

Unpublizierte Daten aus klinischen Studien zu Arzneimitteln

Im AMNOG-Verfahren müssen Hersteller per Gesetz ein umfangreiches Dossier vorlegen, das alle relevanten Studiendaten enthält. Oft bieten die AMNOG-Bewertungen des IQWiG deshalb mehr Informationen als wissenschaftliche Fachzeitschriften.

Rüdig, C.
Kranz, P.
Flintrop, J.
Kaiser, T.

Randomisierte klinische Studien (RCT) sind der Goldstandard in der medizinischen Forschung. Ihre Ergebnisse werden hauptsächlich in hochrangigen Fachzeitschriften veröffentlicht, die in der medizinischen Gemeinschaft große Beachtung finden. Bei der Darstellung medizinischer Erkenntnisse stützen sich Fachleute oft ausschließlich auf diese Publikationen. Andere Quellen wie Berichte der Europäischen Zulassungsbehörde und Einträge in Studienregistern liefern ebenfalls Daten aus klinischen Studien, werden aber seltener genutzt. Allerdings fehlen in diesen Quellen oft relevante Ergebnisse und Informationen zur Studienmethodik. In der frühen Nutzenbewertung von Arzneimitteln in Deutschland nach dem Arzneimittelmarktneuordnungsgesetz (AMNOG) stehen diese Informationen hingegen zur Verfügung. Sie werden in den Bewertungen und dem öffentlichen Teil des Herstellerdossiers veröffentlicht.

Zwei Beispiele verdeutlichen die Problematik:

- **Unpublizierte Daten:** In einer Publikation zur maßgeblichen Studie für die Zulassung von Linagliptin bei Diabetes mellitus Typ 2 (siehe Abbildung 1) berichteten die Autoren, dass unter Linagliptin in einer ausgewählten Patientengruppe signifikant weniger Schlaganfälle auftraten als unter der Vergleichstherapie Glimepirid. Erst ein Blick in den Studienbericht im Rahmen des AMNOG-Verfahrens zeigte, dass dieser vermeintliche Vorteil von Linagliptin gegenüber Glimepirid auf unterschiedliche Therapiestrategien in den beiden Studienarmen zurückzuführen war: In der ersten Studienphase sollten in der Glimepirid-Gruppe die Blutzuckerwerte rasch in den normnahen Bereich abgesenkt werden. Dies führte vermehrt zu Unterzuckerungen und in der Folge zu Schlaganfällen. In der Linagliptin-Gruppe erfolgte die Therapie ohne festgelegten Blutzuckerzielwert, hier kam es nicht zu Hypoglykämien und somit auch nicht zu gehäuften Schlaganfällen. Die zunächst selektive Veröffentlichung der Studienergebnisse zu Linagliptin ist ein gutes Beispiel für einen Reporting Bias.
- **Ungenutzte Daten:** Regorafenib wurde im Rahmen des AMNOG-Verfahrens in der Indikation Kolorektalkarzinom bewertet (siehe Abbildung 2). Die Studiendaten zeigten einen Überlebensvorteil unter Regorafenib, aber auch schwere Nebenwirkungen. Wie es den Patientinnen und Patienten unter der Behandlung mit Regorafenib ging, blieb zunächst unklar, weil die Auswertungen des Herstellers nicht interpretierbar waren. Erst in einer zweiten Bewertung konnten die Auswertungen der erhobenen Daten zur patientenberichteten Symptomatik und zur Lebensqualität sinnvoll analysiert werden. Dabei zeigte sich, dass Patientinnen und Patienten unter Regorafenib ihre gesundheitsbezogene Lebensqualität deutlich schlechter einschätzten als die Vergleichsgruppe. Die deutlich schlechtere Lebensqualität stellt den Zusatznutzen von Regorafenib insgesamt infrage.

