



Dr. PH Dipl.-Stat. Carsten Schwenke
SCO:SSiS Statistical Consulting, Berlin

Implikation von Datenschnitten und Zensierung in klinischen Studien auf die Bewertbarkeit von Arzneimitteln

>> Mit dem Inkrafttreten des Arzneimittelmarktneuordnungsgesetzes (AMNOG) am 1. Januar 2011 wurde in Deutschland das Nutzenbewertungsverfahren neuer Arzneimittel gemäß §35a des SGB V eingeführt. Seitdem wurde in über 500 Verfahren vom Gemeinsamen Bundesausschuss (G-BA) der Zusatznutzen neuer Arzneimittel anhand der vom pharmazeutischen Hersteller vorgelegten Evidenz bewertet und mit dem Beschluss der Preisverhandlung zwischen dem pharmazeutischen Hersteller und dem Spitzenverband der gesetzlichen Krankenkassen zugeführt. Die überwiegende Zahl an Bewertungen erfolgte dabei im Bereich der Onkologie. Die Verfahren werden transparent auf der G-BA-Homepage für die Öffentlichkeit verfügbar gemacht (www.g-ba.de).

Gerade in der Nutzenbewertung ergeben sich immer wieder neue methodische Fragestellungen, die ganz erheblichen Einfluss auf die Bewertung und die Bewertbarkeit der Evidenz für neue Arzneimittel haben. Dazu gehören unter anderem die Verfügbarkeit von Ergebnissen aus randomisierten klinischen Studien verbunden mit der Frage, zu welchem Zeitpunkt Daten in die statistischen Analysen eingeflossen sind und ob diese Zeitpunkte a priori oder post-hoc definiert waren, ob sie z.B. von Zulassungsbehörden wie der europäischen (European Medicines Agency, EMA) oder der amerikanischen (Food and Drug Agency, FDA) im Zulassungsprozess angefordert wurden und ob dabei alle, oder nur ein Teil der patientenrelevanten Endpunkte analysiert wurden. Ein zweiter großer Aspekt in den erforderlichen Ereigniszeitanalysen ist die Bewertbarkeit der Ergebnisse im Lichte von Zensierungen, d.h. der Frage, wie viele Ereignisse beobachtet sein sollten, damit eine Bewertung möglich ist.

Anlass für diesen Kommentar ist die Grundidee, ein gemeinsames Fundament und Basiswissen zu initialisieren und den Diskurs anzustoßen, welche Bedeutung Fachbegriffe wie „prospektiv“ versus „post-hoc“ oder auch die „Zensierung“ im Rahmen der Nutzenbewertung haben. Im Folgenden soll eruiert werden, welche Konstellationen zu Datenschnitten und Zensierungen auftreten können, wie der G-BA damit in Verfahren in neuerer Zeit umgegangen ist und was sich daraus als Erkenntnis gewinnen lässt.

Interpretation verschiedener Datenschnitte in laufenden Studien

Es kommt immer wieder vor, dass statistische Analysen vor Ende der Gesamtstudienlaufzeit durchgeführt werden. Häufigste Gründe sind Interimsanalysen, d.h. Zwischenanalysen für den primären Endpunkt, oder präliminäre Analysen, d.h. finale Analysen für den primären Endpunkt zu einem Zeitpunkt, zu dem z.B. noch keine ausreichende Anzahl Ereignisse für sekundäre Endpunkte beobachtet wurden. Insbesondere in der Onkologie wird ein Datenschnitt prospektiv, in der Regel über eine erforderliche Anzahl an Ereignissen, die eingetreten sein müssen, festgelegt, zu dem die primäre Analyse durchgeführt wird. In der Studie werden aber weiterhin Daten insbe-

