die vom IQWiG empfohlen wird, und dem beschlossenen Zusatznutzen samt seiner „Wahrscheinlichkeit“ durch den G-BA, insbesondere auf der aussagekräftigeren Subpopulationsebene. Das Phänomen ist viel stärker an den Randwerten des Zusatznutzenspektrums (kein Zusatznutzen und erheblicher Zusatznutzen) ausgeprägt, wobei der G-BA mit seinen Beschlüssen glättend zu fungieren scheint. Die Analyse wird mit entsprechenden Übereinstimmungsstatistiken untermauert. Im ordinal logistischen Regressionsmodell zur Untersuchung potenziell diese Unterschiede zwischen IQWiG und G-BA erklärender Einflussfaktoren lassen sich auch solche finden, die nicht gesetzlich genannt werden wie bspw. ungedeckter medizinischer Bedarf, das Votum von Fachgesellschaften aber auch die Kosten der zweckmäßigen Vergleichstherapie. Letzteres stellt den ersten empirischen und nicht nur rein ordnungspolitischen Befund für eine Governanceproblematik in der frühen Nutzenbewertung dar, da der GKV-SV ein genuines Interesse aufgrund der nachfolgenden Erstattungsvertragsverhandlungen an niedrigeren Kosten der zweckmäßigen Vergleichstherapien hat und mit seiner Beteiligung als stimmberechtigte Bank im Unterausschuss Arzneimittel des G-BA zum Zusatznutzen neuer Arzneimittel mitbeschließt. Genau in der Identifizierung und Quantifizierung über das Regressionsmodell weiterer entscheidungsrelevanter Faktoren und deren Einflussstärke zur Abweichung der eher „politischen“ Beschlussfassungen des G-BA vom algorithmischen Vorgehen des IQWiG liegt der Beitrag dieser Publikation.

Chapter 8 führt uns wieder zurück in das externe Beziehungsgeflecht (siehe Abb. 1.2) des G-BA. Es vergleicht die Auflagen des G-BA bei befristeten Beschlüssen zum Zusatznutzen mit entsprechenden Auflagen der Zulassungsbehörden für beauftragte Zulassungen (engl. ‚conditional approvals‘), die auch bei der Nutzenbewertung berücksichtigt werden. Die Publikation trägt dazu bei, durch exzessive Dokumentanalyse Unterschiede bei den Beauftragungen von Zulassungen und Nutzenbewertungen hinsichtlich der weiter zu generierenden Evidenz herauszuarbeiten und die dahinterstehenden Rationales ausfindig zu machen. Die Auflagen des G-BA bei seinen befristeten Beschlüssen übertreffen diejenigen der Zulassung, anders als bei der Zulassung fehlt es beim G-BA allerdings oft an handlungsanleitenden Richtlinien (engl. Guidance) hierzu.

Während Chapter 2 bis Chapter 8 sich hauptsächlich mit der Nutzendimension (in Abb. 1.2 mit V für Value gekennzeichnet) beschäftigen, tangiert Chapter 9 zu den Marktrücknahmen als eine Art Brückenabschnitt neben der Nutzendimension (V) auch die der Arzneimittel-

qualität (Q) und damit einhergehend auch die der Innovationssicherung (I), wobei ätiologisch betrachtet die Nutzendimension (V) und der Preis (P) im Sinne des Erstattungsbetrags in das Geschehen mit involviert sind. Implizit kommt auch ein weiterer Akteur ins Spiel, da die Marktrücknahmen nach bzw. aufgrund der Nutzenbewertung den jeweiligen Empfehlungen der medizinischen Fachgesellschaften auf nationaler und auf europäischer Ebene gegenübergestellt werden. Diese Empfehlungen rühren aus den publizierten Leitlinien dieser Fachgesellschaften her. Die Arbeit liefert einen Beitrag zur Identifizierung potenzieller Versorgungslücken mit mehr oder minder die Versorgungsqualität mit Arzneimitteln beeinflussenden Konsequenzen, was einer Zielkollision gleichkommt, betrachtet man die originäre Zielsetzung des AMNOG, die sowohl wirtschaftliche Arzneimittelpreise als auch die Verbesserung der Versorgungsqualität verfolgt. Wie bei den weiter vorhin beschriebenen Dilemmata zwischen Zulassung und (Zusatz-)Nutzenbewertung als Ausfluss unterschiedlicher Welten mit unterschiedlichen, allerdings auf der gleichen Studienevidenz fußenden Fragestellungen (Wirksamkeit, Sicherheit und Qualität versus Zusatznutzen), kommt es hier zu einem erneuten Dilemma, diesmal zwischen Fachgesellschaften bzw. deren Leitlinienempfehlungen und der Nutzenbewertung. Und auch diesmal zeigen sich trotz aller Heterogenität unter den Leitlinien die gleichen Bruchstellen wie bspw. Endpunktrelevanz aus der herangezogenen Studienevidenz. Ein wichtiger Unterschied besteht in der Ebene, in welcher das Dilemma greift. Während bei der Gegenüberstellung von Zulassung und Nutzenbewertung die Mesoebene der Entscheidungsfindung involviert ist (Zulassungsbehörden und G-BA), werden versorgende Ärzte auf der Mikroebene dem Dilemma einer potenziellen Versorgungslücke aufgrund der Marktrücknahmen ausgesetzt. Die Substitution, der nicht mehr erhältlichen neuen Arzneimittel kann, muss aber nicht je nach vorhandenen Alternativen zu einer Abnahme der Versorgungsqualität führen. Ist der auslösende Grund der Rücknahme alleinig im Preisgeschehen zu finden, so spielt auch die Marktmacht der Kostenträger (GKV-SV) eine ausschlaggebende Rolle hierfür.

In den folgenden vier Abschnitten (Chapter 10 bis Chapter 13) der Dissertation liegt das Augenmerk hauptsächlich auf dem Erstattungsbetrag (P in der Abb. 1.2 zum konzeptionellen Rahmen) und seine konstituierenden Faktoren in Abhängigkeit vom Arzneimittel (Orphan Drugs und Arzneimittel ohne Orphan Drug Designation), den Akteuren (Schiedsstelle) und dem Studiendesign (additiv auf eine vorhandene Intervention versus Substitution einer vorhandenen Intervention), betrachtet immer unter dem Diktum des AMNOG und der für die Erstattungsbetragsverhandlungen ausschlaggebenden Rahmenvereinbarung zwischen dem GKV-Spitzenverband und den pharmazeutischen Verbänden in ihrer gültigen Fassung. Für die Abschnitte Chapter 10 bis Chapter 12 ist die Datenfindung und -aufbereitung mit einem relativ großen Aufwand verbunden, da viele der Daten nicht ohne Weiteres zugänglich sind. Insbesondere europäische Vergleichspreise, wenn auch ohne Rabatte als reine Listenpreise,

bedurften der Sichtung etlicher internationaler Datenbanken und damit einhergehend der Eruierung der jeweiligen Dokumentationslogik dieser Datenbanken. Auch mussten einige auf eigene Heuristiken gründende Annahmen zur Markteintrittsreihenfolge von Arzneimitteln in anderen europäischen Ländern und der dadurch abgeleiteten Preisverfügbarkeit für die Erstattungsbetragsverhandlungen zwischen Herstellern und dem GKV-SV plausibilisiert werden.³⁶

Chapter 10 zu den Prädiktoren der verhandelten und geschiedsten Erstattungsbeträgen von Arzneimitteln ohne Orphan Drug Designation bestätigt die Umsetzung der Kriterien der Rahmenvereinbarung zu einem großen Teil. Die Publikation trägt auch dazu bei, die Nähe des Vorgehens zum Value-based Pricing zu bestimmen. Sowohl der Taxonomie von Sussex et al. (2013) als auch den Kriterien nach Jommi et al. (2020) folgend, lässt sich das AMNOG den Value-based Pricing Ansätzen zuordnen, weil es den Zusatznutzen und den Einfluss auf das zukünftige Ausgabengeschehen berücksichtigt, letzteres implizit in den Regressionsmodellen über die Zielpopulationsgröße und die Jahrestherapiekosten der zweckmäßigen Vergleichstherapie operationalisiert. Darüber hinaus spielen auch andere Faktoren außerhalb der Rahmenvereinbarung eine Rolle beim Verhandlungsgeschehen, die aber auch aufgrund seiner Vertraulichkeit verborgen und damit spekulativ bleiben.

Chapter 11 beschäftigt sich explizit mit den Determinanten der Erstattungsbetragsfindung für Orphan Drugs, die einen besonderen Status im Rahmen der frühen Nutzenbewertung genießen, da ihnen qua Gesetz (§ 35a (1) S.11 SGB V) der Zusatznutzen mit ihrer Zulassung zugesprochen wird. Somit erhalten alle Orphan drugs einen Zusatznutzen unterschiedlicher Ausprägung, aber gleichzeitig entfällt für sie der Preisanker der Jahrestherapiekosten der zweckmäßigen Vergleichstherapie, da sie keinen vom G-BA bestimmten Komparator in ihren Dossiers führen müssen. Die Studie enthält die einzige hierzu publizierte Analyse für den deutschen Arzneimittelversorgungskontext und testet mit zwölf Hypothesen zum Einfluss potenzieller Faktoren auf den verhandelten Erstattungsbetrag für Orphan Drugs eine relativ große Bandbreite an Variablen, die von den Charakteristika des Herstellers (bspw. Hauptsitz und Verhandlungserfahrung mit dem GKV-Spitzenverband) und des Arzneimittels (Indikation und Schwere der Erkrankung etc.), über die bekannten Kriterien der Rahmenvereinbarung bis hin zur Antizipierung von erneuten Verhandlungen und der Marktpenetration reichen. Ihr Beitrag hat auch internationale Tragweite, da sie aufgrund reichlich vorliegender Literatur zu anderen Ländern internationale Vergleiche mit der Preisfindung von Orphan Drugs in einem der größten Arzneimittelmärkte der Welt ermöglicht.

³⁶ durch die langjährige Tätigkeit des Verfassers vorliegender Dissertation für die pharmazeutische Industrie im Bereich Marktzugang (Market Access) sind diese Heuristiken produktbezogen über eine große Anzahl von neuen Arzneimitteln entstanden.

In Chapter 12 wird die Haupthypothese getestet, ob das Studiendesign (additiv vs. substitutiv) als solches einen Einfluss auf das Ergebnis der Erstattungsbetragsverhandlungen nimmt. Dazu werden fünf weitere Subhypothesen abgeleitet und getestet. Die Studie liefert einen doppelten Beitrag zur Transparentmachung von Teilaspekten der Erstattungsbetragsverhandlungen: sie zeigt den Einfluss des Studiendesigns auf die Erstattungsbeträge in Abhängigkeit vom Status des Komparators (palliativ und damit meistens niedrigpreisig versus kurativ und damit meistens hochpreisig) als eigenständigen Faktor neben den in der Rahmenvereinbarung genannten Kriterien und unabhängig vom Ausgabeneinfluss (engl. Budget Impact), auch wenn sein Anteil an der erklärenden Varianz in den berechneten Regressionsmodellen relativ klein bleibt und sie bestätigt das mischkalkulatorische Vorgehen auf Subpopulationsebene für die untersuchten Onkologika. Darüber hinaus präsentiert sie zwei Operationalisierungen von Zuschlägen auf die Jahrestherapiekosten der zweckmäßigen Vergleichstherapie (additiv und multiplikativ), die jeweils unterschiedliche Vorteile in der Dokumentation des Preisgeschehens besitzen und somit auch tiefergehende Erkenntnisse zur untersuchten Fragestellung ermöglichen.

Chapter 13 befasst sich ausschließlich mit den Entscheidungen der Schiedsstelle und ist auch die erste internationale Publikation zu dieser Thematik. Sie trägt dazu bei sowohl das Verständnis zum Vorgehen der Schiedsstelle zu verbessern als auch die Verortung dieses Akteurs in Bezug auf die gesetzliche Zielsetzung zu ermöglichen. Neben den rein preislich relevanten Aspekten werden auch weitere Inhalte der Schiedssprüche im Betrachtungszeitraum analysiert (bspw. Laufzeit der Vereinbarungen). Obwohl eigentlich Schiedssprüche Einzelfallentscheidungen dem Willen des Gesetzgebers folgend sein sollen, lassen sich quasi-algorithmische Elemente erkennen, die durch die relativ enge Orientierung an die Rahmenvereinbarung zwischen Herstellern und GKV-SV sogar einer idealtypischen Umsetzung des Value-based Pricings näherkommen, als die verhandelten Erstattungsbeträge.

Das abschließende Kapitel der kumulativen Dissertation, Chapter 14, ist eine Fallstudie zur Nutzenbewertung und zu den entsprechenden Erstattungsbeträgen von Arzneimitteln, die in der Augenheilkunde angewendet werden. Der Vorteil dieser Fallstudie liegt darin, dass die Anzahl der betrachteten Arzneimittel überschaubar bleibt, aber dennoch alle bekannten Konstellationen im AMNOG abgebildet werden (Orphan Drug Bewertungen, Marktaustritte, Mehrfachbewertungen des gleichen Wirkstoffs in der zuvor bewerteten Indikation und Bewertungen unterschiedlicher Indikationen desselben Wirkstoffs, Schiedsstellenentscheide, Off-label Wettbewerb, indirekte Vergleiche, alternative Studiendesigns zu den etablierten randomisiert kontrollierten Studien, Monotherapien und Kombinationen u.v.a.). Die Fallstudie liefert auch einen Vergleich mit Frankreich als weiteres Land ohne etablierte gesundheitsökonomische Evaluationen in der Bewertung von Arzneimitteln und untersucht inwieweit die Nutzenbewertung in der realen Patientenversorgung ankommt.

1.3 Überblick

Die rein empirisch ausgerichtete Dissertation enthält dreizehn eigenständige Abschnitte, die jeweils eine Publikation zu einer der Dissertation zuzuordnenden Thematik repräsentieren. In den folgenden Unterabschnitten werden die dreizehn eigenständigen Abschnitte kurz zusammengefasst und ihre Hauptergebnisse diskutiert.

Chapter (2): Industry's Experiences with the Scientific Advice offered by the Federal Joint Committee within the Early Benefit Assessment of Pharmaceuticals in Germany

Die Durchführung von Beratungen zwischen Arzneimittelherstellern und Health Technology Assessment-Organisationen ist international tägliche Praxis (Cucho et al. 2014; Longson et al. 2014; Wonder et al. 2013). In Deutschland sind die Beratungen der Hersteller durch den G-BA im Rahmen der frühen Nutzenbewertung von Arzneimitteln in §35a (7) SGB V geregelt und in §8 (1) der dazugehörigen Verordnung (AM-NutzenV) konkretisiert. Hiernach berät der G-BA den jeweiligen Arzneimittelhersteller aufgrund dessen Anforderung auf der Grundlage eingereichter Unterlagen. Von 2011 bis Ende 2014 wurden 260 Beratungsgespräche mit den pharmazeutischen Herstellern geführt (Hecken 2014a). Zu den Erfahrungen der involvierten Akteure wird aus G-BA Seite aufgrund der Vertraulichkeit des Verfahrens nichts berichtet. Die Sicht der Hersteller wurde bisher nur anekdotisch zusammengetragen (Cassel & Heigl 2013; Sattelmeier et al. 2013), womit vorliegende Publikation als erste überhaupt belastbare Daten zu diesem Vorgehen aus der Perspektive eines der involvierten Akteure liefert.

Im Verlauf des Jahres 2012 wurde vorbereitend für die quantitative Erhebung in einer Reihe von Kontakten mit Einzelvertretern forschender Arzneimittelhersteller das entsprechende Meinungsbild zu den Beratungserfahrungen qualitativ festgehalten. Anschließend wurde ein eigens hierfür entwickelter Kurzfragebogen zu unterschiedlichen Zeitpunkten im Zeitabschnitt April 2013 bis März 2015 eingesetzt, um die Informationsbasis zu erweitern und potenzielle Änderungen im longitudinalen Verlauf hinsichtlich des Beratungsvorgehens zu untersuchen. Der Bogen enthielt acht Frageentitäten zur Art der Beratung, den weiteren Teilnehmern, den Beratungsthemen, zu Details hinsichtlich der zweckmäßigen Vergleichstherapie, Subgruppen sowie weiterer relevanter Punkte wie Studiendesign und -dauer und zum Aufwand für die Gesprächsvorbereitung, wobei auch Freitextantworten möglich waren. Es erfolgte eine anonymisierte Auswertung hinsichtlich Hersteller und deren Arzneimittel. Abschließend wurde im Zeitraum April bis Juni 2017 mit 6 Teilnehmern der ersten qualitativen Erhebung und einem weiteren Unternehmensvertreter ein erneutes Interview geführt, um deren Sicht 6 Jahre nach Einführung der Beratungsgespräche zu eruieren.

Im Rahmen der vorgeschalteten qualitativen Erhebung bis Ende 2012 wurden folgende Punkte hervorgehoben: (i) bei vielen Vertretern des G-BA lassen sich fehlende Kenntnisse

hinsichtlich zulassungsrelevanter Anforderungen zu klinischen Studien ausmachen; (ii) es kam durchaus öfter zu inkonsistenten Aussagen während der Beratungsgespräche; (iii) die Argumentation des G-BA wurde oft als dogmatisch empfunden. Als positiv wurde die Möglichkeit der Teilnahme von Vertretern der globalen Zentralen bei simultaner Übersetzung erachtet sowie erzielte Lernkurveneffekte bei mehreren Beratungsgesprächen.

Für den quantitativen Teil gingen insgesamt 61 ausgefüllte Fragebögen von 19 Herstellern bis Anfang 2015 ein (25% der durchgeführten Beratungen). Mit der vom G-BA im Rahmen des Beratungsgesprächs genannten zweckmäßigen Vergleichstherapie waren die Hersteller in 70% der Fälle einverstanden, wobei der G-BA die von den Herstellern vorgeschlagene zweckmäßige Vergleichstherapie auch zu 57% akzeptierte. Bei der Befragung zu den angenommenen Gründen des G-BA hinsichtlich seiner Benennung der zweckmäßigen Vergleichstherapie zeigte sich in 53% der Fälle die Wirtschaftlichkeit als ausschlaggebende Motivation. Zulassungsfragen spielten wiederum in 68% der Fälle auch eine Rolle. Mit Bezug auf die Subpopulationsaufteilung gaben die Hersteller in 24% der Fälle an, dass diese nicht sachgerecht sei. In 69% wurden die Beratungen im Allgemeinen als eindeutig erachtet, in 75% als vollständig. Hinsichtlich der Positionsänderung des G-BA zeigte sich in 4% der Beratungen eine direkte Änderung, in 11% erfolgte sie erst über eine Schleife durch den Unterausschuss Arzneimittel und in 85% beharrte der G-BA auf seine Position auch nach dem Beratungsgespräch.

Der Aufwand für die Gesprächsvorbereitung konnte monetär oder in Vollzeitäquivalenten angegeben werden. Dieser variierte in Abhängigkeit von der Beratungsart zwischen 5.000 € und 90.000 € ohne Einbeziehung der Beratungsgebühren bzw. zwischen 1,6 und 18 Vollzeitäquivalenten (mit Monatsbezug).