Beispiel für fehlende Informationen in Zeitschriftenpublikationen: Linagliptin

Linagliptin wurde 2011 zur Behandlung des Diabetes mellitus Typ 2 zugelassen und in Folgenden im Rahmen einer frühen Nutzenbewertung gemäß AMNOG bewertet. Das dem IQWiG vorgelegte Dossier des Herstellers enthielt Daten für einen Vergleich von Linagliptin mit der zweckmäßigen Vergleichstherapie vor, und zwar für die Zweifachtherapie Linagliptin plus Metformin vs. Glimepirid plus Metformin (Studie 1218.20). In der Publikation zur Studie (Gallwitz B et al. Lancet 2012; 380: 475–83) wurde berichtet, dass unter Linagliptin signifikant weniger nicht tödliche Schlaganfälle aufgetreten waren als unter der Vergleichstherapie Sulfonylharnstoff. Die aufgetretenen Schlaganfälle wurden nicht mit Hypoglykämien in Verbindung gebracht. Für die Bewertung im Rahmen des AMNOG lag der Studienbericht vor. Die Daten hieraus zeigten eine zeitliche Übereinstimmung zwischen dem Auftreten von Hypoglykämien und Schlaganfällen: Diese traten unter Glimepirid insbesondere in der Anfangsphase der Studie auf, wo Glimepirid forciert titriert wurde, und in der Folge der HbA1c-Wert erheblich niedriger war als im Linagliptin-Arm. Auffällig war, dass in der Publikation eine Auswertung der HbA1c-Werte der der Population derjenigen Patientinnen und Patienten, die bis zum Ende der Studie teilgenommen hatten (Completer Set), dargestellt wurde. Der Unterschied im HbA1c zwischen den Behandlungsarmen war hier weniger stark ausgeprägt als in der gesamten Studienpopulation (Full Analysis Set) und suggerierte damit keine Verbindung zwischen Hypoglykämien und Schlaganfällen. In der gesamten Studienpopulation hingegen zeigte der zeitliche Verlauf des HbA1c insbesondere in der Anfangsphase der Studie deutlich geringere Werte in der Glimepirid-Gruppe als in der Linagliptin-Gruppe. Die unübliche forcierte Anwendung von Glimepirid in der Studie führte zu einem verzerrten Bild der Vergleichstherapie, nämlich zu sehr niedrigen HbA1c-Werten und zeitlich korrespondierenden Schlaganfällen.

Abbildung 1: Beispiel Linagliptin

Beispiel für ungenutzte Daten in einer Leitlinie: Regorafenib

Regorafenib wurde 2013 zur Behandlung des metastasierten Kolorektalkarzinoms erstmals zugelassen. Die erste Bewertung im Rahmen des AMNOG stützte sich auf die CORRECT-Studie und zeigte bei schweren unerwünschten Ereignissen einen statistisch signifikanten Nachteil von Regorafenib. Gleichzeitig lag ein Vorteil beim Endpunkt Gesamtmortalität vor. Daten zur Symptomatik (erhoben als patientenberichtete Symptomatik) und zur gesundheitsbezogenen Lebensqualität waren nicht geeignet. Erst im Stellungnahmeverfahren der zweiten Bewertung nach Befristung des G-BA-Beschlusses lagen genügend Daten vor, um eine eigene Berechnung der Effekte vorzunehmen. Die Analysen des Herstellers waren nicht adäquat. Nun zeigte sich für Regorafenib neben dem Vorteil beim Gesamtüberleben und dem erheblichen Nachteil bei schweren unerwünschten Ereignissen zusätzlich ein Nachteil bei der gesundheitsbezogenen Lebensqualität. Das stellt den Zusatznutzen von Regorafenib insgesamt infrage. Diese Ergebnisse zur gesundheitsbezogenen Lebensqualität aus dem AMNOG-Verfahren fanden 2017 keinen Eingang in die Leitlinie zum Kolorektalkarzinom. Damit findet diese Information bislang keine Berücksichtigung bei der Behandlung der Betroffenen. Die Leitlinie wurde bisher nicht aktualisiert, befindet sich nun aber in der Überarbeitung.