sondere für sekundäre Endpunkte für spätere Analysen erhoben. Zum Zeitpunkt der präliminären Analyse des primären Endpunkts, z. B. progressionsfreies Überleben, ist die notwendige Anzahl Ereignisse für den sekundären Endpunkt Gesamtüberleben für eine robuste Analyse häufig noch nicht eingetreten. Im Zulassungsprozess werden dann zunächst die Ergebnisse zum Datenschnitt für den primären Endpunkt zur Erstellung des klinischen Studienreports eingereicht und im Laufe des Verfahrens Analysen zu späteren Datenschnitten nachgereicht. In der Nutzenbewertung nach Zulassung stellt sich die Frage, welcher Datenschnitt für die Bewertung eines Arzneimittels herangezogen werden soll. Ein Datenschnitt meint dabei den Zeitpunkt, bis zu dem Daten in die statistische Analyse eingehen. Man unterscheidet dabei zwischen a priori definierten und post-hoc definierten Datenschnitten, zum einem im Zulassungsprozess, zum anderen im Rahmen der Nutzenbewertungsverfahren. Im Zulassungsprozess ist folgende Definition gebräuchlich: Prospektiv bzw. a priori geplante Datenschnitte sind diejenigen, die im Studienprotokoll detailliert beschrieben sind und für Interimsanalysen, die primäre Analyse für die Darstellung im klinischen Studienbericht und – bei über den Zeitpunkt der Erstellung des Studienberichts hinausgehenden Beobachtungszeiten – finale Analyse zum Abschluss der gesamten Studie.

Die im Zulassungsprozess verwendete Definition der Begriffe a priori und post-hoc sind nicht einfach auf die Nutzenbewertung übertragbar. Für das Nutzendossier ist der Hersteller angehalten, die statistische Analyse durchzuführen, wie sie prospektiv im statistischen Analyseplan der Studie geplant war. Gemäß der Verfahrensordnung des G-BA sind weitere Analysen durchzuführen, die unabhängig von den Ergebnissen der Analysen vom G-BA verlangt werden. Als post-hoc sollten in der Nutzenbewertung nur Analysen, Endpunkte und Subgruppen bezeichnet werden, die nachträglich durch den Hersteller initiiert und in Kenntnis der durchgeführten Analysen definiert und durchgeführt werden.

Anders als im Zulassungsprozess geht es in der Nutzenbewertung bei weiteren Datenschnitten um eine kontinuierlich wachsende Datenbasis aus laufenden Studien, die zu einem Informationsgewinn einer besseren Abschätzbarkeit eines möglichen Zusatznutzens führt. Als Beispiel sei hier der vom Cochrane-Institut verwendete Ansatz der kumulativen Meta-Analyse genannt, bei der Meta-Analysen immer wieder aktualisiert werden, um Therapieentscheidungen auf möglichst aktuellen Informationen basieren zu können. Clarke et al. (Clarke et al. 2014) kommentieren dazu: „This large, unique collection of cumulative meta-analyses highlights how a review of the existing evidence might have helped researchers, practitioners, patients and funders make more informed decisions and choices about new trials over decades of research“. Durch spätere Datenschnitte erhöht sich in der Regel die Datenbasis und damit die Ergebnissicherheit.

Ein Beispiel ist die Bewertung von Ribociclib in Kombination mit Fulvestrant in der Indikation Mammakarzinom. In der Bewertung 2019 lag eine Interimsanalyse mit numerischem, aber nicht stati-

stisch signifikantem Vorteil im Gesamtüberleben vor (G-BA 2019e). Zur Wiedereinreichung nach Fristablauf des Beschlusses (März 2020) lag ein weiterer geplanter Datenschnitt vor. Hier konnte die statistische Signifikanz gezeigt werden (G-BA 2020a). Auch Datenschnitte, die von den Zulassungsbehörden angefordert werden, werden generell berücksichtigt, wie z.B. im Verfahren zu Lenvatinib im fortgeschrittenen Nierenzellkarzinom (G-BA 2017).