Mit Ausnahme der Fragen zu Studienart, Einverstanden sein mit der zweckmäßigen Vergleichstherapie und sachgerechtem „Slicing“ zeigten sich im zeitlichen Verlauf keine statistisch signifikanten Unterschiede. Bei den Antworten zur Studienart scheinen sich beim G-BA Routinen herausgebildet zu haben. Die erneute qualitative Erhebung bei sieben Herstellern bestätigte die Ergebnisse der Fragebogenerhebung.

Die zunehmende Zustimmung der Hersteller mit der vom G-BA in der Beratung genannten zweckmäßigen Vergleichstherapie hängt u. a. auch mit dem 3 AMG-ÄndG vom Juni 2013 zusammen, das die Auswahl der zweckmäßigen Vergleichstherapie durch den Gesetzgeber flexibler gestaltet. Insofern Beratungen für Hersteller den einzigen Weg darstellen, um die konkreten Anforderungen des G-BA an die erforderliche Evidenz für einen Wirkstoff vor Dossiererstellung bzw. Studienplanung zu klären, erlangen sie eine eminent wichtige Bedeutung im Verfahren. Trotz der verbesserten Gesprächsatmosphäre leidet das Vorgehen in Deutschland immer noch zum Teil an einem fehlenden tiefergehenden fachlichen Dialog

zwischen G-BA, Herstellern, Wissenschaft und Zulassungsbehörden. Die mündliche Beratung des Herstellers erfolgt durch nicht entscheidungsbefugte Mitglieder seiner Geschäftsstelle. Die Möglichkeit, dass die Geschäftsstelle im Nachhinein nochmals den Unterausschuss Arzneimittel kontaktiert und dessen Entscheidung ändert, wurde bislang weder in der Verfahrensordnung umgesetzt noch ersetzt sie den fachlichen Dialog. Nachträgliche Änderungen sind nur in seltenen Ausnahmefällen erfolgt. Dadurch, dass studienrelevante Vorgaben bei der Beratung in vielen Fällen nicht klar benannt werden, entsteht neben der Planungsunsicherheit bei den Herstellern ein maximaler Spielraum für den G-BA, relevante Vorgaben post-hoc zu definieren und damit Entscheidungen erst am Verfahrensende der frühen Nutzenbewertung zu begründen. Der G-BA zeigt bislang eine ausgesprochen geringe Bereitschaft, seine Anforderungen als Ergebnis eines wissenschaftlichen Dialogs mit den Herstellern und weiteren Akteuren zu modifizieren. Anders als beim englischen NICE, wo eine strikte Trennung der Beratung und anschließenden Bewertung sowohl auf Verfahrensebene als auch auf Personenebene vorgesehen ist (Maignen et al. 2017), liegen in Deutschland beide Schritte unter der Obhut des Unterausschusses Arzneimittel beim G-BA, was mit einer reduzierten Unabhängigkeit einhergeht und auch auf eine potenzielle Governanceproblematik zurückschließen lässt.

Aus Perspektive der nationalen Zulassungsbehörden BfArM und PEI erzeugt die Konstellation eines europäisch und sogar weitergehend international harmonisierten Zulassungsverfahrens bei weitgehend nationalen Erstattungsregeln zwangsläufig Friktionen, die durch eine gute gemeinsame Beratung im Vorfeld der Studienplanung zwar nicht verhindert, aber doch verringert werden können (Beinlich et al. 2015). Die Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) verweist im Hinblick auf den Verfahrensablauf und Aspekte der frühen Nutzenbewertung einschließlich Festlegung der zweckmäßigen Vergleichstherapie und der Priorisierung von Endpunkten auf Verbesserungsbedarf hin (AWMF 2015). Bei der Festlegung der zweckmäßigen Vergleichstherapie wurde laut AWMF in nicht wenigen Fällen von den Vergleichstherapien der Zulassungsstudien oder der Zulassungsverfahren der EMA abgewichen. Dabei wurden auch Arzneimittel empfohlen, die zwar eine Zulassung für die Indikation haben, aber in aktuellen nationalen oder internationalen Leitlinien aufgrund fehlender Evidenz nicht empfohlen werden.

Eine aktivere Einbindung weiterer Akteure wie nationale Zulassungsbehörden und medizinische Fachgesellschaften sowie prozedurale Anleihen aus anderen Systemen mit seit längerem etablierten und damit zielführenden Beratungsgesprächen, beispielsweise im Sinne eines „scoping workshops“, scheinen zur Optimierung des Verfahrens geboten.

Chapter (3): Gesundheitsbezogene Lebensqualität im Rahmen der frühen Nutzenbewertung von Arzneimitteln nach § 35a SGB V: Ein Endpunkt mit vielen Herausforderungen für alle beteiligten Akteure

Im Rahmen der frühen Nutzenbewertung werden die positiven und negativen Effekte des zu bewertenden Arzneimittels in Relation zu einer vom G-BA definierten zweckmäßigen Vergleichstherapie (zVT) im Anwendungsgebiet von Interesse bemessen. Substanziell ist in diesem Zusammenhang die positive Beeinflussung patientenrelevanter Endpunkte wie Mortalität, Morbidität und gesundheitsbezogene Lebensqualität. Hinsichtlich des Endpunktes gesundheitsbezogene Lebensqualität wird wiederkehrend ein unzureichender Berücksichtigungsgrad bei der Erhebung und Verwertung der Zielgröße beklagt (Faber 2014; Hecken 2014b; Klakow-Franck 2014; Kvitkina et al. 2014). Zugleich werden die methodischen Herausforderungen und Limitationen betreffend die Erhebung dieses Endpunktes an verschiedenen Stellen angeführt und diskutiert. Es stellt sich somit die Frage, welche Geltung diese Zielgröße im bisherigen Verlauf der frühen Nutzenbewertung erfahren hat. Die, einen primär akteurszentrierten Ansatz folgenden, Arbeit liefert neben der deskriptiven Analyse zum Umgang mit dem Endpunkt gesundheitsbezogene Lebensqualität Erkenntnisse zur motivationalen Ausrichtung der involvierten Akteure (IQWiG, G-BA und Hersteller). Dies ist deshalb von Bedeutung, als ein solches Analyseraster in Zusammenhang mit der Verwertung von unterschiedlichen Endpunkten im Rahmen der frühen Nutzenbewertung kaum systematisch Anwendung findet bzw. in der Vergangenheit gefunden hat.

Um die Stellung der gesundheitsbezogenen Lebensqualität im Bewertungsprozess zu bestimmen, wurden die veröffentlichten Dokumente zu den ersten 90 Verfahren untersucht. Dabei wurde zunächst auf unterschiedlichen Ebenen der Umgang mit dem Endpunkt herausgearbeitet, um die Befunde sodann synoptisch zu bewerten.

Im Rahmen der Untersuchung konnten zunächst 59 Verfahren identifiziert werden, in welchen der Endpunkt gesundheitsbezogene Lebensqualität im zugrunde liegenden Studiensetting erhoben wurde. In 53 dieser Verfahren sind die Angaben schließlich zu einer Zusatznutzenaussage herangezogen worden. Anschließend wurde herstellerseitig in 30 Fällen ein Zusatznutzen tatsächlich auch abgeleitet. Davon wiederum 18 Mal in der für die Gesamtaussage maßgeblichen Operationalisierung (AM-NutzenV 2010). Es verbleiben letztlich 4 Verfahren (Crizotinib, Afatinib, Trastuzumab Emtansin sowie das ausschließlich vom G-BA bewertete Orphan Drug Ivacaftor), in denen ein positives Resümee zum Endpunkt gesundheitsbezogene Lebensqualität gezogen wurde.

Insgesamt wurden 47 unterschiedliche Instrumente eingesetzt, darunter auch 3 nicht etablierte Eigenentwicklungen. Neben einer fehlenden statistischen Signifikanz der Ergebnisse werden vielfach methodische Limitationen bei der Erhebung der gesundheitsbezogenen Lebensqualität als Ablehnungsgrund für die Einbeziehung in die Nutzen-

bewertung angeführt: hohe Anteile von in der Auswertung unberücksichtigten Patienten, geringe Rücklaufquoten, Verzerrungspotenziale infolge eines offenen Studiendesigns bzw. Imbalancen zwischen den Behandlungsarmen, nicht validierte (Folge-)Versionen eingesetzter Messinstrumente, die alleinige Darstellung des EQ-5D-Summscores sowie eine nicht valide beurteilbare klinische Relevanz des eingesetzten Responsekriteriums.

Die gewonnenen Erkenntnisse aus der Untersuchung sind vielschichtig. Im Grundsatz zeigt sich die Schwierigkeit, eine qualitative Zielgröße zu quantifizieren und deren Wertigkeit zu verorten. Erkennbar ist, dass der Endpunkt gesundheitsbezogene Lebensqualität in einer Reihe von Verfahren a priori nicht erhoben wurde. Liegen hingegen Daten zur gesundheitsbezogenen Lebensqualität vor, so kann vielfach ein Zusatznutzen nicht abgeleitet werden. In den verbleibenden Fällen schließlich wird ein solcher in der Mehrzahl nicht anerkannt. Gleichwohl die Einbeziehung der gesundheitsbezogenen Lebensqualität im Studiensetting zuzunehmen scheint, ist allzu häufig eine explorative Konzeptualisierung des Endpunktes auszumachen. Entsprechend heterogen stellt sich das vorgefundene Evidenzniveau dar.

Jedenfalls zeugt die Erwartungshaltung des IQWiG bzw. des G-BA zum Zeitpunkt der frühen Nutzenbewertung, zu welchem grundsätzlich die Zulassungsstudien zugrunde zu legen sind, diesen Endpunkt in einer entsprechenden Datenqualität wie zur Mortalität oder zu Morbiditätsempunkten im Rahmen von Zulassungsstudien vorzufinden, im besten Fall von einem unbegründeten Optimismus, in einer weniger positiven Interpretation hingegen von einem systematischen Verkennen der verfügbaren Evidenzlage zu diesem Zeitpunkt. Auch die auf den ersten Blick einleuchtende, den internationalen Kontext von Zulassungsstudien allerdings ausblendende und von der kontextspezifischen Definition der zweckmäßigen Vergleichstherapie(n) getriebene, Forderung des IQWiG nach Durchführung von Studien mit Bezug auf die Versorgungsrealität in Deutschland, scheint in diesem Zusammenhang kaum von Relevanz, scheitert die Einbringung des Endpunktes gesundheitsbezogene Lebensqualität in der frühen Nutzenbewertung bereits an dem verwendeten Instrument – unabhängig vom abgebildeten Versorgungskontext. Damit lässt sich aus den gemachten Erfahrungen auch die marginale Anzahl an postregulatorischen Studien, die eigens für Deutschland zur Erhebung der gesundheitsbezogenen Lebensqualität konzipiert wurden, teilweise erklären.

Bei onkologischen Behandlungsschemata in fortgeschrittenen Linien ist neben der zulassungsbedingten geringeren Relevanz der Lebensqualitätsdaten und der damit einhergehenden geringeren Wahrnehmung durch die Studienärzte („staff oversight“) auch von einer ausgeprägten patientenseitigen Ablehnung des Ausfüllens von entsprechenden Fragebögen – studienbelegt mit einer Rücklaufquote von nur 52% – auszugehen (Bonomi et al. 2000; Fairclough 2010). Dies begründet auch die Forderung der Deutschen Gesellschaft für

Hämatologie und Medizinische Onkologie (DGHO), Auswertungen zu PROs in onkologischen Indikationen zu akzeptieren, wenn mehr als 50% der Fragebögen zu den geforderten Zeitpunkten ausgefüllt wurden (DGHO 2013).

Nach Auffassung der Autoren der Arbeit soll nicht primär das methodische Instrumentarium, das vom IQWiG (und vom G-BA bei Orphan Drugs) zur Anwendung kommt, hinsichtlich seiner Eignung angezweifelt werden, sondern vielmehr eine gebührende Berücksichtigung der zu diesem Zeitpunkt kontextbedingten Evidenzlage zum Endpunkt gesundheitsbezogene Lebensqualität als restringierende Größe einer entsprechend evidenzbasierten Nutzenbewertung breiter unter den involvierten Akteuren zur Diskussion kommen.

Um die Wertigkeit des Endpunktes innerhalb der frühen Nutzenbewertung zu erhöhen, ist es notwendig, im Konsens und unter besonderer Berücksichtigung der Patientenbelange verlässliche (Mess-)Standards zu etablieren. Zudem sollten die Wertungskriterien transparent gemacht und eine Harmonisierung mit dem Zulassungsprozess angestrebt werden. Obwohl es keinen ‚Goldstandard‘ bei der Messung gesundheitsbezogener Lebensqualität geben kann, ist es entscheidend, das Potenzial subjektiv bemessener Parameter in den Bewertungsverfahren stärker zu achten.

Chapter (4): Confirmatory versus explorative endpoint analysis: Decision-making on the basis of evidence available from market authorization and early benefit assessment for oncology drugs

Der Artikel untersucht die Konsequenzen des nicht abgestimmten Zusammenspiels von Zulassung und früher Nutzenbewertung hinsichtlich der jeweiligen Endpunktverwertung aus denselben Evidenzquellen (Zulassungsstudien), auf welche (primär) beide Verfahren fußen, anhand nutzenbewerteter Arzneimittel in onkologischen Indikationen. Er folgt einem akteurszentrierten Ansatz unter Einbeziehung der Zulassungsbehörden neben den zentralen Akteuren des AMNOG.

Die Wirksamkeit, Unbedenklichkeit und Qualität eines neuen Arzneimittels gelten mit der Zulassung als belegt (§ 7 (2) AM-NutzenV). Mit der Einführung des AMNOG am 01.01.2011 muss für die zeitversetzte Erstattungsbeitragsverhandlung eines neuen Arzneimittels die Wirtschaftlichkeit durch die frühe Nutzenbewertung nach §35a SGB V nachgewiesen werden. Die neuen situativen Gegebenheiten stellen den pharmazeutischen Unternehmer vor die Herausforderungen den diskrepanten Anforderungen von Zulassung und früher Nutzenbewertung bezüglich der Konzeption von Studien gerecht zu werden (Beinlich et al. 2015). Dies ist jedoch nicht immer möglich und bedingt häufig eine nachteilige Beurteilung in der frühen Nutzenbewertung, welche sich negativ auf den nationalen sowie internationalen Erstattungspreisen aufgrund der internationalen Referenzierung auswirken kann.

Die Differenzen der Beurteilungsverfahren innerhalb der beiden Prozesse bei der Betrachtung derselben Studien zeigen sich unter anderem im Umgang mit konfirmatorischen und explorativen Endpunkten. Mit der Gleichsetzung dieser beiden Arten von Endpunkten innerhalb der frühen Nutzenbewertung durch die angewandte Methodik (IQWiG 2015) werden konfirmatorische Endpunkte oftmals nicht berücksichtigt, obwohl diese im Rahmen der Zulassung anerkannt werden (Ruof et al. 2014a; Ruof et al. 2014b). Dies führt zu einem Informationsverlust für die Zusatznutzenbewertung, denn konfirmatorische Endpunkte (d.h. primäre bzw. ko-primäre Endpunkte) beinhalten, bedingt durch ihre Präspezifikation, entscheidende hypothesenüberprüfende Ergebnisse zum Primärstudienziel (EMA 1998).

Die Problemstellung wurde weitestgehend in der aktuellen Literatur vor allem für onkologische innovative Arzneimittel diskutiert (Dabisch et al. 2014; Ruof et al. 2014a). Bisher ist jedoch nicht nähergehend untersucht worden, in welchem Ausmaß explorative im Verhältnis zu konfirmatorischen Endpunkten im Kontext der frühen Nutzenbewertung und in der Zulassung verwendet werden. Um einen Vergleich zwischen der Zulassung und der frühen Nutzenbewertung herzustellen, werden die Nutzen-Risiko-Bewertung aus dem Assessment Report der European Medicines Agency (EMA) und die Nutzen-Schaden-Saldierung aus der Dossierbewertung des IQWiG für alle bis zum 17.07.2015 abgeschlossenen onkologischen frühen Nutzenbewertungsverfahren betrachtet und die folgenden zwei Hypothesen überprüft: (i) das Zulassungsverfahren ist in höherem Maße konfirmatorisch als die frühe Nutzenbewertung nach §35a SGB V und (ii) beide Bewertungsverfahren sind prozedural betrachtet nicht-konfirmatorisch.

Die Endpunkterfassung basiert auf einem eigens konzipierten Leitfaden. Die extrahierten Daten werden verwendet, um den Einfluss von explorativen Endpunkten im Verhältnis zu konfirmatorischen Endpunkten für die Zulassung und die frühe Nutzenbewertung separiert und zusammenhängend durch eine Gegenüberstellung von Nutzen-Schaden-Saldierung und Nutzen-Risiko-Bewertung darzustellen. Hierbei werden alle verfahrensübergreifenden eingeschlossenen Endpunkte nach der Relevanz für die beiden Bewertungsverfahren entsprechend kategorisiert. Um die Zuverlässigkeit und die Robustheit der Ergebnisse zu überprüfen, wird die Konkordanz durch den Cohens Kappa-Koeffizient (k) im Sinne einer Interrater-Reliabilität überprüft (Grouven et al. 2007) und für die Interpretation der Ergebnisse an den Richtwerten nach Altmann (1991) orientiert. Um den Einfluss von nicht eindeutig zuordnungsfähigen Endpunkten auf die Ergebnisse der Endpunktanalyse zu untersuchen, werden darüber hinaus kategoriespezifische Sensitivitätsanalysen durchgeführt.

Einundzwanzig von 41 identifizierten Bewertungen werden in der Analyse berücksichtigt. Das Zulassungsverfahren ist in einem höheren Maße konfirmatorisch als die frühe Nutzenbewertung, wobei beide Verfahren rein prozedural betrachtet nicht konfirmatorisch sind, da

jeweils explorative Endpunkte auch berücksichtigt werden. Bei der frühen Nutzenbewertung sind für die Nutzen-Schaden-Saldierung nur in 67% der Fälle primäre Endpunkte herangezogen ($k = 0,078$) worden. Explorative Mortalitätsendpunkte erreichten die höchste Übereinstimmung bei der beidseitigen Berücksichtigung von Endpunkten für die Nutzen-Risiko-Bewertung und Nutzen-Schaden-Saldierung ($k = 0,000$). Für explorative Morbiditätsendpunkte ($k = -0,600$), Lebensqualitätsendpunkte ($k = -0,600$) und Nebenwirkungsendpunkte ($k = -0,949$) wurde keine Übereinstimmung vorgefunden.

Die frühe Nutzenbewertung nach § 35a SGB V und die Zulassung streben unterschiedliche Zielsetzungen an (Kaiser et al. 2015). Dies führt zu einem abweichenden Verständnis des Stellenwertes von konfirmatorischen Endpunkten. Vor allem im Therapiegebiet der Onkologie werden primäre Endpunkte im Kontext der frühen Nutzenbewertung oftmals nicht berücksichtigt (Ruof et al. 2014a), was unter anderem mit dem methodischen Vorgehen des IQWiG zu begründen ist, laut welchem anhand indikationsunabhängiger patientenrelevanter Endpunkte eine endpunktbezogene Prioritätensetzung zur Beurteilung des Zusatznutzens vorgenommen wird (IQWiG 2015).