Abbildung 2: Beispiel Regorafenib

Die Beispiele zeigen: Informierte Entscheidungen zu Therapien, sei es durch Ärztinnen und Ärzte, Patientinnen und Patienten oder kostentragende Institutionen, können nur auf der Basis vollständiger Daten gemacht werden. Stehen nur unvollständige Daten zur Verfügung, wie im Beispiel Linagliptin aufgrund von selektiver Publikation oder im Beispiel Regorafenib aufgrund der Nichtnutzung von verfügbaren Informationen, besteht das Risiko von Fehlentscheidungen zum Nachteil der Patientinnen und Patienten. Informationsquellen, wie Patientenleitlinien oder Gesundheitsinformationen, die sich auf Behandlungsleitlinien stützen, können Patientinnen und Patienten nur eingeschränkte Informationen bieten, wenn die Behandlungsleitlinien öffentlich verfügbare Daten nicht berücksichtigen.

Systematische Untersuchung des Informationsgehalts öffentlich verfügbarer Quellen zu Studiendaten

Um die Ergebnisse einer Studie richtig zu interpretieren, braucht es Informationen zur geplanten und tatsächlichen Durchführung der Studien. Diese stehen in den Studienberichten, die die Hersteller allerdings selten veröffentlichen. Wichtige Angaben sind detaillierte Studienbeschreibungen, geplante und tatsächliche Nachbeobachtungszeiten, Bewertungen möglicher Verzerrungen, Folgetherapien in onkologischen Studien, Übertragbarkeit der Ergebnisse, kritische Evaluationen des Komparators sowie Ergebnisse zu in der Versorgung relevanten Teilpopulationen. Teilweise tauchen solche Informationen inzwischen in Anhängen zu Publikationen auf, dies ist aber noch nicht durchgängig der Fall. Im AMNOG-Verfahren sind diese Informationen hingegen verfügbar und werden in den Bewertungen und dem öffentlichen Teil des Herstellerdossiers veröffentlicht.

Das IQWiG hat systematisch untersucht, wie wissenschaftliche Publikationen, Zulassungsberichte und Studienregister (zusammengefasst als Non-AMNOG) im Vergleich zu AMNOG-Bewertungen informieren. Es prüfte, wie umfassend Methodik und Ergebnisse zu relevanten Endpunkten in Non-AMNOG-Quellen im Vergleich zu den AMNOG-Dossiers berichtet werden (1). Die Ergebnisse sind in Abbildung 3 dargestellt.

Die systematische Untersuchung zeigte, dass Non-AMNOG-Unterlagen deutlich weniger Informationen enthalten als die Bewertungen und Dossiers im AMNOG-Verfahren. Dies betrifft sowohl die Studienmethodik als auch die Ergebnisse von Endpunkten. Dabei werden die patientenberichteten Endpunkte zur Symptomatik und insbesondere zur gesundheitsbezogenen Lebensqualität in öffentlich zugänglichen Quellen kaum dargestellt, selbst wenn sie in Studien erhoben wurden (1, 2).

Reporting Bias

Reporting Bias entsteht, wenn Studiendaten unvollständig veröffentlicht werden und dadurch ein verzerrtes Bild der Ergebnisse oder des untersuchten Wirkstoffs entsteht. Der Oberbegriff umfasst zwei unterschiedliche Sachverhalte:

- Publikationsbias bezeichnet die Tatsache, dass zu Studien gar keine Publikationen erstellt werden, und dies eher Studien mit „negativem“ Ausgang betrifft.
- Outcome Reporting Bias bezeichnet die Tatsache, dass es zwar eine oder mehrere Publikationen zu einer Studie gibt, diese aber nur unvollständige Informationen ent-