Dementsprechend sollten in der Nutzenbewertung Datenschnitte herangezogen werden, wenn der einzige Unterschied zu früheren Datenschnitten die Länge der Beobachtungszeit ist, sich aber an den Analysemethoden, den Endpunkten und der Darstellung der Ergebnisse nichts ändert.

Allen Datenschnitten gemein ist der erforderliche Aufwand. So wird die Datenbank zum Datum des Datenschnitts geschlossen, die dann in der Datenbank enthaltenen Daten auf Plausibilität und Fehler geprüft, gegebenenfalls korrigiert und dann der statistischen Analyse zugeführt. Unter Berücksichtigung des Aufwandes sollte ein Datenschnitt wohlüberlegt sein. Insbesondere bei unverblindeten Studien ist im Falle ungeplanter Datenschnitte ohne Anforderung der Zulassungsbehörden zu prüfen, welche Relevanz der Datenschnitt für die Nutzenbewertung hat.

So heißt es in der Formatvorlage zum Modul 4: „Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich.“ (G-BA 2019a).

Der Hinweis in der Verfahrensordnung, dass in der Regel nur vorgeplante und von den Zulassungsbehörden angeforderte Daten-

schnitte zu berichten seien, lässt den Schluss zu, dass auch andere Datenschnitte relevant sein können. So wurde im Fall Pomalidomid eine Interimsanalyse, die primäre Analyse sowie ein nicht geplanter Datenschnitt etwa 6 Monate nach der primären Analyse für die Herleitung des Zusatznutzens herangezogen (G-BA 2016). Im Verfahren zu Apalutamid wurde hingegen ein ungeplanter Datenschnitt, der mit der Stellungnahme eingereicht wurde, im Beschluss nicht herangezogen (G-BA 2019f). Hier verwies der G-BA auf die Unsicherheit in der Validität der Ergebnisse und berücksichtigte den signifikanten Vorteil im Gesamtüberleben für die Herleitung des Zusatznutzens nicht.

Unklar bleibt, nach welchen Kriterien relevante Datenschnitte zu identifizieren sind. Zudem ist es in der Praxis eine Herausforderung, auch bei Datenschnitten für einzelne Endpunkte, in der Onkologie in der Regel das Gesamtüberleben, aber auch bei Updates zu Sicherheitsendpunkten, auch alle Analysen für alle weiteren relevanten Endpunkte durchzuführen. Hier gilt es nachzuweisen, welche Endpunkte überhaupt noch erhoben werden. So sind bei späten Datenschnitten viele Patienten außerhalb des Beobachtungsfensters für z.B. unerwünschte Ereignisse und patientenberichtete Endpunkte zu Symptomatik und Lebensqualität. Der Stand der Information ändert sich mit späteren Datenschnitten nicht mehr. Das Gesamtüberleben hingegen kann auch nach Studienende bzw. nach Teilnahmeende des Patienten in Erfahrung gebracht werden. In Situationen, in denen das sogenannte Data Cleaning, d.h. die Qualitätskontrolle der eingegebenen Daten und die Bereinigung von Fehlern in der Datenbank, nur für den Endpunkt durchgeführt wird, führen in der Praxis zu Problemen, wenn Analysen für alle Endpunkte verlangt werden.

Eine höhere Anzahl relevanter Datenschnitte führt dabei nicht zwangsläufig zu einer einfacheren Bewertbarkeit der Ergebnisse. Liegen z.B. in onkologischen Studien Ergebnisse über viele Daten-