Eine Vergleichbarkeit der beiden Bewertungsverfahren auf Endpunktebene ist nur bedingt möglich, wenn man ein abweichendes Populationsverständnis sowie die Unterschiede im methodischen Umgang mit vorliegenden Studien unberücksichtigt lässt. Im Rahmen der frühen Nutzenbewertung kommt es teilweise dazu, dass nicht die gesamte Studienpopulation für den Nachweis des Zusatznutzens berücksichtigt wird (Köhler et al. 2015; Lange 2015). Diese eingrenzende Betrachtung von Studienpopulationen kann zu einem abweichenden Komparatorverständnis zwischen früher Nutzenbewertung und Zulassung führen. Dieser Faktor kann jedoch nicht in der Endpunktanalyse berücksichtigt werden, da im Rahmen der Nutzen-Risiko-Bewertung eine solche Eingrenzung in der Studienpopulation wie in der frühen Nutzenbewertung nicht vorgesehen ist. Im Hinblick auf die beidseitige Endpunktrelevanz stellt sich jedoch die Frage, ob und inwieweit ein Endpunkt, der sowohl in der Zulassung als auch in der frühen Nutzenbewertung eingeschlossen wird, für die Nutzen-Risiko-Bewertung mit dem einzugrenzenden Populationsverständnis der frühen Nutzenbewertung tatsächlich relevant ist. Womöglich hätte in der Zulassung ein eingrenzendes Populationsverständnis dazu geführt, dass einige für die gesamte Studienpopulation relevante Endpunkte im Zuge dessen keine Berücksichtigung erfahren hätten. Darüber hinaus führt die mangelnde Informationstransparenz des EMA Assessment Reports im Vergleich zu den vorliegenden Dokumenten aus der frühen Nutzenbewertung dazu, dass Endpunkte aus der Nutzen-Risiko-Bewertung nicht immer eindeutig der beidseitig relevanten Studie zugeordnet werden können.

Um langfristig eine breitere konfirmatorische Absicherung von Health Technology Assessment-Entscheidungen zu gewährleisten, ist eine institutionsübergreifende Annäherung erforderlich. Zu diesem Anlass haben der G-BA, das Paul-Ehrlich-Institut und das Bundesinstitut für Arzneimittel und Medizinprodukte im April 2016 eine strukturierte Zusammenarbeit vereinbart. Ziel der entwickelten Maßnahmen ist es, eine frühzeitige strukturierte Zusammenarbeit herzustellen, um eine gute Evidenz aus Zulassungsstudien zu generieren, die beiden Bewertungsverfahren gerecht wird (German Federal Ministry of Health 2016). Allerdings bleibt immer noch abzuwarten, inwieweit das gesetzte Ziel erreicht wird. Zumindest aus Sicht der pharmazeutischen Industrie scheint dies noch nicht der Fall zu sein (siehe zweite qualitative Erhebung bis Mitte 2017 in Dintsios & Schlenkrich (2018)).

Chapter (5): Failure due to formal reasons within German benefit assessment of medicinal products: the dilemma between marketing authorization and HTA

Die Arbeit untersucht und klassifiziert auf Indikationsebene (formale) Gründe für das Scheitern von Bewertungen im Rahmen der frühen Nutzenbewertung. Sie liefert hierbei eine Übersicht zu diesen Gründen bis Ende 2017 unter Verwendung unterschiedlicher Raster aus prozeduraler und biometrischer Sicht.

Mit der Einführung des AMNOG 2011 zur Kostendämpfung im Gesundheitssektor und zur qualitativen Verbesserung der Arzneimittelversorgung wird nun nach Zulassung neuer Arzneimittel (Wirksamkeit, Sicherheit, Qualität) der Zusatznutzen vom G-BA unter Berücksichtigung der vorausgegangenen IQWiG Bewertung bestimmt. Dabei unterscheiden sich beide Verfahren aufgrund ihrer spezifischen Anforderungen. Das Ergebnis der frühen Nutzenbewertung ist fundamental für die nachgelagerten Erstattungsbeitragsverhandlungen zwischen Kostenträger und pharmazeutischen Unternehmen. Bei keinen durch den G-BA attestierten Zusatznutzen dürfen nach dem fünften Sozialgesetzbuch die Jahrestherapiekosten des bewerteten Arzneimittels nicht diejenigen seines Komparators, der vom G-BA festgesetzten zweckmäßigen Vergleichstherapie, übersteigen. Somit wird die Bedeutung der Bewertungen ohne Zusatznutzen und deren Begründung evident.

Es wurden die relevanten Unterlagen (G-BA Beschlüsse und deren tragende Gründe) gesichtet und die Gründe auf Basis von vorher entwickelten Arbeitsdefinitionen klassifiziert und gruppiert. Diese Arbeitsdefinitionen sind aus einem dreischrittigen Vorgehen abgeleitet worden: zuerst wurden 100 Dossiers gesichtet und Ablehnungsgründe für einen Zusatznutzen extrahiert. Im folgenden Schritt wurden daraus thematisch zusammenhängende Gruppen gebildet und anschließend im letzten Schritt um Inhalte aus der vorhandenen Literatur, insbesondere zu indirekten Vergleichen, ergänzt. Abschließend wurden die Erfolgswahrscheinlichkeiten der Binomialverteilung (Binomial Proportions) für

definierte Indikationsgruppen sowohl auf Verfahrens- als auch auf Subgruppenebene berechnet.

Es zeigte sich, dass alle Gruppen der Kategorien von Ablehnungsgründen für einen Zusatznutzen relevant waren. Dies beinhaltete Ablehnungsgründe in Zusammenhang mit der Umsetzung der Prüfintervention und der zweckmäßigen Vergleichstherapie, der Umsetzung von indirekten Vergleichen sowie dem Studiendesign und der Operationalisierung von Endpunkten. Insgesamt zeichnen auf Subgruppenebene formale Gründe für keinen Zusatznutzen in Höhe von 90% verantwortlich, womit eine fehlende Überlegenheit der neuen Arzneimittel gegenüber ihren zweckmäßigen Vergleichstherapien nur in 10% vorkommt. Als mit Abstand wichtigster Ablehnungsgrund erwiesen sich Abweichungen des Zulassungslabels im Vergleich zu den zugrunde gelegten Zulassungsstudien (59%): In vielen Fällen deckten die Populationen der Zulassungsstudien nicht das gesamte Label ab. In der gesonderten Analyse der G-BA Bewertungen auf Subgruppenebene zeigte sich, dass für viele dieser Subgruppen keine Studienevidenz vorlag mit der Konsequenz keinen Zusatznutzen hierfür zu vergeben. Es zeigte sich auch das Gegenteil mit derselben Konsequenz: Subgruppen scheiterten in der Nutzenbewertung, weil die vorhandene Studienevidenz größere Populationen als diejenige der entsprechenden Subgruppen abdeckte. Es folgten die Ablehnungsgründe falscher Komparator (20,5%) und methodische Unzulänglichkeiten bei der Durchführung von indirekten Vergleichen (12,3%). Die Anteile des Scheiterns sind somit relativ heterogen und unterscheiden sich in Abhängigkeit von der Indikationsgruppe signifikant. Allerdings beinhaltete das Konfidenzintervall der Erfolgswahrscheinlichkeit der Binomialverteilung in den meisten Indikationsgruppen sowohl auf Verfahrens- als auch auf Subgruppenebene die 50%.

Aus den Ergebnissen der Analyse geht hervor, dass potenzielle Ablehnungsgründe in allen betrachteten Bereichen der Nutzenbewertung Hersteller-seitig untersucht werden sollten. Unter Berücksichtigung der spezifischen Anforderungen, insofern studientechnisch möglich, können die Hersteller ihre Erfolgsrate im Rahmen der Nutzenbewertung erhöhen. Dies sollte allerdings zu einem frühen Zeitpunkt der Studienplanung erfolgen. Ferner ist es wichtig, für alle aus dem intendierten Label postulierten Subgruppen Evidenz zu generieren. Die Subgruppen sollten mit der gelabelten Population übereinstimmen. Aber auch weitergehende Klarstellungen durch den G-BA innerhalb seiner Verfahrensordnung zur Auswahl der zweckmäßigen Vergleichstherapie, zu den Methoden der indirekten Vergleiche und zur Operationalisierung von patientenrelevanten Endpunkten würden das Scheitern wegen formaler Gründe reduzieren. Ferner könnte der Evidenztransfer aus der best-verfügbaren Evidenz einzelner Subgruppen auf andere ermöglicht werden, zu welchen keine direkte Evidenz aus den Zulassungsstudien vorliegt. Letztlich würde eine stärkere Angleichung der

verwendeten Anforderungen und eingesetzten Methodik von Zulassung und Nutzenbewertung das Scheitern aufgrund formaler Gründe reduzieren. Das Format hierfür könnten Beratungsgespräche vor oder während der Planung der Zulassungsstudien zu den strittigen Punkten (Studiendesign, Subgruppen, zweckmäßige Vergleichstherapie und Endpunkte) liefern (Dintsios & Schlenkrich 2018).

Während eine studienbelegte fehlende Überlegenheit gegenüber dem Komparator (zweckmäßige Vergleichstherapie) eindeutig und somit weniger angreifbar ist und keine Zweifel hinsichtlich eines im Verfahrensablauf strategisch motivierten Verhaltens des G-BA bzw. seiner Bänke und seines Auftragsinstituts (IQWiG) aufkommen lässt, sind formale Ablehnungsgründe immer mit der Frage ihrer potenziellen Vermeidung sowohl durch den Hersteller als auch durch den G-BA verbunden.

Chapter (6): Methodological problems in the method used by IQWiG within early benefit assessment of new pharmaceuticals in Germany

Die Arbeit untersucht auf biometrisch-statistischer Basis das in dem Methodenpapier des IQWiG und in ergänzender Literatur festgehaltene Vorgehen zur endpunktbezogenen Quantifizierung des Zusatznutzens im Rahmen der frühen Nutzenbewertung.

Im Nutzenbewertungsverfahren eines Arzneimittels zieht das IQWiG zur Beurteilung des Ausmaßes eines Effekts ein Schema mit insgesamt sechs, davon drei quantifizierbaren Kategorien (erheblicher, beträchtlicher und geringer Zusatznutzen) heran (AMNOG 2010; IQWiG 2013; Skipka et al. 2016). Dieses Bewertungsschema stellt in seiner einmaligen Form ein Eigenkonstrukt im internationalen Vergleich dar. Trotz der Verankerung von internationalen Standards der Evidenz-basierten Medizin (EbM) im entsprechenden Paragraphen des fünften Sozialgesetzbuches zur frühen Nutzenbewertung (§ 35a SGB V) scheint das vom IQWiG umgesetzte Vorgehen in keiner Weise vergleichbar mit dem Vorgehen anderer Health Technologie Assessment (HTA) Agenturen zu sein.

Das IQWiG geht von zwei pivotalen Zulassungsstudien aus und leitet für den Endpunkt der gepoolten Daten obere Konfidenzintervallgrenzen für den Zusatznutzen ab. Hierbei wird der in den Studien gewählte Endpunkt in drei Kategorien eingeteilt: Mortalität, Morbidität oder gesundheitsbezogene Lebensqualität, wobei die Nebenwirkungen (Sicherheitsendpunkte) der gleichen Operationalisierung der Endpunktkategorie Morbidität folgen.

Dieses Vorgehen bildet aus der oberen Konfidenzintervallgrenze und der Kategorie des Endpunktes eine Bewertungsmatrix, um den Zusatznutzen zu ermitteln. Diese Matrix ist so angelegt, dass der Zusatznutzen aus Studien mit dem Endpunkt Mortalität einfacher am höchsten, also als ‚erheblich‘, bewertet werden kann, da diesem Endpunkt ein sehr hohes Gewicht zuerkannt wird. Für Studien, die Mortalität als Endpunkt untersucht haben, sind in

der IQWiG Matrix folglich die abgeleiteten Schwellengrößen für die höchstmögliche Bewertung, im Vergleich zu den Vorgaben für andere Endpunkte, mit kleineren Effekten bzw. breiteren Konfidenzintervallen zu erreichen. Zudem hat jede statistisch signifikante Verlängerung der Überlebenszeit per definitionem mindestens einen geringen Zusatznutzen.

Auch schwerwiegende Morbidität erreicht, wenn auch viel schwieriger als Mortalität, immer noch eine höhere Bewertung, weil dies über das Vorgehen des IQWiG allerdings unter der Bedingung fast dramatischer Effekte noch möglich ist. Dagegen kann ein anderer Endpunkt als Mortalität oder schwerwiegende Morbidität, also nicht schwerwiegende Morbidität oder gesundheitsbezogene Lebensqualität in keinem Fall einen erheblichen Zusatznutzen zuerkannt bekommen, unabhängig davon wie groß die in den Studien beobachteten Effekte sind. Für die nicht schwerwiegende Morbidität gilt dies unabhängig vom klinisch erzielten Ergebnis während für die gesundheitsbezogene Lebensqualität dramatische Effekte aus klinischen Studien zumindest unbekannt sind.

Grundsätzlich muss ein Institut wie das IQWiG, das verschiedene Arzneimittel möglichst standardisiert bewerten soll und hierbei an Vorgaben des Gesetzgebers gebunden ist (Röhmel 2012; Vach 2014), ein wie auch immer geartetes Beurteilungsschema anwenden. Allerdings gibt es verschiedene Kritiken am Vorgehen des IQWiG. Unter anderem wird kritisiert, dass: (i) das derzeitige auf einen arbiträren Orientierungsanker für den Endpunkt Mortalität fußende Bewertungsschema zur endpunktbezogenen Zusatznutzenquantifizierung aufgrund der implizit höheren Gewichtung dieses Endpunktes Arzneimittel, deren Nutzen sich mit dem Endpunkt „Überleben“ messen lässt, bevorzugt. Es benachteiligt eventuell implizit solche, deren Nutzen sich nicht in den Endpunktkategorien Mortalität oder schwerwiegende Morbidität messen lässt, auch wenn das fragliche Arzneimittel in seiner Indikation einen Therapiedurchbruch darstellen sollte. In einer Extremkonstellation könnte dies dazu führen, dass ein Arzneimittel in einer Indikation, in der es derzeit gar keine Therapie gibt (Solistenstatus), keinen erheblichen Zusatznutzen zuerkannt bekommt, weil der Therapieerfolg in dieser Indikation nicht in einem vom IQWiG in seiner Bewertungsmatrix implizit höher gewichteten Endpunkt abgebildet wird bzw. werden kann (IQWiG 2013; Vach 2014). (ii) die Praxis zeigt, dass in den meisten Nutzenbewertungsverfahren nicht zwei Studien die Entscheidungsgrundlage bilden, sondern meistens nur eine Studie. Trotzdem verwendet das IQWiG hier die oberen Konfidenzintervallgrenzen, die es aus der Annahme von zwei vorliegenden Studien abgeleitet hat. Dies führt dann allerdings dazu, dass die Schwelle am oberen Konfidenzintervall leichter überschritten wird, da es bei nur einer vorliegenden pivotalen Studie zwangsläufig zu einer geringeren Präzision und hiermit einhergehend zu einem breiteren Konfidenzintervall kommt (IQWiG 2013; Vach 2014), es sei denn diese Studie ist überpowert. (iii) das IQWiG immer von einem dichotomen Endpunkt ausgeht und in

seinem Methodenpapier (IQWiG 2015) die Dichotomisierung nicht dichotomer Endpunkte vorsieht. Dies bedeutet allerdings einen Informationsverlust, falls die ursprünglichen Studiendaten nicht darauf ausgelegt waren, die Überlegenheit zwangsläufig in einem dichotomen bzw. im Nachgang dichotomisierten Endpunkt zu zeigen. Dieser Informationsverlust kann sogar dazu führen, dass ein eigentlich in einem stetigen Endpunkt überlegenes Arzneimittel keinen Zusatznutzen zuerkannt bekommen könnte, nur weil die Überlegenheit in einem nicht dichotomen Endpunkt gezeigt wurde (IQWiG 2013), sich diese Überlegenheit in dem dichotomisierten Endpunkt jedoch nicht mehr zeigen lässt.

Um die Frage zu beantworten, ob durch die unterschiedliche Gewichtung der Endpunkte bestimmte Indikationen bevorzugt werden, wurden die bis Ende 2016 verfügbaren Dossiers entsprechend ausgewertet. Aufgrund der Tatsache, dass zu wenig Dossiers auswertbar waren, konnte eine Bevorzugung von Indikationen, deren Wirksamkeit anhand des Endpunktes Überleben gemessen wird, nicht statistisch untermauert werden. Aus den vorhandenen Dossiers ging allerdings hervor, dass bisher etwa dreiviertel aller Dossiers nicht auf zwei Studien fußen, sondern nur auf einer Studie. Dass somit die Kritik, dass die Annahme des IQWiG rein theoretisch sei, diese in der Praxis allerdings selten erfüllt wird, konnte somit anhand der Dossiers belegt werden. Daraufhin wurde untersucht, welche Auswirkungen diese Diskrepanz zwischen Annahme und Realität hat. Durch Berechnung der Power, wenn nur einer Studie vorliegt, obwohl zwei erwartet werden, wurde gezeigt, dass die Power von angenommenen 80% auf Werte um die 50% einbricht. In Simulationen konnte außerdem gezeigt werden, dass das Dichotomisieren von nicht dichotomen Endpunkten eine starke Reduktion der Power bewirkt und dass somit auch der dritte Kritikpunkt an der IQWiG Methodik gerechtfertigt ist. Hinzu kommt noch die Tatsache, dass das IQWiG alle verfügbaren Endpunkte auswertet, unabhängig davon, ob die zugrundeliegende Fallzahl für diese Endpunkte adäquat ist. Für sekundäre oder explorative Studienendpunkte ist somit die statistische Power unklar und kann auch so niedrig liegen, dass der Befund auch rein zufallsbedingt sein könnte. Es verfestigt sich somit die Erkenntnis, dass die statistische Qualität der einzelnen Endpunktbewertungen im Rahmen des IQWiG Vorgehens durchaus als zweifelhaft angesehen werden kann.

Alternativ wird vorgeschlagen, für HTA relevante Endpunkte ein ähnliches Vorgehen, wie es bei der Zulassung angewandt wird, zugrunde zu legen. Dies würde im Gegensatz zum aktuellen Vorgehen des IQWiG dazu führen, dass eine höhere statistische Qualität der, im Rahmen der Bewertung, zu berücksichtigenden Endpunkte – bei insgesamt niedriger Anzahl an Endpunkten – erzielt wird. Insgesamt kann festgehalten werden, dass das Vorgehen des IQWiG, insbesondere hinsichtlich der statistischen Validität der einzelnen Endpunktbewertungen, modifiziert werden sollte.

Chapter (7): Different interpretation of additional evidence for HTA by the commissioned HTA body and the commissioning decision maker in Germany: whenever IQWiG and Federal Joint Committee disagree

In der Arbeit werden die unterschiedlichen Ergebnisse der IQWiG Bewertungen (assessments) und der G-BA Beschlüsse (appraisals) abgehandelt. Die Arbeit orientiert sich an einem akteurszentrierten Ansatz und arbeitet Gemeinsamkeiten und Dissensen des Auftraggebers (G-BA) und Auftragnehmers (IQWiG) von frühen Nutzenbewertungen nach § 35a SGB V seit Beginn des AMNOG bis Ende 2017 heraus. Es handelt sich um die erste Arbeit bis zum Zeitpunkt ihrer Publikation, die sich vertieft mit der Analyse der IQWiG Addenda befasst.