Literatur

- 1 Köhler M, Haag S, Biester K et al. Information on new drugs at market entry: retrospective analysis of health technology assessment reports versus regulatory reports, journal publications, and registry reports. *BMJ* 2015; 350: h796. doi: 10.1136/bmj.h796.
- 2 Wieseler B, Wolfram N, McGauran N et al. Completeness of reporting of patient-relevant clinical trial outcomes: comparison of unpublished clinical study reports with publicly available data. *PLoS Med* 2013; 10(10):e1001526. doi: 10.1371/journal.pmed.1001526.

halten. Dies kann zum einen vermeintlich weniger relevante oder wiederum „negative“ Ergebnisse betreffen (z.B. zur Lebensqualität), es kann sich aber auch um nachträglich angepasste Auswertungen handeln, aus denen sich dann ein vermeintlich „positives“ Ergebnis ergibt.

Wie oben beschrieben, kann die selektive Publikation von klinischen Studiendaten zu einer verzerrten Gesamteinschätzung der untersuchten Behandlung und zu falschen Schlussfolgerungen führen. Im schlimmsten Fall verschwenden Ärzte und Patienten Zeit und Geld für unwirksame oder gefährliche Therapien.

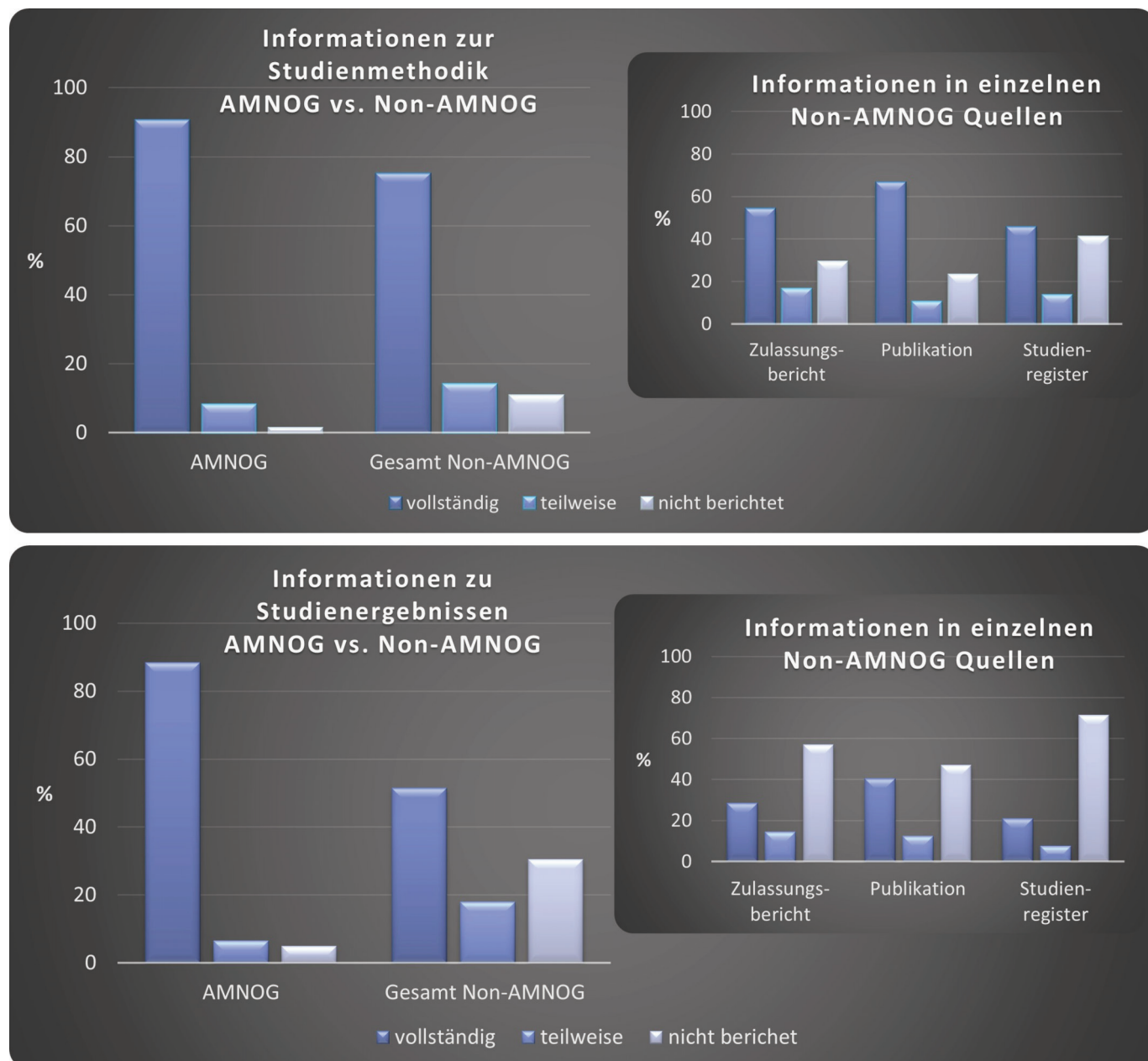


Abbildung 3: Verfügbarkeit von Informationen zur Studienmethodik und Studienergebnissen in öffentlich verfügbaren Quellen zu Studiendaten (1)

Öffentlich verfügbaren Quellen zu Studiendaten im Detail

Studien zu neuen Arzneimitteln werden inzwischen nahezu regelhaft publiziert, sodass sich insbesondere die Frage des Outcome Reporting Bias stellt. Denn in wissenschaftlichen Publikationen in Fachzeitschriften fehlen oft relevante Informationen zum Studiendesign und zu Ergebnissen. Zwar veröffentlichen Autorinnen und Autoren zunehmend weitere Studienunterlagen wie Studienprotokolle und ergänzende Daten in Anhängen, doch dies geschieht zum einen noch nicht flächendeckend. Zum anderen sind auch mit diesen Anhängen die Informationen zumeist unvollständig. Zu oft beschränken sich Publikationen auf die Patientencharakteristika, die Ergebnisse des primären Endpunkts und einige sekundäre Endpunkte wie das Gesamtüberleben sowie auf eine deskriptive Beschreibung der unerwünschten Ereignisse.