Durch die einzigartige Konstellation im deutschen Gesundheitssystem, dass das IQWiG als Auftragsinstitut des G-BA neben der Evidenzaufbereitung für den Auftraggeber auch seine eigenen Empfehlungen zu den eingereichten Dossiers der Hersteller publiziert, welche nach § 139b (4) SGB V vom G-BA im Rahmen seiner Aufgabenstellung sogar zu berücksichtigen sind, erlangen IQWiG Bewertungen einen durchaus prädisponierenden Charakter. Allerdings kann der direkte Vergleich von IQWiG Bewertungen und G-BA Beschlüssen – so wie er in der Vergangenheit praktiziert und auch publiziert wurde (Fischer & Stargardt 2014; Hörn et al. 2014; Ruof et al. 2014b), ohne eine Berücksichtigung der im Rahmen der Kommentierungen von IQWiG Bewertungen durch Stellungnahme-berechtigte Akteure (Hersteller, deren Wettbewerber, Fachgesellschaften, AKdÄ usw.) bzw. deren Anhörung beim G-BA hinzugewonnenen Erkenntnisse und potenziell zusätzlicher Evidenz (meistens durch Bereitstellung und Aufarbeitung durch den Hersteller), die eine Nachbeauftragung des IQWiG durch den G-BA induzieren (IQWiG Addenda), irreführend sein, da somit für einen sehr großen Anteil an frühen Nutzenbewertungen ein solcher Vergleich auf potenziell unterschiedlichen Evidenzkörpern und somit teils überholten Erkenntnissen fußt.

Es wurden für die Arbeit alle vorhandenen und zugänglichen Dokumente für die Bewertungen von Arzneimitteln, die eine Beauftragung des IQWiG mit einem Addendum auslösten, gesichtet und entscheidungsrelevante Inhalte (Ausmaß und Wahrscheinlichkeit des Zusatznutzens, akzeptierte Endpunkte, vorhandene Alternativen im Indikationsgebiet, Jahrestherapiekosten der zweckmäßigen Vergleichstherapie, Zielpopulationsgröße und Budget Impact, Positionierung der Fachgesellschaften und der AKdÄ etc.) in einem eigens hierfür entwickelten elektronischen Bogen überführt. Neben der deskriptiven Statistik auf Verfahrens- und Subgruppenebene zu den Unterschieden unterteilt nach Indikationsgebieten und aggregierten Kontingenztabellen wurden unterschiedliche Übereinstimmungsstatistiken angewendet: Cohen's kappa für den nominalen dichotomen Übereinstimmungsvergleich (Zusatznutzen ja/nein) von zwei Ratern (IQWiG Addenda versus G-BA Beschlüsse), Fleiss' kappa für den ordinalen Übereinstimmungsvergleich von drei Ratern (IQWiG Bewertung versus IQWiG Addendum versus G-BA Beschluss) und gewichteter Cohen's kappa für den

ordinalen Vergleich (fünf Zusatznutzenkategorien und drei Wahrscheinlichkeitsklassen) von zwei Ratern (IQWiG Addenda versus G-BA Beschlüsse). Zusätzlich wurden ordinale logistische Regressionsanalysen unter Anwendung einer schrittweise Rückwärtsauswahl für die unabhängigen Variablen und des Akaike Informationskriteriums (AIC) zur Modellgüte berechnet, um den Einfluss identifizierter potenziell entscheidungsrelevanter Einflussfaktoren auf den Unterschied zwischen den IQWiG Empfehlungen und G-BA Beschlussfassungen abzuschätzen.

Von den 90 in die Betrachtung eingeschlossenen Addenda betrafen die meisten onkologische Arzneimittel. Auf Subpopulationsebene (N = 124) übertrafen sowohl für die Zusatznutzenausprägungen als auch hinsichtlich der Wahrscheinlichkeit eines Zusatznutzens die positiven Änderungen („upgrading“) die negativen („downgrading“) im Vergleich IQWiG Addenda Empfehlungen versus G-BA Beschlussfassungen. Fleiss' ordinal kappa für die Übereinstimmung von IQWiG Bewertungen, IQWiG Addenda und G-BA Beschlussfassungen erreichte eine moderate Stärke für die Zusatznutzenausprägungen (0,474, 95%-KI: 0,408–0,540) und für die Wahrscheinlichkeit des Zusatznutzens (0,520, 95%-KI: 0,454–0,586). Die Gesamtübereinstimmung zwischen IQWiG Addenda und G-BA Beschlussfassungen auf binärer Basis fällt gering aus (Cohen's kappa 0,183; 95%-KI: 0,010–0,357) und variiert indikationsabhängig zwischen „geringer als zufallsbedingt“ (Atemwegserkrankungen) und „perfekt“ (neurologische Erkrankungen). Folgende Faktoren zeigten im ordinalen logistischen Regressionsmodell einen hohen Einfluss unterschiedlicher Richtung hinsichtlich der zu den IQWiG Empfehlungen abweichenden G-BA Beschlussfassungen: Mortalität als Endpunkt, ungedeckter medizinischer Bedarf, eine unterschiedliche Positionierung der AKdÄ und der Fachgesellschaften im Vergleich zur IQWiG Empfehlung sowie die Jahrestherapiekosten der zweckmäßigen Vergleichstherapie, wobei der G-BA beim Mortalitätsendpunkt die hohe Zusatznutzenempfehlungen des IQWiG aufgrund des algorithmischen IQWiG Vorgehens bei der endpunktbezogenen Zusatznutzenausprägung (Herpers & Dintsios 2019) in seinen Beschlussfassungen eher herabstufte. Das OR der Jahrestherapiekosten der zweckmäßigen Vergleichstherapie fällt mit 1,017 auf den ersten Blick relativ niedrig aus, allerdings bezieht sich dieser Einflussfaktor auf jeweils 1000 € der entsprechenden Kosten, womit er bei hohen Jahrestherapiekosten der zweckmäßigen Vergleichstherapie einen großen Zuwachs erfahren kann.

Für viele Onkologika und einige Antidiabetika scheint der G-BA einen pragmatischeren Ansatz bei der Bestimmung des Zusatznutzens im Sinne einer größeren Berücksichtigung des ungedeckten medizinischen Bedarfs und weiterer Kontextfaktoren im Vergleich zum IQWiG zu folgen, wie aus den Heraufstufungen ersichtlich wird. Andererseits stellt sich für neurologische Erkrankungen mit einer perfekten Übereinstimmung zwischen G-BA und

IQWiG bei fast ausschließlich negativen Nutzenbewertungen die Frage, ob der G-BA die Aufträge für die Durchführung von Addenda an das IQWiG erteilt hat, um eine weitere Rationale für seine Entscheidung bereitzustellen. Letztlich scheint der G-BA mit seinen glättenden Entscheidungen – die Divergenz zwischen IQWiG und G-BA wird insbesondere bei den Extremausprägungen des Zusatznutzens (erheblicher und kein Zusatznutzen bzw. niedrigerer Nutzen) ersichtlich – als eine Art Korrektiv zu fungieren, wobei das IQWiG bedingt durch seine eigene algorithmisch strukturierte Matrix zum endpunktbezogenen Zusatznutzen, die vom G-BA aufgrund ihrer impliziten Werturteile nicht akzeptiert wird (G-BA 2011), häufiger zu Ausreißern gelangt.

Das ordinale logistische Regressionsmodell konnte weitere entscheidungsrelevante Einflussfaktoren aufzeigen und untermauert somit das pragmatischere Vorgehen des G-BA im Vergleich zum IQWiG. Allerdings zeigt sich auch beim Einfluss der Jahrestherapiekosten der zweckmäßigen Vergleichstherapie auf den Zusatznutzen, dass der G-BA zumindest ansatzweise das Kostengeschehen antizipativ in die Nutzenbewertung einbringt, während dies eigentlich Gegenstand der Erstattungsbetragsverhandlungen zwischen Herstellern und dem GKV SV sein sollte. Dies kann auch als potenzielles Indiz einer „bad Governance“ gewertet werden, da es zu einer systematischen Zusatznutzenherabstufung im Vergleich zum IQWiG bei teuren Komparatoren führt. Andererseits spiegelt die jeweilige Preisstruktur in unterschiedlichen Indikationsgebieten auch den Innovationsstand dieser wider, womit bei hochpreisigen zweckmäßigen Vergleichstherapien davon auszugehen ist, dass in diesen Indikationsgebieten der Therapiestandard aus neueren Arzneimitteln besteht und somit ein großer Zusatznutzen vergleichend schwieriger zu erzielen ist.

Chapter (8): Comparison of post-authorisation measures from regulatory authorities with additional evidence requirements from the HTA body in Germany – are additional data requirements by the Federal Joint Committee justified?

Im Rahmen dieser Arbeit werden die Unterschiede zwischen Auflagen der europäischen Zulassungsbehörde (EMA) und zusätzlichen Evidenzanforderungen des G-BA bei befristeten Beschlüssen zum Zusatznutzen herausgearbeitet und kritisch hinterfragt. Da letztlich sowohl Zulassung als auch frühe Nutzenbewertung primär auf pivotalen Studien fußen, ist der Umgang hinsichtlich der (parallelen) Generierung zusätzlicher Evidenz für beide Verfahren fundamental und als Folge des bekannten Dilemmas zwischen Zulassung und früher Nutzenbewertung friktionsgeladen. Die Arbeit verfolgt somit einen vergleichenden Ansatz zum Vorgehen zweier für die Nutzenbewertung zentraler Akteure, der Zulassung als Institution, die den Rahmen für die Evidenzgenerierung aus pivotalen Studien setzt und des G-BA, der die frühe Nutzenbewertung durchführt.

In einem schrittweisen Vorgehen wurden alle befristeten Beschlussfassungen zu frühen Nutzenbewertungen von neuen Arzneimitteln ab Einführung des AMNOG bis Mitte 2014 identifiziert und die dazugehörenden Europäischen Öffentlichen Beurteilungsberichte (engl. EPAR: European Public Assessment Reports) bzw. deren Variationen gesichtet sowie die dort enthaltenen Datenforderungen bei bedingten Zulassungen bzw. Postzulassungsmaßnahmen (engl. PAM: Post-Authorisation Measures) extrahiert und hinsichtlich therapeutischer Indikation, Kategorien fehlender Daten, Typen der geforderten Daten sowie Klarheit der Datenforderungen analysiert und den Datenforderungen des G-BA gegenübergestellt.

Zwanzig bedingte Zulassungen zu 19 Produkten und 34 EPARs mit PAMs zu 33 Produkten konnten für den Betrachtungszeitraum identifiziert werden und daraus (i) Datenkategorien (zur Efficacy, Sicherheit, Effektivität, Pharmakologie und mit Referenz zur EMA konsistente Datenforderungen des G-BA mit den PAMs), (ii) Datentypen der geforderten Daten (RCTs, nicht-RCTs, neue Analysen existierender Daten, andere wie Simulationen oder in vitro Studien und mit Referenz zur EMA wiederum konsistente Datenforderungen des G-BA mit den PAMs) und (iii) Inhalte aus den G-BA Datenforderungen, die über die entsprechenden Anforderungen der EMA PAMs hinausgehen wie Endpunkte, Komparatoren, Langzeitdaten und Studienpopulationen sowie (iv) die Klarheit der Datenforderungen bestimmt werden. Befristete Zusatznutzenbewertungen betrafen insbesondere Onkologika (13/19 Produkten mit Befristung) und weniger PAMs (14/33 Produkten mit Postzulassungsmaßnahmen). In den öffentlich zur Verfügung stehenden Quellen konnte oft keine klare Rationale zur Befristung identifiziert werden. Sowohl die EMA als auch der G-BA forderten hauptsächlich Daten zur Efficacy und zu Sicherheitsaspekten (jeweils 44/54 und 23/35 geforderte Datenkategorien). Allerdings gingen 28/35 (80%) Anforderungen des G-BA an die nachzureichenden Daten über die entsprechenden Anforderungen der EMA hinaus. Fast die Hälfte der Datenforderungen des G-BA (9/20), dagegen keine einzige PAM, waren unklar, ohne dass der G-BA eine methodische Anleitung zur Erfüllung seiner Datenforderungen bereitstellte.

Der G-BA erachtet sich selbst als nicht an die Entscheidungen der Zulassungsbehörden gebunden im Hinblick auf Komparatoren (zweckmäßige Vergleichstherapien) und Definition von patientenrelevanten Endpunkten (Grüne 2013), zwei Bereiche, die für die Bestimmung des Zusatznutzens ausschlaggebend sind. Die offensichtlichen Unterschiede im Umgang mit Datenforderungen zwischen G-BA und EMA, wobei letztere eindeutig eine höhere Flexibilität bei der Datenakzeptanz zeigt, erklären auch weshalb etwa die Hälfte der Produkte mit befristeten Beschlussfassungen des G-BA keine PAMs Auflagen durch die EMA erhalten hatten. In diesem Zusammenhang sollte hervorgehoben werden, dass die EMA durchaus ähnliche Bedenken wie der G-BA hatte, aber zusätzliche Daten für die Beantwortung offener Fragen einsetzt. Dies hängt auch teilweise damit zusammen, dass sie in ihre Betrachtung

Studien mit unterschiedlichen Komparatoren bzw. Studien, in welchen die zu bewertende Intervention nicht exakt nach ihrem Label umgesetzt wird, einbezieht. Viel wichtiger noch, die EMA ist in ihren Bewertungen nicht ausschließlich auf RCTs angewiesen. Während der G-BA in seiner frühen Nutzenbewertung den methodischen Vorgaben des IQWiG im Großen und Ganzen folgt, die die Durchführung von RCTs wo immer möglich favorisieren, verwendet die EMA auch Evidenz aus nicht-RCTs bzw. sogar Expertenmeinungen, um Daten aus pivotalen Studien in einen Kontext zu bringen, ähnlich wie es das englische 'National Institute for Health and Clinical Excellence (NICE)' für HTA Zwecke macht (Ivandic 2014). Das G-BA Vorgehen weist auch direkte Implikationen auf die Preisgestaltung auf, da die Erstattungsbeitragsverhandlungen auf den Zusatznutzen fußen und somit bei der Studienevidenz aus bedingten Zulassungen konservativere Erstattungsbeiträge zu erwarten sind, welche wiederum aufgrund des im Vergleich mit der EMA rigiden Vorgehens des G-BA bei den Anforderungen zu den nachzureichenden Daten im Falle von befristeten Beschlüssen kaum mit einer Erhöhung im Nachgang rechnen können. Auch in dieser Konstellation lassen sich Indizien für eine „bad Governance“ dahingehend ausmachen, dass durch die strikten und damit auch schwer erfüllbaren Datenanforderungen für die erneute Bewertung nach Fristablauf der befristeten Beschlüsse, nicht mit einer Heraufstufung des bereits in der ersten Bewertung wegen der bedingten Zulassung konservativ ausfallenden Zusatznutzens zu rechnen ist. Schließlich ist das primäre Ziel des AMNOG die Kostendämpfung für neue Arzneimittel und einer der aktiv an der frühen Zusatznutzenbewertung involvierten Akteure als maßgebliche Bank des G-BA die Kostenträgerseite. Obwohl ein Ausweg aus dem Datendilemma zwischen Zulassung und früher Nutzenbewertung in Form einer strukturierten Zusammenarbeit zwischen BfArM bzw. PEI und G-BA über eine nationale gemeinsame Beratung der Hersteller gesucht wurde, hat man von einer weiteren Harmonisierung des Vorgehens abgesehen und somit das Dilemma zwischen Zulassung und HTA, auch wenn die Zielsetzung zwischen beiden Prozessen – allerdings bei (primär) gleicher zugrunde gelegter Studienevidenz – sich unterscheidet, weiterhin perpetuiert.

Jedenfalls würde eine Angleichung der Datenanforderung zwischen europäischer Zulassung (EMA) und (nationalem) HTA (G-BA) zum Zeitpunkt der Zulassung und im Verlauf des Produktlebenszyklus, insbesondere bei bedingten Zulassungen und damit einhergehenden befristeten Beschlussfassungen des G-BA den Prozess der (zusätzlichen) Datengenerierung um einiges vereinfachen und gezielter Antworten auf wichtige Fragen zu den Arzneimitteln von Interesse liefern.

Chapter (9): “Market withdrawals” of medicines in Germany after AMNOG: a comparison of HTA ratings and clinical guideline recommendations

Dieser Artikel untersucht die Marktrücknahmen von neuen Arzneimitteln während oder nach der frühen Nutzenbewertung und Erstattungsbeitragsverhandlung entweder als Außer Vertrieß (AV) Setzung oder als im Rahmen der Rahmenvereinbarung zwischen den Pharmaverbänden und den GKV SV möglichen Opt-out spätestens 14 Tage nach dem ersten Verhandlungstermin (GKV-SV et al. 2016), wobei im letzteren Fall anders als bei AV kein Erstattungsbeitrag vereinbart bzw. von der Schiedsstelle festgesetzt wird. Der Artikel stellt die Ergebnisse der frühen Nutzenbewertung der entsprechenden Arzneimittel mit den Empfehlungen innerhalb der von den medizinischen Fachgesellschaften veröffentlichten Leitlinien zu diesen Arzneimitteln bzw. deren Wirkstoffklassen gegenüber. Da neben der Ausgabeneingrenzung für hochpreisige neue Arzneimittel als Primärziel des AMNOG auch die qualitative Verbesserung der ambulanten Arzneimittelversorgung als Ziel angeführt wird, ist die Marktrücknahme von Arzneimitteln – insbesondere insofern diese einen ungedeckten medizinischen Bedarf bedienen sollten – von großer Bedeutung, weil sie potenziell dieses Ziel konterkarieren könnte. Die Arbeit ist vergleichend akteurszentriert angelegt.

Es wurden alle Opt-outs und AV von nutzenbewerteten Arzneimitteln inklusive der bis Anfang 2014 bewerteten Bestandsmarktprodukte im Zeitraum vom Beginn des AMNOG bis zum 1. Juni 2016 in unterschiedlichen Quellen identifiziert (AMNOG Berichten, GKV SV Datenbank, Arzneimittel-Verordnungsreport usw.). Arzneimittel, die im Betrachtungszeitraum nur vorübergehend nicht vertrieben wurden, wurden nicht weiter in die Analyse eingeschlossen. Arzneimittel für seltene Leiden (Orphan Drugs) wurden ebenfalls aus der weiteren Analyse ausgeschlossen, da der G-BA bei deren Bewertung keine zweckmäßige Therapie (Komparator) benennt und somit keine systematische Sichtung von entsprechenden Leitlinien durch den G-BA erfolgt. Für Arzneimittel mit mehreren Indikationen bzw. mehreren unterschiedlichen Bewertungen von Subgruppen wurde als konservativer Ansatz die höchste Zusatznutzenausprägung berücksichtigt. Die unterschiedlichen Gründe für einen nicht vergebenen Zusatznutzen wurden in folgender Reihenfolge nach abnehmenden Grad des demonstrierten klinischen Vorteils einbezogen: das betroffene Arzneimittel konnte im Rahmen der Nutzenbewertung keine ausreichende klinische Überlegenheit zeigen, es wurden keine ausreichenden Daten für die Nutzenbewertung zur Verfügung gestellt oder es wurde kein Dossier vom Hersteller eingereicht.

Da der G-BA bei der Festsetzung der zweckmäßigen Vergleichstherapie unter anderem auch eine systematische Literaturrecherche zu Leitlinien im entsprechenden Indikationsgebiet durchführt (G-BA 2012), wurden die Ergebnisse dieser Literaturrecherche zum Status der Empfehlungen zum Zeitpunkt der Nutzenbewertung herangezogen. Bei erneuter Bewertung

wurde die erste Bewertung als Referenzpunkt für die Leitlinienempfehlung zum Zeitpunkt der Nutzenbewertungen genommen. Um die Leitlinienempfehlungen zum Zeitpunkt der Analysedurchführung in die Betrachtung mit einzubeziehen, wurden alle vom G-BA identifizierten Leitlinien hinsichtlich ihrer aktuellen Version (Status 1 Juni 2016) recherchiert und gemäß dem vom G-BA verwendeten Suchalgorithmus manuell nach neuen evidenzbasierten Leitlinien gesucht. Die identifizierten Leitlinien wurden nach beiden folgenden Kriterien klassifiziert: (i) enthalten systematische Evidenzbewertung, die eine Rangfolge der Empfehlungen ermöglicht und (ii) eine grafische Darstellung des vorgeschlagenen Behandlungsalgorithmus. Für Leitlinien aus Deutschland wurde untersucht, ob diese den Kriterien einer S3 Leitlinie entsprechen und damit den höchsten methodischen Standard widerspiegeln, wie sie von der Arbeitsgemeinschaft der Medizinischen Fachgesellschaften (AWMF) definiert werden (AWMF 2016).

Unter 139 bewerteten Arzneimitteln konnten 10 Opt-outs und 12 AV Setzungen gefunden werden. Einundzwanzig von 22 nicht mehr vertriebenen Arzneimitteln (95%) wurde vom G-BA im Rahmen der frühen Nutzenbewertung kein Zusatznutzen attestiert. Über alle bewerteten Arzneimittel liegt der Anteil für solche ohne einen Zusatznutzen bei 47% im Betrachtungszeitraum. Von den 22 nicht mehr vertriebenen Arzneimitteln wurden 15 (68%) mindestens von einer Leitlinie zum Zeitpunkt ihrer Nutzenbewertung empfohlen und 18 (82%) zum Stichtag der Analyse (1. Juni 2016). Es zeigte sich eine relativ stark ausgeprägte Heterogenität unter den Leitlinien, wobei diese teilweise auch auf unterschiedliche Gesundheitssysteme zurückzuführen ist. Im Allgemeinen gilt für Leitlinien, dass – anders als HTA, das direkt auf die verfügbare Evidenz zum Zeitpunkt der Bewertung zurückgreifen kann – oft ihre Aktualität zeitlich hinterherhinkt. Ein genereller Unterschied zwischen Leitlinien und G-BA Nutzenbewertung lag auch in der Akzeptanz von verschiedenen klinischen Endpunkten bei der Beurteilung der zugrunde gelegten Studienevidenz.

Als eine spezifische Herausforderung für HTA können sogenannte beschleunigte Zulassungen angesehen werden, bei welchen die Evidenz noch teilweise als unreif gilt, aber Zulassungen aufgrund des medizinischen Bedarfs erteilt werden. Ferner scheint das Ordnungsverhalten von Ärzten auch mit einer potenziellen Marktzurücknahme negativ assoziiert zu sein, weswegen Ärzte bis zum erfolgreichen Abschluss des AMNOG Verfahrens inklusive der Erstattungsbeitragsverhandlungen zurückhaltender verordnen. Beides kann einen negativen Einfluss auf die Qualität der Arzneimittelverordnung nehmen. Insbesondere in großen Indikationen wie Diabetes Mellitus ist eine Marktrücknahme mit einem nicht zu unterschätzenden Aufwand für Umstellungen und damit einhergehend auch durch die entsprechende Leistungsanspruchnahme entstehenden Kosten verbunden; so geschehen im Fall Vildagliptin.

Eine unterschiedliche Beurteilung von Arzneimitteln zwischen HTA und Leitlinien an sich stellt kein Paradox dar, da das zugrunde gelegte methodische Instrumentarium und Bewertungsvorgehen bzw. die Interpretation der zur Verfügung stehenden Evidenz auch unter den unterschiedlichen Akteuren variieren kann. Eine bessere gemeinsame Ausrichtung der klinischen Perspektive in der Ableitung von zukünftigen Behandlungsalgorithmen und eine engere Zusammenarbeit zwischen allen involvierten Akteuren (G-BA, IQWiG, Fachgesellschaften und Patientenvertretern) könnte eine Optimierung der Arzneimittelversorgung gewährleisten, insbesondere da der Einfluss der Nutzenbewertung auf diese auch über andere Kanäle außerhalb der Preisbildung als solcher zunimmt (bspw. Arztinformationssystem mit Bereitstellung der Nutzenbewertungsergebnisse in der Praxissoftware). Hinsichtlich der Zielerreichung einer qualitativ verbesserten Arzneimittelversorgung kann auf Basis der Marktrücknahmen mit einem Anteil von 16% keine gesicherte Aussage getroffen werden. Dennoch sollten nach Auffassung der Autoren der Arbeit Arzneimittel, die in Leitlinien empfohlen werden, auch für die Patientenversorgung verfügbar sein und kein Dilemma zwischen Leitlinienempfehlung und Nutzenbewertung für die verordnende Ärzteschaft forciert werden.

Chapter (10): Predictors of negotiated prices for new drugs in Germany

Der Artikel analysiert den Einfluss von elf erklärenden Variablen auf die – auf Basis der geschlossenen Rahmenvereinbarung zwischen Pharmaverbänden und dem Spitzenverband der GKV verhandelten – Jahrestherapiekosten von neuen Arzneimitteln, die eine frühe Nutzenbewertung bis Juni 2016 durchlaufen haben und deren verhandelte Erstattungsbeträge in der Lauer-Taxe bis November 2017 einsehbar waren. Er verfolgt somit den Ansatz einer empirischen Preisfindungsanalyse im Rahmen des gesetzgeberisch mit dem AMNOG intendierten „value based pricing“ als Kostensteigerungs-dämpfungsmaßnahme und reiht sich ein in die (gesundheits-)ökonomische Methodik als einer der Pfeiler der umgesetzten multidisziplinären Methodentriangulation der Dissertation.

Die als Grundlage für die Erstattungsbetragsverhandlungen zwischen Herstellern und GKV SV dienende Rahmenvereinbarung sieht vor, dass (i) die Jahrestherapiekosten der vom G-BA bestimmten zweckmäßigen Vergleichstherapie, (ii) das Ausmaß und die Wahrscheinlichkeit des Zusatznutzens wie in der Beschlussfassung des G-BA festgehalten, (iii) die Preise der vergleichbaren Arzneimittel in der zugelassenen Indikation des nutzenbewerteten Arzneimittels sowie (iv) europäische Vergleichspreise aus den definierten Referenzländern herangezogen werden. Dadurch, dass in der Rahmenvereinbarung auf den Beschluss des G-BA in Gänze rekurriert wird, impliziert dies auch die Berücksichtigung der Populationsgrößen der betrachteten Subpopulationen in der Nutzenbewertung. Der verfolgte Ansatz

kann somit als subgruppengewichtet (Volumen) Mischpreis-kalkulatorisch bezeichnet werden. Während anfänglich in der Literatur Rabatte ohne belastbare erklärende Faktoren (kein erkennbarer Zusammenhang zwischen Zusatznutzenausprägung und Rabatt) analysiert wurden (Theidel & von der Schulenburg 2016), untersuchen aktuellere Studien (Hammerschmidt 2017; Lauenroth & Stargardt 2017; Radic et al. 2018) den monetären Aufschlag auf die zweckmäßige Vergleichstherapie als abhängige Variable und zeigten eine monotone Beziehung in Zusammenhang mit dem Ausmaß des Zusatznutzens. Dennoch bleibt ihre Aussage aufgrund des geringen Stichprobenumfangs und einer relativ großen Varianz der anteiligen Aufschläge limitiert. Um diesen Limitationen zu entgehen, wurden im Artikel die Jahrestherapiekosten auf Basis des verhandelten Erstattungsbetrags als abhängige Variable in den – aufgrund der ursprünglichen Verteilung einiger kontinuierlichen Variablen (untersucht mit dem Shapiro–Wilk Test) zur Approximation der Normalverteilung log-transformierten – OLS Regressionsmodellen mit den Jahrestherapiekosten der zweckmäßigen Vergleichstherapie als unabhängige Variable bestimmt und folgten damit dem Ansatz einer aktuellen Analyse zu den Preisen nicht-onkologischer Orphan Drugs in Deutschland (Schlander et al. 2018). Zur Untersuchung potenziell vorliegender Multikollinearität zwischen den kontinuierlichen unabhängigen Variablen wurde eine Korrelationsmatrix konstruiert. Dabei wurde die Unabhängigkeit der Variablen mit dem Pearson χ^2 Test überprüft. Zusätzlich wurden Varianzinflationsfaktoren für die geschätzten Regressionskoeffizienten berechnet. Um eine potenzielle Heteroskedastizität zu berücksichtigen, wurden die Regressionen mit robusten Standardfehlern geschätzt.

In die Analyse wurden für die betrachtete Zeitspanne 106 nicht-Orphans eingeschlossen. Es zeigte sich in den berechneten vier Regressionsmodellen eine signifikante und positive Assoziation der log-transformierten verhandelten Jahrestherapiekosten neuer Arzneimittel mit den log-transformierten Jahrestherapiekosten ihrer Komparatoren (zweckmäßigen Vergleichstherapien), dem Ausmaß des Zusatznutzens und der log-transformierten Zielpopulationsgröße. Die Verbesserung um eine Zusatznutzenkategorie führte zu einer Erhöhung der Jahrestherapiekosten um 23%, die Verdoppelung der Zielpopulationsgröße hingegen zu einer Verminderung der Jahrestherapiekosten um 13%. Die Erhöhung der Wahrscheinlichkeit des Zusatznutzens hatte einen gleichgerichteten aber etwas schwächer ausgeprägten Einfluss auf die Jahrestherapiekosten im Vergleich zum Ausmaß des Zusatznutzens. Bei Betrachtung der einzelnen Endpunkte anstelle des aggregierten Zusatznutzens in den Regressionsmodellen zeigte sich unerwartet, dass die Nebenwirkungen den größten positiven Effekt auf den Erstattungsbetrag ausübten. Die verschiedenen durchgeführten Subgruppen- und Sensitivitätsanalysen bestätigten die Robustheit der Ergebnisse. Das adjustierte R^2 lag für alle Modelle bei über 80%. Bei ausschließlicher

Berücksichtigung der 24 Fälle mit einem von der Schiedsstelle festgesetzten Erstattungsbetrag erreichte das adjustierte R^2 mit 91% höhere Werte im Vergleich zu den verhandelten Erstattungsbeträgen. Dies lässt den Rückschluss zu, dass die Schiedsstelle in ihrem Vorgehen die Vorgaben der Rahmenvereinbarung in einem größeren Ausmaß berücksichtigt als der GKV SV.

In der Analyse konnten die in der Rahmenvereinbarung genannten Verhandlungskriterien als Haupttreiber der Erstattungsbetragsfindung zwischen Herstellern und GKV SV empirisch bestätigt werden. Widererwartend zeigte auf einzelner Endpunktebene die Zunahme der Nebenwirkungen einen signifikant positiven Effekt auf die Preisgestaltung ungeachtet der Kontrolle für positive Effekte anderer klinischer Endpunkte, insbesondere der Reduktion der Mortalität. Dies bedeutet, dass der Effekt der Nebenwirkungen die positiven Effekte der anderen Endpunkte übertrifft. Dennoch wird eine Zunahme von Nebenwirkungen in der klinischen Entwicklung immer durch die Zunahme anderer positiver Effekte ausgeglichen. Somit ist es durchaus möglich, einen reduzierteren Nutzen durch die Nebenwirkungen zu erhalten, aber dennoch auf aggregierter Ebene über alle Endpunkte einen Zusatznutzen zu erzielen. Wir erachten aus diesem Grunde die statistischen Ergebnisse als in sich konsistent und gehen nicht von einem statistischen Artefakt aus. Das Phänomen könnte entweder durch eine Korrelation mit weiteren, im Regressionsmodell nicht einbezogenen, Variablen erklärt werden oder es handelt sich um einen kontra-intuitiven (irrationalen), aber realen Zusammenhang. Letzteres scheint sogar aufgrund der hohen adjustierten R^2 Werte der Regressionsmodelle wahrscheinlicher. Limitationen der Analyse betreffen die Annahmen zu einigen Daten sowie die Tatsache, dass europäische Vergleichspreise und vergleichbare Arzneimittel nicht in die Betrachtung einbezogen wurden. Ferner kann eventuell von weiteren unbekanntem Einflussfaktoren ausgegangen werden, die potenziell zu Einsparungen führen, aber offiziell nicht benannt werden und unberücksichtigt bleiben.

Chapter (11): Determinants of Orphan Drug Prices in Germany

Die Arbeit befasst sich mit den Einflussfaktoren auf den Erstattungsbetrag von Arzneimitteln für seltene Leiden (Orphan Drugs), für welche hinsichtlich der frühen Nutzenbewertung besondere Regeln gelten und somit nicht von den gleichen Grundvoraussetzungen wie bei Arzneimitteln ausgegangen werden kann, die eine konventionelle frühe Nutzenbewertung durchlaufen. Es handelt sich um die zweite Arbeit, die einen empirischen Ansatz zur Preisfindungsanalyse im Rahmen des intendierten „value based pricing“ verfolgt und sich in die (gesundheits-)ökonomische Methodik der umgesetzten Methodentriangulation der Dissertation einreicht.

Für Orphan Drugs gemäß Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999, gilt der medizinische Zusatznutzen durch die Zulassung als belegt nach § 35a SGB V, womit auch keine zweckmäßige Vergleichstherapie vom G-BA bestimmt wird und somit ein essentieller Orientierungspunkt aus der Rahmenvereinbarung (GKV-SV et al. 2012) für die Erstattungsbetragsverhandlungen der Hersteller mit dem GKV SV entfällt. Daraus ergibt sich zwangsläufig die Frage, welche Einflussfaktoren und wie diese spezifisch auf die Erstattungsbetragsverhandlungen von Orphan Drugs wirken und ob es noch weitere Einflussfaktoren außer den restlichen in der Rahmenvereinbarung genannten gibt.

Zwölf Hypothesen zu potenziellen Einflussfaktoren auf den Erstattungsbetrag von Orphan Drugs wurden aus der vorhandenen Literatur und der Rahmenvereinbarung zwischen den Pharmavereinigungen und dem GKV SV abgeleitet: i) die Hauptniederlassung des Herstellers außerhalb Deutschlands hat einen negativen Einfluss auf den Erstattungsbetrag, ii) die Verhandlungserfahrung der Hersteller beeinflusst positiv das Verhandlungsergebnis, iii) der Therapiebereich beeinflusst die Erstattungsbeträge in unbekannter Richtung, iv) die Krankheitsschwere übt einen positiven Einfluss auf den Erstattungsbetrag aus, v) pädiatrische Zulassungen zeichnen sich durch einen positiven Einfluss auf den Erstattungsbetrag aus, vi) die Zielpopulationsgröße übt einen negativen Einfluss auf den Erstattungsbetrag aus, vii) ein vom G-BA quantifizierter Zusatznutzen beeinflusst den Erstattungsbetrag positiv, viii) das Vorhandensein vergleichbarer Arzneimittel und deren Preis haben einen positiven Einfluss auf den Erstattungsbetrag, ix) Europäische Preise des Orphan Drugs wirken sich positiv auf den Erstattungsbetrag aus, x) die Antizipierung einer Wiederverhandlung übt einen Einfluss auf den Erstattungsbetrag in unbekannter Richtung aus, xi) das aktuelle und antizipierte Verordnungsvolumen haben einen negativen Einfluss auf den Erstattungsbetrag und xii) das budgetäre Geschehen wirkt sich negativ auf den Erstattungsbetrag aus. Univariate Analysen wurden angewendet, um statistisch signifikante Korrelationen zwischen den Jahrestherapiekosten des verhandelten Orphan Drugs auf Basis des Erstattungsbetrags und den hypothetischen Einflussfaktoren zu finden. Bivariate Analysen wurden zur potenziellen Confounderbestimmung durchgeführt. Zusätzlich wurde aufgrund der Verteilung einiger Variablen ein log-log-transformiertes OLS Regressionsmodell mit Rückwärtsauswahl der unabhängigen Variablen und dem Akaike Informationskriterium für die Modellgüteeinschätzung zur Bestimmung potenziell signifikanter erklärender Variablen gerechnet. Abschließend bewerteten Sensitivitätsanalysen die Robustheit der Ergebnisse.

Fünfunddreißig Orphan Drugs im Betrachtungszeitraum vom Anfang des AMNOG bis zum 31. Januar 2018 mit jeweils durchschnittlichen Jahrestherapiekosten in einer Spannweite von 7.321 € und 2.363.670 € pro Patienten wurden in die Analyse einbezogen. Die univariaten

Analysen und anschließenden Sensitivitätsanalysen validierten fünf der formulierten 12 Hypothesen. Es konnte eine starke statistisch signifikante Assoziation zwischen Orphan Drug Erstattungsbetrag und i) dem Therapiebereich, in dem das Orphan Drug zur Anwendung kommt, ii) eine pädiatrisch zugelassene Indikation, iii) der Zielpopulationsgröße, iv) den Kosten vergleichbarer Orphan Drugs und v) den europäischen Preisen gefunden werden.

Die nachfolgenden Analysen zur Korrelation des therapeutischen Bereichs und anderer Einflussfaktoren in Zusammenhang mit dem verhandelten Erstattungsbetrag zeigten eine Multikollinearität für den Therapiebereich. Die behandelte seltene Erkrankung gilt als Trigger für diesen Effekt. Weil der Therapiebereich ein Proxy für die zugrundeliegende seltene Erkrankung ist, wird geschlossen, dass der Therapiebereich eine Mediatorvariable darstellt, die mit allen mit der Krankheits zusammenhängenden Variablen als auch mit dem Erstattungsbetrag assoziiert ist.

Das log-log-transformierte OLS-Regressionsmodell identifizierte die europäischen Preise als die Variable unter den unabhängigen Variablen mit der stärksten Assoziation mit den Erstattungsbeträgen für Orphan Drugs. Wenn sich die europäischen Preise um 1% verändern, wird eine Veränderung des verhandelten Erstattungsbetrags von Orphan Drugs um 0,97% erwartet. Allerdings könnte diese starke Assoziation auch durch zwei Limitationen der Studie erklärt werden: Erstens wurden europäische Listenpreise verwendet, da vertrauliche Rabatte nicht ohne Weiteres zugänglich sind und zweitens waren alle Variablen mit einer signifikanten Assoziation mit dem Erstattungsbetrag mit den europäischen Preisen verbunden, und somit kann man zurückverfolgen, dass die starke Korrelation zwischen europäischen Preisen und Erstattungsbetrag darauf zurückzuführen ist, dass andere europäische Länder wiederum ähnliche Kriterien für die Preisbildung von Orphan Drugs ansetzen.