Ergebnisse zu weiteren, patientenrelevanten sekundären Endpunkten (insbesondere Symptomatik und Lebensqualität), die für Therapieentscheidungen wichtig sein können, fehlen meist oder werden erst in späteren Publikationen berichtet. Auch relevante Angaben, die für die Interpretation der Ergebnisse notwendig sind – detaillierte Studienbeschreibung, geplante und tatsächliche Dauer der Nachbeobachtung, Folgetherapien in onkologischen Studien, Übertragbarkeitsaspekte, kritische Evaluation des eingesetzten Komparators, Ergebnisse zu in der Versorgung relevanten Teilpopulation – werden oft nicht berichtet.

Zulassungsberichte der europäischen (EMA) und US-amerikanischen (FDA) Zulassungsbehörden liefern zwar oft weitergehende Daten, doch auch diese sind oft lückenhaft. Sie beschreiben das Studiendesign detailliert, beschreiben dessen Schwächen und ordnen die Studienergebnisse unter Einbezug dieser Informationen ein. Die Darstellung der Ergebnisse ist umfangreich. Die Zulassungsbehörde wägt jedoch primär zwischen Wirksamkeit und Sicherheit eines Arzneimittels ab. Ein Vergleich mit dem aktuellen Therapiestandard, der für Therapieentscheidungen in der Versorgung entscheidend ist, gehört nicht zur Bewertung der Zulassungsbehörde. Patientenberichtete Endpunkte und Angaben zur gesundheitsbezogenen Lebensqualität werden nicht immer dargestellt; insbesondere dann nicht, wenn diese als explorative Endpunkte geplant wurden. Auch Analysen von Teilpopulationen und Subgruppen, die für einen Versorgungskontext relevant sind, bleiben unberücksichtigt, wenn sie nicht für die Zulassung relevant sind.