Unabhängig von ihren Limitationen liefert die Arbeit einen entscheidenden Beitrag hinsichtlich der Transparenzmachung des Preisgeschehens bei Orphan Drugs in Deutschland. Letztlich zeigt die Arbeit auf, dass weder die vom unabhängigen G-BA Vorsitzenden, Prof. Hecken, artikulierte Gefahr einer „Orphanisierung“ aufgrund der Zusatznutzenregelung im AMNOG mit nur einem Drittel unter den neu zugelassenen Arzneimitteln besteht, noch mit einem Anteil in Höhe von nur 4% an den Arzneimittelausgaben der GKV Orphan Drugs einen gravierenden Ausgabenfaktor darstellen. Auch zeigt sich, dass die Preise von Orphan Drugs in den Ländern mit frühem Zugang zum Arzneimittelmarkt ähnlich ausfallen und somit die Aussage, dass Deutschland ein Hochpreisland für Orphan Drugs sei, relativiert wird.

Die Arbeit konnte zeigen, dass die verhandelten Erstattungsbeträge für Orphan Drugs ein multivariates Phänomen darstellen. Allerdings sind aufgrund unterschiedlicher Interdependenzen die Ergebnisse mit Vorsicht zu interpretieren.

Chapter (12): The impact of additive or substitutive clinical study design on the negotiated reimbursement for oncology pharmaceuticals after early benefit assessment in Germany

Die Arbeit untersucht empirisch den Einfluss des Studiendesigns auf den Erstattungsbetrag von Onkologika. Wie in den vorausgegangenen zwei Arbeiten handelt es sich auch hier um eine Arbeit, die einen empirischen Ansatz zur Preisfindungsanalyse im Rahmen des intendierten „value based pricing“ verfolgt und sich in die (gesundheits-)ökonomische Methodik der umgesetzten Methodentriangulation der Dissertation einreicht.

Zum Nachweis eines Zusatznutzens reicht der Hersteller bei Markteintritt sein Dossier mit Daten der klinischen Studien beim G-BA ein. Das Design dieser klinischen Studien kann dabei variieren: Es gibt additive Studiendesigns, bei denen im Interventionsarm der Studie zusätzlich zur Baseline Therapie aus dem Kontrollarm eine neue Therapie angewendet wird (bspw. Intervention + Best Supportive Care vs. Placebo + Best Supportive Care) und es kommen aber genauso substituierende Studiendesigns zur Anwendung, bei welchen im Interventionsarm die Therapie des Kontrollarms durch die zu prüfende Therapie ersetzt wird. Im Rahmen der Analyse wird zwischen unterschiedlichen Studiendesigns differenziert und geprüft, ob das Design der klinischen Studien als solches einen Einfluss auf den verhandelten Erstattungsbetrag nimmt. Die Analyse wird explizit im Therapiegebiet der Onkologie durchgeführt, weil hier insbesondere viele unterschiedliche Studiendesigns in klinischen Studien verwendet werden. In der Onkologie werden neue Therapien häufig erst in späten Therapielinien eingesetzt und aus diesem Grund oft ergänzend zu bestehenden Therapien angewendet, sodass additive Studiendesigns zur klinischen Prüfung gewählt werden. Es werden allerdings ebenso klassisch substituierende Studiendesigns zum Nachweis eines Zusatznutzens verwendet. Bei Orphan Drugs in der Onkologie kann zwischen solchen mit einem solistischen Status unterschieden werden und solchen, die trotz einer Anwendung in einem seltenen Leiden in der Onkologie im Wettbewerb mit anderen Orphan Drugs stehen. Auch bei Orphan Drugs kann zwischen additiven und substituierenden Studiendesigns unterschieden werden. Da das Studiendesign das Therapieregime vorgibt, hat es eine direkte Auswirkung auch auf den Budget Impact eines Arzneimittels.

Die zu überprüfende Haupthypothese lautet: Das Design der im Rahmen der frühen Nutzenbewertung eingegangenen Studien hat einen Einfluss auf den Erstattungsbetrag. Daraus lassen sich folgende untergeordnete Hypothesen zur weiteren Überprüfung ableiten: i) bei einem additiven Studiendesign und einem nicht belegten Zusatznutzen liegen die Jahrestherapiekosten der Therapie nicht oberhalb derjenigen der zweckmäßigen Vergleichstherapie (es ist zu erwarten, dass ein Preis von 0 € für das neue Arzneimittel entsteht, da sich die Kosten für die Baseline Therapie rauskürzen); ii) bei einem substituierenden Studiendesign und einem nicht belegten Zusatznutzen liegen die Jahrestherapiekosten nicht oberhalb

derjenigen der zweckmäßigen Vergleichstherapie; iii) bei einem additiven Studiendesign und einem nachgewiesenen Zusatznutzen erfolgt ein Zuschlag auf die Jahrestherapiekosten (der kleiner als bei einem substituierenden Studiendesign ist, da der größere Budget Impact von Kombinationstherapien mit additivem Studiendesign durch einen geringeren Zuschlag ausgeglichen wird); iv) bei einem substituierenden Studiendesign und einem nachgewiesenen Zusatznutzen erfolgt ein Zuschlag auf die Jahrestherapiekosten der zweckmäßigen Vergleichstherapie (der wiederum im Umkehrschluss zu Hypothese 3 höher ist, als bei einem additiven Studiendesign); v) bei Orphan Drugs können in Abhängigkeit vom Studiendesign Konstellationen eintreten, die analog zu den beschriebenen Hypothesen 3 und 4 sind, allerdings unter der Prämisse, dass es keine zweckmäßige Vergleichstherapie als solche gibt und vergleichbare Arzneimittel – insofern in der Indikation vorhanden – als potenzieller Preisanker dienen, wobei für die Konstellation eines Solisten das Ergebnis der Erstattungs-betragsverhandlungen die direkte Zahlungsbereitschaft des GKV SV widerspiegelt.

Für alle 55 im Betrachtungszeitraum von der Einführung des AMNOG bis September 2016 abgeschlossenen onkologischen Verfahren wurde anhand der von den Herstellern beim G-BA eingereichten Dossiers bestimmt, welches Studiendesign in den entsprechenden Studien verwendet wurde (47 von 55 Verfahren wurden nach Ausschluss wegen unterschiedlicher Gründe in die Analyse einbezogen). Der Einfluss auf den Erstattungsbetrag wurde Subgruppen-spezifisch mittels additiver und multiplikativer Aufschläge auf die (gewichteten) Jahrestherapiekosten der zweckmäßigen Vergleichstherapien (mit Ausnahme von Orphan Drugs) untersucht. Weitere Einflussfaktoren auf den Erstattungsbetrag wurden nebst dem Budget Impact in der abschließenden Betrachtung mithilfe univariater Analysen und aufgrund der Verteilung einiger stetiger Variablen (Shapiro-Wilk W Statistik) multivariater log-transformierter OLS-Regressionsanalysen untersucht.

Hypothese 1 konnte aufgrund der geringen Fallzahl nicht abschließend überprüft werden, wird aber durch die geltenden rechtlichen Regelungen stark unterstützt. Hypothesen 2 und 3 können grundsätzlich angenommen werden, wobei Hypothese 3 noch der Hypothese 4 gegenübergestellt werden muss. Die berechneten additiven und multiplikativen Aufschläge stützen Hypothese 3 und 4, dass die Aufschläge bei substituierenden Studiendesigns höher sind, als bei additiven Designs, allerdings nur für den Fall, dass es sich um additive Studiendesigns mit einer kurativen zweckmäßigen Vergleichstherapie handelt. Für additive Studiendesigns mit einer palliativen zweckmäßigen Vergleichstherapie fallen die Aufschläge aufgrund der niedrigeren Jahrestherapiekosten palliativer Interventionen deutlich höher als bei substituierenden Studiendesigns aus und führen somit zunächst zu einem kontraintuitiven Ergebnis, welches sich jedoch durch die palliative zweckmäßige Vergleichstherapie erklären lässt. Unter der Annahme, dass nur additive Studiendesigns mit kurativer Zielsetzung

berücksichtigt werden, kann Hypothese 4 dann auch angenommen werden. Der Mittelwert (50.477,68 €) und der Median (49.841,24 €) für additive Aufschläge bei substitutiven Studiendesigns lagen höher als bei additiven Studiendesigns (48.750,00 € und 42.820,44 €), wenn keine palliative zweckmäßige Vergleichstherapie zur Anwendung kam. Der durchschnittliche multiplikative Aufschlag für substitutive Studiendesigns war 15,07 und der für additive Studiendesigns 2,29. Hypothese 5 muss erneut bei Vorliegen weiterer Fälle überprüft werden, da die Ergebnisse nicht eindeutig waren (für Orphan Drugs in Indikationen mit vergleichbaren Arzneimitteln als Preisanker gab es sowohl Auf- als auch Abschläge).

Die Kriterien aus der Rahmenvereinbarung (GKV-SV et al. 2016) zwischen den Pharma-Verbänden und dem GKV SV für die Erstattungsbetragsverhandlungen (Ausmaß des Zusatznutzens, vergleichbare Arzneimittel, europäische Preise) sowie weitere mögliche Einflussfaktoren auf den additiven (bzw. multiplikativen) Aufschlag wie die Jahrestherapiekosten der zweckmäßigen Vergleichstherapie und die Zielpopulationsgröße wurden nebst dem Studiendesign in den Regressionen als unabhängige Variablen modelliert. Trotz der niedrigen Anzahl an Fällen (es wurden nur solche mit vollständigem Datensatz einbezogen und Fälle mit mehreren Indikationen sowie zwei Ausreißer ausgeschlossen) wurden in den univariaten Analysen zu den additiven Aufschlägen europäische Preise und Zielpopulationsgröße statistisch signifikant (bei den multiplikativen Aufschlägen vergleichbare Arzneimittel). Das adjustierte R^2 in den log-transformierten OLS Regressionsmodellen mit additivem Aufschlag als abhängige Variable erreichte 0,708 bei Einschluss aller erklärenden Variablen und Berücksichtigung der Interaktion zwischen Zielpopulationsgröße und Jahrestherapiekosten der zweckmäßigen Vergleichstherapie. Die Einbeziehung des Studiendesigns als unabhängige Variable in das Modell ging mit einem relevanten Zuwachs des adjustierten R^2 einher und zeigte, dass das Studiendesign selber neben der Zielpopulationsgröße und den Jahrestherapiekosten der Vergleichstherapie einen relevanten Anteil an Varianz erklären konnte und somit angenommen werden kann, dass es nicht implizit durch den potenziellen Budget Impact abgebildet wird.

Das gewählte Vorgehen der Aufschlagsberechnung ist komplexer, da die Jahrestherapiekosten von Intervention und Kontrolle berechnet werden müssen, was insbesondere bei onkologischen Verfahren aufgrund patientenindividueller Therapieschemata, die bspw. an die Körperoberfläche angepasst sind, aufwendig ist. Für die Berechnung der Jahrestherapiekosten sind Daten zu Versorgungsanteilen und Kostendaten zur zweckmäßigen Vergleichstherapie nötig, die allerdings auch Annahmen notwendig machen. Eine Rabattbetrachtung, bei der Packungspreise vor und nach Verhandlung miteinander verglichen werden, lässt sich sicherlich mit geringerem Aufwand durchführen (Theidel & von der Schulenburg 2016), allerdings wird dabei nicht berücksichtigt, dass der Hersteller antizipativ Rabatte einpreist

und möglicherweise den Markteintrittspreis entsprechend anpasst. Des Weiteren und abweichend zu anderen Analysen (Hammerschmidt 2017; Radic et al. 2018), werden die Aufschläge subgruppenspezifisch berechnet. Bei vielen Verfahren hat eine Subgruppe keinen belegten Zusatznutzen, während eine andere Subgruppe einen attestierten Zusatznutzen aufweist, sodass eine Differenzierung zwischen den Subgruppen und die Berücksichtigung ihrer Anteile an der gesamten Zielpopulation des Arzneimittels wichtig zur Vermeidung von Verzerrungen wird.

Chapter (13): Arbitration Board Setting Reimbursement Amounts for Pharmaceutical Innovations in Germany When Price Negotiations between Payers and Manufacturers Fail: An Empirical Analysis of 5 Years' Experience

Der Artikel untersucht als erste internationale Publikation die Funktion der Schiedsstelle nach § 130b (5) SGB V, die mit dem AMNOG ins Leben gerufen wurde, falls Hersteller und GKV SV sich im Rahmen der Erstattungsbetragsverhandlungen nicht einigen können. Im Fall der Nichteinigung muss die Schiedsstelle innerhalb von drei Monaten entscheiden, wobei ihr Schiedsspruch rückwirkend ab dem dreizehnten Monat nach Initiierung der frühen Nutzenbewertung gilt. Gegen einen Schiedsstellenentscheid kann eine gerichtliche Klage erhoben werden, womit die Entscheidung der Schiedsstelle rechtsmittelfähig ist, allerdings hat eine Klage keine aufschiebende Wirkung, sodass der festgesetzte Erstattungsbetrag weiterhin gilt. Die Schiedsstelle nimmt somit eine zentrale Rolle im AMNOG-Prozess ein. Der Artikel verfolgt einen akteurszentrierten Ansatz, in welchem die Positionen der Hersteller und des GKV SV vor der Schiedsstelle fallbezogen analysiert werden und die finalen Entscheidungen der Schiedsstelle diesen gegenübergestellt wird. Ziel war es, eine Übersicht über alle sechzehn Schiedsverfahren im Betrachtungszeitraum von Beginn des AMNOG bis zum Januar 2016 zu schaffen, den impliziten Grund des Anrufens der Schiedsstelle zu ermitteln und Entscheidungsgrundlagen der Schiedsstelle zu beleuchten sowie gegebenenfalls Heuristiken für ihre Rationale abzuleiten.

Für die Schiedsstelle, die im Rahmen des AMNOG geschaffen wurde, lassen sich gesetzliche Vorbilder in anderen Schiedsämtern des fünften Sozialgesetzbuches finden. Im § 130b SGB V, welcher die Vereinbarungen zwischen dem GKV SV und Herstellern über Erstattungsbeträge für Arzneimittel regelt, wird auf das Schiedsamt für die vertragsärztliche Versorgung (§ 89 SGB V) und die Schiedsstelle für den Rahmenvertrag der Arzneimittelversorgung (§ 129 Abs. 7 bis 10 SGB V) verwiesen. Die Schiedsämter für die vertragsärztliche Versorgung nach § 89 SGB V sind ebenfalls in das Recht der GKV eingebunden. Der Ablauf des Schiedsverfahrens der AMNOG-Schiedsstelle und des Schiedsamtes für die vertragsärztliche Versorgung, in der die AMNOG Schiedsstelle ihre Grundlage findet, ist sehr ähnlich. Ein Unterschied zwischen der im Jahr 2011 durch das AMNOG neu ins Leben gerufenen

Schiedsstelle und des Schiedsamtes für die vertragsärztliche Versorgung stellt der Schlichtungsgrund für die Schiedsstelle dar. Die AMNOG Schiedsstelle versucht zwischen Interessen einer Körperschaft des öffentlichen Rechts, dem GKV SV und einem privatwirtschaftlichen Unternehmen zu schlichten, was im Sozialrecht ein Novum darstellt. Dagegen schlichtet das Schiedsamt für die vertragsärztliche Versorgung Streitigkeiten zwischen den Verbänden der Krankenkassen und den Verbänden der Leistungserbringer.

Vorgaben für Kriterien, die für die Schiedsstelle bindend sind, lassen sich sowohl im § 130b SGB V, als auch in der Rahmenvereinbarung nach § 130b Abs. 9 SGB V (GKV-SV et al. 2012) finden. Das Gesetz legt fest, dass die Schiedsstelle für Arzneimittel, für die vom G-BA ein Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie festgestellt wurde, bei der Ermittlung des Erstattungsbetrags neben der Monetarisierung des Zusatznutzens die Jahrestherapiekosten vergleichbarer Arzneimittel und die Höhe des tatsächlichen Abgabepreises in anderen europäischen Ländern zu berücksichtigen hat. Hat der G-BA keinen Zusatznutzen feststellen können, so darf der Erstattungsbetrag nicht zu höheren Jahrestherapiekosten führen als die der zweckmäßigen Vergleichstherapien. In dem Fall, dass es mehrere zweckmäßige Vergleichstherapien gibt, darf der Erstattungsbetrag nicht zu höheren Jahrestherapiekosten führen als die wirtschaftlichste Alternative. Erst im Zuge des 3. AMG-Änderungsgesetzes von 2013 wurde im SGB V mit aufgenommen, dass es mehrere zweckmäßige Vergleichstherapien geben kann. Die Entscheidungen der Schiedsstelle sind Ermessensentscheidungen. In § 130b Abs. 4 Satz 2 SGB V ist geregelt, dass die Schiedsstelle unter freier Würdigung aller Umstände des Einzelfalls entscheidet und dabei die Besonderheiten des jeweiligen Therapiegebietes berücksichtigt. Nichtsdestotrotz muss sich das Ermessen im Rahmen der gesetzlichen Regelungen bewegen und besonders den G-BA Beschluss als bindend betrachten. Die Schiedsstelle muss in ihrer Entscheidungsfindung erkennen lassen, auf welchen Argumenten sich das Ermessen stützt. Sie ist an den G-BA Beschluss gebunden, kann aber das Ausmaß des Zusatznutzens hinsichtlich seiner Monetarisierung dennoch anders einschätzen als die Vertragspartner. Sie betont mehrfach, dass sie keinem Algorithmus in ihrer Entscheidungsfindung folgt (und laut § 130b (5) Satz 2 SGB V auch nicht folgen soll), sondern vielmehr die Anteile für die Kriterien je nach Fall unterschiedlich gewichtet.

In einem mehrschrittigen Vorgehen wurden zunächst neben den jeweiligen Schiedssprüchen der Schiedsstelle (für die interessierte Öffentlichkeit einsehbar in Berlin bei der Schiedsstelle) Nutzenbewertungen, G-BA Beschlüsse und deren Tragende Gründe sowie die zusammenfassenden Dokumentationen (erhältlich über die Internetseiten des G-BA) der entsprechenden Fälle gesichtet und alle relevanten Daten zum Inhalt des Schiedsspruchs extrahiert (u.a. Zusatznutzenausprägung, Streitpunkte zwischen den Verhandlungsparteien

und Erstattungsbetragsforderungen). Anschließend folgte eine Häufigkeitsanalyse der strittigen Punkte sowie ein Vergleich der Erstattungsbetragsforderungen der Verhandlungsparteien und der Entscheidungen der Schiedsstelle hierzu. In den Schiedssprüchen ist nicht nur von Erstattungsbeträgen, Rabatten und Abgabepreise die Rede, es wird auch zwischen unterschiedlichen Preisbasen gewechselt. Es sind Angaben zu Jahrestherapiekosten, Tagestherapiekosten (DDD) und Kosten je Pharmazentralnummer (PZN), das heißt auf Basis der Packungen zu finden. Damit eine einheitliche Analyse gewährleistet werden kann, werden fast alle Kostenangaben in eine Einheit transformiert. Die kleinste Einheit, die gleichzeitig eine gute Vergleichbarkeit ermöglicht, sind die Tagestherapiekosten. Anhand von Kasuistiken werden die Verfahren mit und ohne Zusatznutzen separat voneinander betrachtet, da unterschiedliche rechtliche Regelungen zum Erstattungsbetrag hierfür gelten.

Im Betrachtungszeitraum kam es zu 16 Schiedssprüchen (5 mit Zusatznutzen) bei 106 abgeschlossenen (15%) und noch 20 laufenden Erstattungsbetragsverhandlungen sowie 6 laufenden Schiedsverfahren. In 4 Fällen wurden die Arzneimittel direkt in eine Festbetragsgruppe überführt, während 20 Arzneimittel nicht mehr vertrieben wurden. Als Hauptstreitpunkte bemessen an ihrer Häufigkeit erwiesen sich neben dem Erstattungsbetrag: Verordnungsmengen, Vertragslaufzeit, Ablösung der Herstellerrabatte sowie Zeitpunkt der Geltung des Erstattungsbetrags. Es ist sehr schwierig die Vertragslaufzeit als singulären Faktor zu betrachten, da sie immer im direkten Zusammenhang mit dem Erstattungsbetrag steht. Hersteller haben ein großes Interesse daran, den Rabatt auf den von ihnen festgelegten Abgabepreis möglichst gering zu halten und gleichzeitig aus Planungssicherheitsgründen eine lange Vertragslaufzeit zu vereinbaren. Auch die Verordnungsmengen hängen direkt mit dem Erstattungsbetrag zusammen und müssen aus Herstellerperspektive umso höher angesetzt werden, je höher der Rabatt auf den Abgabepreis ist, da das Ordnungsverhalten der Ärzte wiederum häufig stark mit dem Preis zusammenhängt, sodass bei therapeutisch vergleichbaren Arzneimitteln das preisgünstigste Arzneimittel eher verschrieben wird. Eine Ablösung des Herstellerabschlages (Zwangsrabatt nach § 130a (1a) SGB V) bedeutet, dass dieser Rabatt mit im Erstattungsrabatt eingepreist wird. Für den Hersteller hat die Überlegung einer Ablösung des Abschlages strategische Gründe, da er sich gegen mögliche Erhöhungen schützen kann, wobei er allerdings bei einer Reduzierung nicht profitiert. Während das Bild bei den Verordnungsmengen gemischt ausfällt, entschied sich die Schiedsstelle bei der Vertragslaufzeit eher im Sinne des GKV SV und bei der Ablösung der Herstellerrabatte ausschließlich zugunsten der Hersteller. In fast allen Fällen handelte es sich bei der zweckmäßigen Vergleichstherapie um generische Produkte. In den Fällen ohne Zusatznutzen folgte die Schiedsstelle strikt den gesetzlichen Regelungen zum Erstattungsbetrag. Im Durchschnitt lagen die Erstattungsbeträge 38,4% unter dem Mittel der geforderten Höhe

beider Verhandlungsparteien (69,2% unter den Forderungen der Hersteller). Generell galt: Je höher der Zusatznutzen ist, desto größer sein Gewicht in der Gesamtbewertung. Je geringer der Aussagewert der mitgeteilten europäischen Abgabepreise, desto geringer ihr Gewicht in der Gesamtbewertung. Preise vergleichbarer Arzneimittel schienen wiederum ein Indiz in den Augen der Schiedsstelle für die Angemessenheit des Erstattungsbetrages zu sein.

Es liegt eine schwache Evidenz für ein quasi-algorithmisches Vorgehen der Schiedsstelle hinsichtlich der Gewichtung der Faktoren aus der Rahmenvereinbarung in den Fällen mit Zusatznutzen vor. Jedenfalls hat die Analyse gezeigt, dass die synthetisch erzeugten Erstattungsbeträge deutlich näher an den Vorstellungen des GKV SV liegen. Der Interessensausgleich zwischen Herstellern und GKV SV spielt sich somit eher in einem eng abgesteckten Spielraum ab, obwohl es sich um Einzelfallentscheidungen handelt. In dieser Hinsicht scheint die Zielsetzung von AMNOG Schiedssprüchen nicht ausreichend erreicht.

Chapter (14): A decade of early benefit assessment of ophthalmic drugs in Germany: success story or not?

In der abschließenden Publikation der kumulativen Dissertation wird an einer Fallstudie zu Arzneimitteln zur Behandlung von Augenkrankheiten (Ophthalmika), die eine frühe Nutzenbewertung in der ersten Dekade des AMNOG durchlaufen haben, der gesamte Fragekomplex aus der Nutzenbewertung per se über die Erstattungsbetragsverhandlungen bis hin zur Schiedsstelle exemplarisch aufgearbeitet und auch das reale Marktgeschehen betrachtet. Ophthalmika bieten sich für diese Analyse aus verschiedenen Gründen an: überschaubare Anzahl an Fällen, Fälle mit und ohne Zusatznutzen, Subgruppenanalysen, indirekte Vergleiche, unterschiedliche Studiendesigns, konventionelle Zulassungen in sehr großen Indikationen und Orphan Drugs, patientenrelevante und Surrogat- bzw. Laborendpunkte, Mono- und Kombinationstherapien, Arzneimittel und Medizintechnologien als zweckmäßige Vergleichstherapien, Bewertungskaskaden durch Erweiterungen der Zulassungen sowie Neubewertungen nach Ablauf der Befristung des G-BA Beschlusses, Marktrücknahmen, Off-label Wettbewerb und zukünftige Biosimilars. Die Arbeit führt auch einen Vergleich der Ergebnisse der Bewertungen in Deutschland und in Frankreich durch. Beide Länder zeichnen sich durch eine dominierende vergleichende Effektivitätsanalyse von neuen Arzneimitteln aus, gelangen dennoch zu unterschiedlichen Ergebnissen. Darüber hinaus wird auch Bezug zu einer aktuellen internationalen Publikation von IQWiG Mitarbeitern genommen, in welcher die These aufgestellt wird, dass die Arzneimittelentwicklung bei mehr als der Hälfte von neuen Arzneimitteln in Deutschland zu einer Bewertung ohne Zusatznutzen geführt hat und deswegen reformiert werden müsste.

Jedes abgeschlossene Bewertungsverfahren zu Ophthalmika im Betrachtungszeitraum vom Anfang des AMNOG bis zum Ende März 2021 wurde identifiziert und hinsichtlich der zugrunde gelegten Studienevidenz, seiner Ergebnisse aus der frühen Nutzenbewertung sowie der verhandelten Erstattungsbeträge und seines realen Budget Impacts analysiert. Als primäre Datenquellen diente die G-BA Internetseite zur frühen Nutzenbewertung (IQWiG Bewertungen, Anhörungsprotokolle, IQWiG Addenda, G-BA Beschlussfassungen und deren Tragende Gründe) und die entsprechenden Daten der Haute Autorité de Santé (HAS) für den Vergleich mit Frankreich sowie die produktspezifischen Eintragungen in der Lauer-Taxe für die Erstattungsbeträge neben Publikationen der GKV zu den Verordnungsvolumina und zum Ausgabengeschehen.

Es konnten 13 Nutzenbewertungen zu 9 Ophthalmika (14 Subgruppen) identifiziert werden (eine Neubewertung nach Ablauf der Befristung ersetzte eine ältere Bewertung). In nur 4 von 14 Subgruppen (28,6%) konnte ein Zusatznutzen gezeigt werden (darunter drei Orphan Drugs mit Zusatznutzen qua Gesetz). Das Odds Ratio (Chancenverhältnis) für keinen Zusatznutzen bei Ophthalmika im Vergleich zu allen anderen bewerteten Arzneimitteln lag auf Subgruppenebene bei 1,667 (95%-KI: 0,519 – 5,356). Die Übereinstimmung zwischen dem von den Herstellern eingeforderten Zusatznutzen und den G-BA Beschlüssen gestaltete sich auf Subgruppenebene eher moderat (Cohen's kappa 0,462; 95%-KI: 0,086 – 0,837). Sowohl für das OR als auch für Cohen's kappa konnte aufgrund der relativ niedrigen Fallzahl keine statistische Signifikanz erzielt werden. Für die 12 (Tafluprost/Timolol nur in Deutschland bewertet) in den Vergleich zwischen Deutschland und Frankreich einbezogenen Fälle ergab sich eine mäßige Übereinstimmung (Cohen's kappa 0,333, 95%-KI: -0,170 – 0,836). In allen Fällen wurde Evidenz aus RCTs abgeleitet (es lag auch Evidenz niedrigerer Level wie bspw. einarmige Studien und Anwendungsbeobachtungsstudien vor), auch wenn nicht alle einen direkten Vergleich mit der vom G-BA festgesetzten Vergleichstherapie (neben Arzneimitteln die Medizintechnologien Laser und Photokoagulation) aufgrund eines anderen Studienkomparators oder einer nicht aktuellen fachinformationsgerechten Umsetzung in der Studie ermöglichten. Von 3 indirekten Vergleichen wurde nur einer vom G-BA anerkannt. In den Studien wurden sowohl aus Sicht des G-BA patientenrelevante Endpunkte wie bspw. Sehschärfe oder krankheitsspezifische Lebensqualität als auch nicht patientenrelevante Endpunkte wie intraokulärer Druck erhoben. Die verhandelten Erstattungsbeträge variierten zwischen 6,8% und 47,4% unter den Markteintrittspreisen, wobei die Abschläge für Ophthalmika (Median 14,5%) deutlich unter denen für alle verhandelten neuen Arzneimittel lagen (Median 24%). Der auf Verordnungen zu Lasten der GKV fußende reale Budget Impact für die zwei am häufigsten verordneten und zu den 10 Arzneimitteln mit den höchsten Umsätzen gehörenden Ophthalmika (Aflibercept und Ranibizumab) lag mit 663 Millionen €

weit unter dem vom G-BA in seinen Beschlussfassungen festgehaltenen virtuellen Umsatz in Höhe von 3,13 Milliarden €, wobei ein Großteil der Patienten (40%) mit dem preisgünstigeren Off-label Medikament Bevacizumab (onkologische Zulassung in anderen Organsystemen) versorgt wird.

In der Arbeit werden Opportunitätskosten nicht berücksichtigt, da deren Erhebung für das deutsche Gesundheitssystem, das anders als viele europäische Gesundheitssysteme ohne Kosten-Nutzwert-Analysen und entsprechenden Schwellenwerten zur Erstattung im Sinne impliziter Zahlungsbereitschaften auskommt, kaum bewältigbar ist. Dennoch kann mit Bezug auf Ophthalmika davon ausgegangen werden, dass neben Erstattungsbetragsverhandlungen und Preiswettbewerb durch gegenläufige Preiseffekte wie Rabatte auf Selektivvertragsebene und die in näherer Zukunft erwartete Einführung von Biosimilars parallel zur Off-label Anwendung von Bevacizumab sowie effizientere Therapieregime, die durch Ophthalmika induzierten Opportunitätskosten reduziert werden.

Hinsichtlich der IQWiG These, dass aufgrund (i) immer noch möglicher pivotaler Studien mit Plazebokontrolle, (ii) fehlenden Vergleichen neuer Arzneimittel mit Therapiestandards und (iii) einer inadäquaten Umsetzung des Komparators (Off-label Anwendungen oder inadäquaten Dosisregimen) sich die aktuelle Arzneimittelentwicklung ineffizient gestaltet, lässt sich entgegen, dass zumindest für Ophthalmika im Vergleich zum rigideren französischen Vorgehen die Ergebnisse der Nutzenbewertungen in Frankreich positiver ausfallen und dass wegen der gesetzlichen Zielsetzung der Kostendämpfung von neuen Arzneimitteln und dem damit verbundenen Bewertungsauftrag an HTA Institutionen, die Betrachtung von neuen Arzneimitteln nicht frei von finanziellen Hintergedanken gestaltet wird. Letztlich zeitigt die Versorgungsrealität ein anderes Bild, wenn trotz des sehr geringen Anteils an Ophthalmika mit einem Zusatznutzen aus dem Bewertungsverfahren – primär bedingt durch formale Gründe bei der Umsetzung des Studienkomparators – diese im Markt mit Ausnahme der wenigen Marktrücknahmen erfolgreiche Ergebnisse erzielen, weil sie bei den Verordnern einen positiven therapeutischen Stellenwert einnehmen, der nicht mit dem Ergebnis der frühen Nutzenbewertung einhergeht. Ob das 2017 mit dem GKV-AMVSG (Arzneimittelversorgungsstärkungsgesetz) eingeführte Arztinformationssystem (AIS) zur direkten Information der Verordner hinsichtlich des Ergebnisses der Nutzenbewertung, dessen wesentliche Eckpunkte durch das Bundesministerium für Gesundheit (BMG) in der Elektronischen Arzneimittelinformationen-Verordnung (EAMIV) 2019 festgelegt wurden, auf das Verordnungsverhalten Einfluss nehmen wird, bleibt noch abzuwarten.

1.4 Liste der Publikationen der kumulativen Dissertation

1. Dintsios, C.M., Schlenkrich, S. (2018). Industry's Experiences with the Scientific Advice offered by the Federal Joint Committee within the Early Benefit Assessment of Pharmaceuticals in Germany. *International Journal of Technology Assessment in Health Care* 34: 196–204, DOI: 10.1017/S0266462317004536
2. Bender, H., Dintsios, C. M. (2018). Gesundheitsbezogene Lebensqualität im Rahmen der frühen Nutzenbewertung von Arzneimitteln nach § 35a SGB V: Ein Endpunkt mit vielen Herausforderungen für alle beteiligten Akteure. *Gesundheitswesen* 80: 132–143, DOI: 10.1055/s-0042-116433
3. Niehaus, I., Dintsios, C. M. (2018). Confirmatory versus explorative endpoint analysis: Decision-making on the basis of evidence available from market authorization and early benefit assessment for oncology drugs. *Health Policy* 122: 599–606, DOI: 10.1016/j.healthpol.2018.03.017
4. Molitor, M., Dintsios, C. M., (2021). Failure due to formal reasons within German benefit assessment of medicinal products: the dilemma between marketing authorization and HTA. *Expert Review of Pharmacoeconomics & Outcomes Research* 21: 145–157, DOI: 10.1080/14737167.2020.1729131
5. Herpers, M., Dintsios, C. M. (2019). Methodological problems in the method used by IQWiG within early benefit assessment of new pharmaceuticals in Germany. *The European Journal of Health Economics* 20: 45–57, DOI: 10.1007/s10198-018-0981-3
6. Dintsios, C. M., Worm, F., Ruof, S., Herpers, M. (2019). Different interpretation of additional evidence for HTA by the commissioned HTA body and the commissioning decision maker in Germany: whenever IQWiG and Federal Joint Committee disagree. *Health Economics Review* 9:35 1–15, DOI: 10.1186/s13561-019-0254-6
7. Ruof, J., Staab, T., Dintsios, C.M., Schröter, J., Schwartz, F.W. (2016). Comparison of post-authorization measures from regulatory authorities with additional evidence requirements from the HTA body in Germany – are additional data requirements by the Federal Joint Committee justified? *Health Economics Review* 6:46 1–11, DOI: 10.1186/s13561-016-0124-4
8. Staab, T.R., Walter, M., Mariotti Nesurini, S., Dintsios, C.M., Graf von der Schulenburg, J.M., Amelung, V.E., Ruof, J. (2018). “Market withdrawals” of medicines in Germany after AMNOG: a comparison of HTA ratings and clinical guideline recommendations. *Health Economics Review* 8:23 1–11, DOI: 0.1186/s13561-018-0209-3
9. Gandjour, A., Schüßler, S., Hammerschmidt, T., Dintsios, C.M. (2020). Predictors of negotiated prices for new drugs in Germany. *The European Journal of Health Economics* 21:1049–1057, DOI: 10.1007/s10198-020-01201-z
10. Worm, F., Dintsios, C. M. (2018). Determinants of Orphan Drug Prices in Germany. *PharmacoEconomics* 38: 397–411, DOI: 10.1007/s40273-019-00872-8
11. Dintsios, C. M., Beinhauer, I. (2020). The impact of additive or substitutive clinical study design on the negotiated reimbursement for oncology pharmaceuticals after early benefit assessment in Germany. *Health Economics Review* 10:7 1–25, DOI: 10.1186/s13561-020-00263-2

12. Ludwig, S., Dintsios, C. M. (2016). Arbitration Board Setting Reimbursement Amounts for Pharmaceutical Innovations in Germany When Price Negotiations between Payers and Manufacturers Fail: An Empirical Analysis of 5 Years' Experience. *Value in Health* 19: 1016–1025, DOI: 10.1016/j.jval.2016.05.016
13. Dintsios, C. M. (2022). A decade of early benefit assessment of ophthalmic drugs in Germany: success story or not? *Expert Review of Pharmacoeconomics & Outcomes Research* 22: 283–297, DOI: 10.1080/14737167.2021.1930532

Literatur

- AM-NutzenV. (2010). Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V. https://www.bgbl.de/xaver/bgbl/start.xav#_bgbl_%2F%2F*%5B%40attr_id%3D%27bgbl1110s2324.pdf%27%5D_1523126802497 [Abgerufen 09 April 2018]
- AMNOG. (2010). Arzneimittelneuordnungsgesetz-AMNOG. *BGBl.* Teil I, Nr. 67, S. 2262-2277.
- AWMF. (2015). Stellungnahme zum Verfahren der Frühen Nutzenbewertung von Arzneimitteln nach § 35a SGBV und aufgrund des AMNOG von 2010 vom 24.02.2015. http://www.awmf.org/fileadmin/user_upload/Stellungnahmen/Medizinische_Versorgung/AWMF-Stellungnahme_AMNOG_2015-02-24.pdf [Abgerufen 27 Mai 2015]
- AWMF. (2016). Aktuelle Leitlinien. <https://www.awmf.org/awmfonline-das-portal-der-wissenschaftlichen-medizin/awmf-aktuell.html> [Abgerufen 15 Oktober 2016]
- Bauer, C., May, U., Wasem, J. (2016). IBES DISKUSSIONSBEITRAG. In: *Analyse und Beschreibung des AMNOG- Umsetzungsproblems in die Versorgungspraxis*. Universität Essen, Essen.
- Beinlich, P., Müller-Berghaus, J., Sudhop, T., Vieths, S., Broich, K. (2015). [Interplay between marketing authorization and early benefit assessment of drugs]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 58(3):227-231.
- Blankart, K. E. & Stargardt, T. (2020). The impact of drug quality ratings from health technology assessments on the adoption of new drugs by physicians in Germany. *Health Econ* 29 Suppl 1:63-82.
- Bonomi, P., Kim, K., Fairclough, D., Cella, D., Kugler, J., Rowinsky, E., et al. (2000). Comparison of survival and quality of life in advanced non-small-cell lung cancer patients treated with two dose levels of paclitaxel combined with cisplatin versus etoposide with cisplatin: results of an Eastern Cooperative Oncology Group trial. *J Clin Oncol* 18(3):623-631.
- Brouwer, W. B. & Rutten, F. F. (2010). The efficiency frontier approach to economic evaluation: will it help German policy making? *Health Econ* 19(10):1128-1131.
- Bungenstock, J. M. (2011). *Innovative Arzneimittel in der Gesetzlichen Krankenversicherung: Eine normativ-ökonomische Analyse zu Versorgung und Finanzierung*. Bd. 34, 1 Aufl. Europäische Schriften zu Staat und Wirtschaft. Nomos Verlagsgesellschaft mbH & Co. KG, Baden-Baden.
- Caro, J. J., Nord, E., Siebert, U., McGuire, A., McGregor, M., Henry, D., et al. (2010a). The efficiency frontier approach to economic evaluation of health-care interventions. *Health Econ* 19(10):1117-1127.
- Caro, J. J., Nord, E., Siebert, U., McGuire, A., McGregor, M., Henry, D., et al. (2010b). IQWiG methods-a response to two critiques. *Health Econ* 19(10):1137-1138.
- Cassel, D. & Heigl, A. (2013). AMNOG in der Umsetzung: Preisregulierung als Innovationsbremse? *RPG* 19(1):1-18.