Eine weitere Datenquelle sind Studienregister. Sie wurden in den letzten Jahren ausgebaut und liefern Informationen zu Studiendesign und Ergebnissen. Doch ihr Umfang ist begrenzt, die Daten erscheinen verspätet und sind nicht immer zuverlässig.

In der frühen Nutzenbewertung nach AMNOG stehen neben den öffentlich verfügbaren Informationen auch unveröffentlichte Daten aus den Herstellerdossiers zur Verfügung. Unter anderem liegt regelhaft ein umfangreicher Studienbericht¹ vor. Darüber hinaus fordern die Dossievorgaben des G-BA unter anderem auch vollständige Analysen zu allen patientenrelevanten Endpunkten. Zudem enthalten die Unterlagen Analysen von Teilpopulationen und Subgruppen, die für den deutschen Versorgungskontext relevant sind. Das Dossier enthält damit Informationen und Daten, die weit über den Informationsgehalt anderer Veröffentlichungen hinausgehen. Die relevanten Daten werden in den AMNOG-Bewertungen des IQWiG und im Beschluss des G-BA veröffentlicht. Ein Großteil des Dossiers erscheint zudem auf der Website des G-BA. Die Ergebnisse der frühen

¹ Studienberichte bestehen aus einem Kerntext von ca. 200 Seiten, in dem die Studienplanung, -durchführung, Analyse und Ergebnisse beschrieben sind. In Anhängen zum Kerntext, die mehrere 1000 Seiten umfassen, sind die Dokumente zur Studienplanung (u. a. sämtliche Versionen des Studienprotokolls und des Analyseplans, Datenerhebungsformulare und Fragebögen) und weitere Ergebnisse (u. a. Auswertungen und Rohdaten, Fallbeschreibungen von schwerwiegenden unerwünschten Ereignissen) zu finden. Im Vergleich dazu umfassen wissenschaftliche Publikationen von Studien ca. 15 Seiten. Teilweise sind in Anhängen Tabellen und Abbildungen zu finden, deren Daten narrativ im Haupttext beschrieben werden, sowie das Studienprotokoll und der statistische Analyseplan.

Nutzenbewertung fließen auch in das Arztinformationssystem ein und stehen damit den Ärztinnen und Ärzten direkt zur Verfügung.

Auch für die Erstellung von medizinischen Leitlinien stehen diese Daten natürlich zur Verfügung. Allerdings werden die Daten aus dem AMNOG-Verfahren bisher nicht systematisch und umfassend bei der Leitlinienerstellung berücksichtigt (3). Das IQWiG arbeitet derzeit in einem Projekt im Rahmen des Generalauftrags daran, diese Lücke zu schließen (4).

Ausblick auf die europäische Nutzenbewertung

Die Einführung der europäischen Nutzenbewertung hat das Potenzial, nahezu vollständige Datensätze zu klinischen Studien verfügbar zu machen. Ein Vorteil: Die Daten liegen in englischer Sprache vor und sind somit international zugänglich. Anders als die AMNOG-Bewertungen, die sich auf den deutschen Versorgungskontext beziehen, bilden die EU-Bewertungen verschiedene Fragestellungen (PICO [patient, intervention, comparator, outcome]) verschiedener Versorgungskontexte ab. Noch offen ist, ob die pharmazeutischen Hersteller die Veröffentlichung von Daten als Betriebs- und Geschäftsgeheimnisse blockieren dürfen. Eine solche Zensur würde die angestrebte Transparenz in der medizinischen Forschung untergraben.