- Cassel, D. & Ulrich, V. (2015). AMNOG auf dem ökonomischen Prüfstand: Funktionsweise, Ergebnisse und Reformbedarf der Preisregulierung für neue Arzneimittel in Deutschland. *Gesundheitsökonomische Beiträge*. Bd. 56, 1 Aufl. Nomos Verlag, Baden-Baden.
- Claxton, K., Briggs, A., Buxton, M. J., Culyer, A. J., McCabe, C., Walker, S., et al. (2008). Value based pricing for NHS drugs: an opportunity not to be missed? *Bmj* 336(7638):251-254.
- Coenen, M., Haucap, J., Loebert. (2016). BESTANDSAUFNAHME ZUM GEMEINSAMEN BUNDESAUSSCHUSS (G-BA). Stiftung Münch, München.
- Cook, J. P. & Golec, J. (2017). How excluding some benefits from value assessment of new drugs impacts innovation. *Health Econ* 26(12):1813-1825.
- Cuche, M., Beckerman, R., Chowdhury, C. A., van Weelden, M. A. (2014). Early dialogue with health technology assessment bodies: a European perspective. *Int J Technol Assess Health Care* 30(6):571-578.
- Dabisch, I., Dethling, J., Dintsios, C. M., Drechsler, M., Kalanovic, D., Kaskel, P., et al. (2014). Patient relevant endpoints in oncology: current issues in the context of early benefit assessment in Germany. *Health Econ Rev* 4(1):2.
- Danzon, P. M. (1997). Price Discrimination for Pharmaceuticals: Welfare Effects in the US and the EU. *International Journal of the Economics of Business* 4(3):301-322.
- Deutscher Bundestag. (2010). [Entwurf eines Gesetzes zur Neuordnung des Arzneimittelmarktes in der gesetzlichen Krankenversicherung (Arzneimittelmarktneuordnungsgesetz-AMNOG)]. <http://dipbt.bundestag.de/dip21/btd/17/031/1703116.pdf> [Abgerufen 31 August 2021]
- DGHO. (2013). Nutzenbewertung von Arzneimitteln der Onkologie und Hämatologie 2013. Eine Standortbestimmung. http://www.dgho.de/informationen/gesundheitspolitische-schriftenreihe/dgho_gpsr_arzneimittelnutzenbewertung.pdf [Abgerufen 31 August 2021]
- Dintsios, C. M. & Gerber, A. (2010). Some essential clarifications: IQWiG comments on two critiques of the efficiency frontier approach. *Health Econ* 19(10):1139-1141.
- Dintsios, C. M. & Schlenkrich, S. (2018). Industry's Experiences with the Scientific Advice offered by the Federal Joint Committee within the Early Benefit Assessment of Pharmaceuticals in Germany. *Int J Technol Assess Health Care* 34(2):196-204.
- EMA. (1998). ICH Topic E 9 Statistical Principles for Clinical Trials (CPMP/ICH/363/96).
- Faber, U. (2014). Drei Jahre frühe Nutzenbewertung aus Patientensicht. Drei Jahre frühe Nutzenbewertung: Erfahrungen - Standortbestimmung - Weiterentwicklung. AMNOG-Fachtagung des G-BA am 30.04.2014 in Berlin. https://www.g-ba.de/downloads/17-98-3850/04-2014-04-30_AMNOG-Fachtagung-G-BA_Faber.pdf [Abgerufen 09 April 2017]
- Fairclough, D. (2010). *Design and Analysis of Quality of Life Studies in Clinical Trials*, 2nd ed. Chapman & Hall, Boca Raton.
- Fischer, K. E. & Stargardt, T. (2014). Early benefit assessment of pharmaceuticals in Germany: manufacturers' expectations versus the Federal Joint Committee's decisions. *Med Decis Making* 34(8):1030-1047.
- Fritsch, M. (2014). *Marktversagen und Wirtschaftspolitik: Mikroökonomische Grundlagen staatlichen Handelns*, 9 Aufl. Vahlens Handbücher der Wirtschafts- und Sozialwissenschaften. Verlag C.H.Beck, München.

- G-BA. (2011). Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Ticagrelor.
- G-BA. (2012). Chapter 5: Assessment of the benefits of pharmaceuticals according to §35a SGB V. http://www.english.g-ba.de/downloads/17-98-3042/2011-05-18_5%20Kapitel%20VerfO%20Englisch.pdf [Abgerufen 15 Oktober 2020]
- Gandjour, A. & Chernyak, N. (2011). A new prize system for drug innovation. *Health Policy* 102(2-3):170-177.
- Gandjour, A. (2018). Underuse of innovative medicines in Germany: A justification for government intervention? *Health Policy* 122(12):1283-1286.
- Gandjour, A. (2020). Willingness to pay for new medicines: a step towards narrowing the gap between NICE and IQWiG. *BMC Health Serv Res* 20(1):343.
- Gerber, A. & Dintios, C. M. (2010). A distorted picture of IQWiG methodology. *Health Aff (Millwood)* 29(1):220-221; author reply 221.
- Gerhardus, A. & Dintios, C. M. (2005). The impact of HTA reports on health policy: a systematic review. *GMS Health Technol Assess* 1:Doc02.
- German Federal Ministry of Health. (2016). Strukturierte Zusammenarbeit zwischendem Gemeinsamen Bundesausschuss dem Bundesinstitut für Arzneimittel und Medizinprodukte und dem Paul-Ehrlich-Institut. http://www.bmg.bund.de/fileadmin/dateien/Downloads/P/Pharmadialog/Vereinbarung_G-BA-BOB_2016-04-12.pdf [Abgerufen 15 Oktober 2020]
- GKV-Schätzerkreis. Pressemitteilung des GKV-Schätzerkreises vom 9. Dezember 2009. http://www.gkv-spitzenverband.de/media/dokumente/presse/pressemitteilungen/2009/PM_Schaetzerkreis091209_10561.pdf [Abgerufen 20 September 2017]
- GKV-SV, Bundesverband der Arzneimittel-Hersteller e.V., Bundesverband der Pharmazeutischen Industrie e.V., Pro Generika e.V., Verband Forschender Arzneimittelhersteller e.V. (2012). Rahmenvereinbarung nach §130b (9) SGB V. http://www.pharmatrix.de/cms/upload/pdf/Recht/arzneimittelrecht/12/RahmenvereinbErstattBetrag_130_b_SGB_V_us_1205.pdf [Aufgerufen 19 März 2012].
- GKV-SV, Bundesverband der Arzneimittel-Hersteller e.V., Bundesverband der Pharmazeutischen Industrie e.V., Pro Generika e.V., Verband Forschender Arzneimittelhersteller e.V. (2016). Rahmenvereinbarung nach §130b (9) SGB V. https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/arzneimittel/rahmenvertraege/pharmazeutische_unternehmer/Rahmenvereinbarung_130b_Abs9_SGB_V_2016.pdf [Aufgerufen 30.06.2016].
- Greiner, W., Kuhlmann, A., Schwarzbach, C. (2010). Ökonomische Beurteilung des Effizienzgrenzenkonzeptes. *Gesundheitsökonomie & Qualitätsmanagement* 15(05):241-250.
- Grouven, U., Bender, R., Ziegler, A., Lange, S. (2007). Der Kappa-Koeffizient. *Dtsch Med Wochenschr* 132(S 01):e65-e68.
- Grüne, M. (2013). „Bedeutung der Zulassung für die Nutzenbewertung“- Stellenwert des Zulassungsstatus bei G-BA- Entscheidungen. https://www.iqwig.de/download/13-06-21_IQWiG_im_Dialog_Maximilian_Gruene_Bedeutung_der_Zulassung_Nutzenbewerту

- [ng Stellenwert des Zulassungsstatus bei G-BA-Entscheidungen.pdf](#) [Abgerufen 20 Oktober 2015]
- Hammerschmidt, T. (2017). Analysis of German new drug prices after AMNOG in relation to European prices. *Gesundh ökon Qual manag* 22(01):43-53.
- Hecken, J. (2014a). Rückblick 2014 – Ausblick 2015. *Schütze-Brief: Gesundheitspolitischer Info-Dienst* 98:1-20.
- Hecken, J. (2014b). Begrüßung und Einführung des Unparteiischen Vorsitzenden. Drei Jahre frühe Nutzenbewertung: Erfahrungen - Standortbestimmung - Weiterentwicklung. AMNOG-Fachtagung des G-BA am 30.04.2014 in Berlin. https://www.g-ba.de/downloads/17-98-3691/00-2014-04-30-AMNOG-Fachtagung-G-BA_Hecken.pdf [Abgerufen 25 Februar 2016]
- Herndon, J. B. (2002). Health insurer monopsony power: the all-or-none model. *J Health Econ* 21(2):197-206.
- Herpers, M. & Dintios, C. M. (2019). Methodological problems in the method used by IQWiG within early benefit assessment of new pharmaceuticals in Germany. *Eur J Health Econ* 20(1):45-57.
- Hinterhuber, A. (2004). Towards value-based pricing—An integrative framework for decision making. *Industrial Marketing Management* 33(8):765-778.
- Hörn, H., Nink, K., McGauran, N., Wieseler, B. (2014). Early benefit assessment of new drugs in Germany - results from 2011 to 2012. *Health Policy* 116(2-3):147-153.
- Hüer, T., Abels, C., Wasem, J. (2021). 10 Jahre AMNOG aus Sicht der Gesundheitsökonomik. *Gesundheits- und Sozialpolitik (G&S)* 75(1):7-12.
- IQWiG. (2013). Documentation and evaluation of comments on the Draft of the General Methods 4.1. https://www.iqwig.de/download/Dokumentation_und_Wuerdigung_der_Stellungnahmen_IQWiG_Methoden_4-1.pdf [Abgerufen 25 April 2015]
- IQWiG. (2015). General Methods. Version 4.2 of 22 April 2015. https://www.iqwig.de/download/IQWiG_General_Methods_Version_4-2_no_longer_valid.pdf [Abgerufen 17 Juni 2016]
- IQWiG. (2013). Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medikamentösen Behandlungen.
- Ivandic, V. (2014). Requirements for benefit assessment in Germany and England - overview and comparison. *Health Econ Rev* 4(1):12.
- Jommi, C., Armeni, P., Costa, F., Bertolani, A., Otto, M. (2020). Implementation of Value-based Pricing for Medicines. *Clin Ther* 42(1):15-24.
- Kaiser, T., Vervölgyi, V., Wieseler, B. (2015). Nutzenbewertung von Arzneimitteln. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 58(3):232-239.
- Kanavos, P., Nicod, E., Espin, J., Aardwg, S. (2010). Short- and long-term effects of value-based pricing vs. external price referencing. In, *EMINET*. LSE, London.
- Kanavos, P., Fontrier, A.-M., Gill, J., Efthymiadou, O. (2020). Does external reference pricing deliver what it promises? Evidence on its impact at national level. *The European Journal of Health Economics* 21(1):129-151.

- Kifmann, M. (2010). Indikationsspezifische Kosten-Nutzen-Bewertung auf Grundlage eines sozialen Gesundheitsindexes. *PharmacoEconomics German Research Articles* 8(1):5-19.
- Klakow-Franck, R. (2014). [The relevance of quality of life for the work of the Federal Joint Committee]. *Z Evid Fortbild Qual Gesundheitswes* 108(2-3):151-156.
- Köhler, M., Haag, S., Biester, K., Brockhaus, A. C., McGauran, N., Grouven, U., et al. (2015). Information on new drugs at market entry: retrospective analysis of health technology assessment reports versus regulatory reports, journal publications, and registry reports. *Bmj* 350:h796.
- Kvitkina, T., ten Haaf, A., Reken, S., McGauran, N., Wieseler, B. (2014). Patientenrelevante Endpunkte und Surrogate in der frühen Nutzenbewertung von Arzneimitteln: erste Erfahrungen. *Z Evid Fortbild Qual Gesundheitswes* 108(8-9):528-538.
- Lange, S. (2015). Besonderheiten von Onkologischen Studien im Rahmen der frühen Nutzenbewertung von Arzneimittel. *Z Evid Fortbild Qual Gesundheitswes* 109(6):417-430.
- Lauenroth, V. D. & Stargardt, T. (2017). Pharmaceutical Pricing in Germany: How Is Value Determined within the Scope of AMNOG? *Value in Health* 20(7):927-935.
- Liozu, S. M. (2017). Value-based pricing special issue: Editorial. *Journal of Revenue and Pricing Management* 16(1):1-3.
- Longson, C., Crabb, N., Osipenko, L. (2014). NICE supports allowing technology assessment bodies to provide advice to drug industry. *Bmj* 349:g5035.
- Maignen, F., Osipenko, L., Pinilla-Dominguez, P., Crowe, E. (2017). Integrating health technology assessment requirements in the clinical development of medicines: the experience from NICE scientific advice. *Eur J Clin Pharmacol* 73(3):297-305.
- Mayntz, R. & Scharpf, F. (1995). *Gesellschaftliche Selbstregulierung und Steuerung*. Campus, Frankfurt a. M.
- Moïse, P. & Docteur, E. (2007). Pharmaceutical Pricing and Reimbursement Policies in Sweden. In, *Health Working Papers*. OECD, Paris.
- Mühlbacher, A. C. & Sadler, A. (2017). The Probabilistic Efficiency Frontier: A Framework for Cost-Effectiveness Analysis in Germany Put into Practice for Hepatitis C Treatment Options. *Value Health* 20(2):266-272.
- Pauly, M. V. (1988). Market power, monopsony, and health insurance markets. *J Health Econ* 7(2):111-128.
- Pindyck, R. S. & Rubinfeld, D. L. (2009). *Mikroökonomie*. Pearson Deutschland.
- Porter, M. E. & Teisberg, E. O. (2006). *Redefining Health Care: Creating Value-Based Competition on Results*. Harvard Business School Press, Boston.
- Porter, M. E. (2010). What is value in health care? *N Engl J Med* 363(26):2477-2481.
- Prieto-Pinto, L., Garzón-Orjuela, N., Lasalvia, P., Castañeda-Cardona, C., Rosselli, D. (2020). International Experience in Therapeutic Value and Value-Based Pricing: A Rapid Review of the Literature. *Value Health Reg Issues* 23:37-48.
- Radic, D., Haugk, S., Radic, M. (2018). Nutzenbewertung und Preisverhandlung unter AMNOG: Berechenbares Verfahren oder unfaire Pokerpartie?. *Gesundheitswesen* 80(6):573-579.

- Raftery, J. (2013). Value based pricing: can it work? *Bmj* 347:f5941.
- Rémuzat, C., Urbinati, D., Mzoughi, O., El Hammi, E., Belgaied, W., Toumi, M. (2015). Overview of external reference pricing systems in Europe. *J Mark Access Health Policy* 3.
- Röhmel, J. (2012). Gutachten zum Vorschlag des IQWiG zur Bewertung des Ausmaßes des Zusatznutzens im Rahmen der Nutzenbewertung von Arzneimitteln nach §35a SGB V. In.
- Ruof, J., Knoerzer, D., Dunne, A. A., Dintsios, C. M., Staab, T., Schwartz, F. W. (2014a). Analysis of endpoints used in marketing authorisations versus value assessments of oncology medicines in Germany. *Health Policy* 118(2):242-254.
- Ruof, J., Schwartz, F. W., Schulenburg, J. M., Dintsios, C. M. (2014b). Early benefit assessment (EBA) in Germany: analysing decisions 18 months after introducing the new AMNOG legislation. *Eur J Health Econ* 15(6):577-589.
- Sandmann, F. G., Mostardt, S., Lhachimi, S. K., Gerber-Grote, A. (2018). The efficiency-frontier approach for health economic evaluation versus cost-effectiveness thresholds and internal reference pricing: combining the best of both worlds? *Expert Rev Pharmacoecon Outcomes Res* 18(5):475-486.
- Sattelmeier, J., Prenzler, A., Frank, M. (2013). Das Arzneimittelmarktneuordnungsgesetz (AMNOG) in der Praxis – erste Erfahrungen, Kritikpunkte und Weiterentwicklungsmöglichkeiten. *Gesundh ökon Qual manag* 18(05):213-220.
- Schlender, M., Dintsios, C. M., Gandjour, A. (2018). Budgetary Impact and Cost Drivers of Drugs for Rare and Ultrarare Diseases. *Value Health* 21(5):525-531.
- Schwabe, U. & Ludwig, W.-D. (2020). Arzneiverordnungen 2019 im Überblick. In Schwabe, U. & Ludwig, W.-D. (Hrsg.), *Arzneiverordnungs-Report 2020*. Berlin, Heidelberg, Springer Berlin Heidelberg. S. 3-41.
- Schwarzbach, C., Kuhlmann, A., Vauth, C., von der Schulenburg, J. M. G. (2009). Die Effizienzgrenze des IQWiG: Eine gelungene Adaption der Modelle von Markowitz und Koopmans? *Sozialer Fortschritt* 58(7):137-147.
- Sculpher, M. & Claxton, K. (2010). Sins of omission and obfuscation: IQWiG's guidelines on economic evaluation methods. *Health Econ* 19(10):1132-1136.
- Shiroiwa, T. (2020). Cost-effectiveness evaluation for pricing medicines and devices: A new value-based price adjustment system in Japan. *Int J Technol Assess Health Care* 36(3):270-276.
- Skipka, G., Wieseler, B., Kaiser, T., Thomas, S., Bender, R., Windeler, J., et al. (2016). Methodological approach to determine minor, considerable, and major treatment effects in the early benefit assessment of new drugs. *Biom J* 58(1):43-58.
- Sussex, J., Towse, A., Devlin, N. (2013). Operationalizing value-based pricing of medicines : a taxonomy of approaches. *Pharmacoeconomics* 31(1):1-10.
- Theidel, U. & von der Schulenburg, J. M. (2016). Benefit assessment in Germany: implications for price discounts. *Health Econ Rev* 6(1):33.
- Towse, A. & Barnsley, P. (2013). Approaches to identifying, measuring, and aggregating elements of value. *Int J Technol Assess Health Care* 29(4):360-364.

- Vach, W. (2014). Quantifizierung des Ausmaßes des Zusatznutzens von neuen Arzneimitteln: „gering“ – „beträchtlich“ – „erheblich“ – 6 Anmerkungen aus der Sicht eines Biometrikers. *Gesundheitswesen* 76(11):757-762.
- Vermeulen, K. M. & Krabbe, P. F. M. (2018). Value judgment of health interventions from different perspectives: arguments and criteria. *Cost Eff Resour Alloc* 16:16.
- Voigt, S. (2009). *Institutionenökonomik: Neue Ökonomische Bibliothek*. UTB GmbH, München.
- Vorderwülbecke, U. (2013). Gemeinsamer Bundesausschuss und Arzneimittelhersteller: Ein Governance-Problem. *PharmR* 4:149-152.
- Wonder, M., Backhouse, M. E., Hornby, E. (2013). Early scientific advice obtained simultaneously from regulators and payers: findings from a pilot study in Australia. *Value Health* 16(6):1067-1073.
- Wright, J. S., Barron, A. J., Shah, S. M., Klingler, C. (2017). Convergence, divergence and hybridity: a regulatory governance perspective on health technology assessment in England and Germany. *Global Policy* 8:69-75.