Ein Beispiel für nicht nutzbare Daten durch Zensur liefert die Bewertung von Palbociclib, die zeitgleich durch das britische National Institute for Health and Care Excellence (NICE) und das IQWiG erfolgte (siehe Abbildung 4): Aufgrund der Dossievorgaben des G-BA konnten im AMNOG-Verfahren Daten verwendet werden, die der Hersteller in der Bewertung des NICE zensiert hatte. Die Bewertungen unterschieden sich insofern, dass NICE trotz der Unsicherheit von einem möglichen Überlebensvorteil ausging, während die Bewertung des IQWiG diesen Vorteil nicht zeigte.

3 Kaiser T. Der „allgemein anerkannte Stand der medizinischen Erkenntnisse“: Die Wissensgrundlagen der Regelbildung in der gesetzlichen Krankenversicherung. Vortrag bei der Tagung aus Anlass des 75. Geburtstags von Friedhelm Hase, Universität Bremen, 28.3.2025.

4 Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. [GA25-01] Berücksichtigung von Daten aus dem AMNOG-Verfahren in Leitlinien. Köln, 3.7.2025. Verfügbar unter: <https://www.iqwig.de/projekte/ga25-01.html>.

Beispiel für nicht nutzbare Daten aufgrund von Zensurierung durch Hersteller : Palbociclib

Die britische Health Technology Assessment (HTA) Organisation NICE (National Institute for Health and Care Excellence) durfte in ihrer Bewertung von Palbociclib bei Brustkrebs keine Daten zum Gesamtüberleben einer Studie darstellen. Der Hersteller hatte angegeben, dass die Daten noch nicht zur Verfügung stünden, da noch nicht genügend Ereignisse aufgetreten seien, und schwärzte die entsprechenden Passagen. Aufgrund der Ergebnisse einer zweiten Studie ging NICE von einem möglichen Überlebensvorteil unter Palbociclib aus. Etwa zeitgleich wurde Palbociclib im Rahmen der frühen Nutzenbewertung nach dem AMNOG in Deutschland bewertet. Auch hier gab der Hersteller im öffentlichen Teil des Dossiers an, dass die Daten zum Gesamtüberleben noch nicht zur Verfügung stünden. Im Studienbericht, den der Hersteller miteinreichen musste, lagen Daten vor, aus denen die Mortalität berechnet werden konnte. In der Gesamtschau beider Studien ergab sich kein Vorteil für Palbociclib beim Gesamtüberleben.

Abbildung 4: Beispiel Palbociclib

Fazit

- Publikationen als einzige Informationsquelle bergen aufgrund des Outcome Reporting Bias das Risiko von Fehlentscheidungen. Andere öffentliche Quellen wie Zulassungsberichte oder Studienregister liefern zwar ggf. zusätzliche Informationen und Daten, relevante Informationen fehlen teilweise aber auch in diesen Quellen.
- Über das AMNOG-Verfahren ist eine vollständige und transparente Datenbasis verfügbar. Dennoch nutzt die Fachwelt diese Quelle bislang nicht regelhaft, etwa bei der Erstellung medizinischer Leitlinien.
- Informationsquellen wie Patientenleitlinien oder Gesundheitsinformationen, die sich auf Behandlungsleitlinien stützen, können nur eingeschränkte Informationen bieten, wenn die Behandlungsleitlinien öffentlich verfügbare Daten nicht berücksichtigen.
- Wer die Daten aus dem AMNOG-Verfahren ignoriert, trifft Entscheidungen ohne Berücksichtigung der gesamten Datenbasis. Die umfassenden Daten des AMNOG-Verfahrens unterstützen dabei, informierte Entscheidungen zu treffen.
- Die transparente Veröffentlichung von Studienergebnissen muss auch für EU-HTA gelten. Andernfalls ist zu befürchten, dass die Erfolge des AMNOG verwässert werden.

Interessenkonflikte

Die Autorin und die Autoren geben an, keine Interessenkonflikte zu haben.

Dr. rer. nat. Cornelia Rüdig

Dr. med. Philip Kranz

Jens Flintrop

Dr. med. Thomas Kaiser