Literatur

- Clarke M, Brice A, Chalmers I (2014) Accumulating Research: A Systematic Account of How Cumulative Meta-Analyses Would Have Provided Knowledge, Improved Health, Reduced Harm and Saved Resources. *PLoS ONE* 9(7): e102670. doi:10.1371/journal.pone.0102670
- Gemeinsamer Bundesausschuss (2015): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Afatinib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/170/ (abgerufen am 26.06.2020)
- Gemeinsamer Bundesausschuss (2016): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Lenvatinib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/194/ (abgerufen am 26.06.2020)
- Gemeinsamer Bundesausschuss (2017): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Lenvatinib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/261/ (abgerufen am 26.06.2020)
- Gemeinsamer Bundesausschuss (2019a): Dossier zur Nutzenbewertung gemäß § 35a SGB V Modul 4, Dokumentenvorlage, Version vom 21.02.2019. Verfügbar unter: www.g-ba.de/download/40-268-5599/2019-02-21_VerfO_Aenderung-Anlage-I-II_Kapitel-5_Anlage-II-7_WZ.pdf (abgerufen am 15.05.2020)
- Gemeinsamer Bundesausschuss (2019b): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Abemaciclib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/409/ (abgerufen am 11.09.2020)
- Gemeinsamer Bundesausschuss (2019c): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Dacomitinib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/460/ (abgerufen am 15.05.2020)
- Gemeinsamer Bundesausschuss (2019d): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Rucaparib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/446/ (abgerufen am 26.06.2020)
- Gemeinsamer Bundesausschuss (2019e): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Ribociclib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/430/ (abgerufen am 26.06.2020)
- Gemeinsamer Bundesausschuss (2019f): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Apalutamid, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/437/ (abgerufen am 26.06.2020)
- Gemeinsamer Bundesausschuss (2020a): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Ribociclib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/527/ (abgerufen am 11.09.2020)
- Gemeinsamer Bundesausschuss (2020b): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Abemaciclib, Verfügbar unter: www.g-ba.de/bewertungsverfahren/nutzenbewertung/535/ (abgerufen am 11.09.2020)
- Hecken J. (2017): Erfahrungen aus der frühen Nutzenbewertung onkologischer Arzneimittel - Blickwinkel des G-BA. In: Grandt, D./Schubert, I. (Hrsg.) *Arzneimittelreport 2017* (2017): 248 – 251
- Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2017): *Allgemeine Methoden*. Version 5.0 vom 10.07.2017. Verfügbar unter: www.iqwig.de/de/methoden/methodenpapier.3020.html (abgerufen am 15.05.2020)
- Juhnke, C. (2019): Nutzenbewertung europäisch gestalten?. In: *Monitor Versorgungsforschung* 2019, 04: 46-47.
- Schwenke C., Schwenke S. (2018) Die frühe Nutzenbewertung von Arzneimitteln gemäß § 35a SGB V. In: Pfannstiel M., Jaeckel R., Da-Cruz P. (eds) *Innovative Gesundheitsversorgung und Market Access*. Springer Gabler, Wiesbaden
- Unkel S, Amiri M, Benda N, Beyersmann J, Knoerzer D, Kupas K, Langer F, Leverkus F, Loos A, Ose C, Proctor T, Schmoor C, Schwenke C, Skipka G, Unnebrink K, Voss F, Friede T (2018): On estimands and the analysis of adverse events in the presence of varying follow-up times within the benefit assessment of therapies. *Pharmaceutical Statistics*. 18: 166-183
- Zwiener I, Blettner M, Hommel G: Überlebenszeitanalyse. *Deutsches Ärzteblatt*, Jg. 108(10): 163-169

schnitte vor, die zwar gleichgerichtet, nicht aber zwangsläufig zu jedem Datenschnitt statistisch signifikant sind, so kann das zu einer Abstufung der Ergebnissicherheit führen (Schwenke 2018).

Zensierung in Ereigniszeitanalysen im Kontext der Nutzenbewertung

Im Kontext der Nutzenbewertung spielen Ereigniszeitanalysen immer dann eine zentrale Rolle, wenn die Beobachtungszeiten in den untersuchten Behandlungsgruppen unterschiedlich sind (Unkel et al. 2018). Eine Analyse ohne Berücksichtigung der unterschiedlichen Beobachtungszeiten würde zu einer Verzerrung zugunsten der Behandlung mit der kürzeren Beobachtungszeit führen, da die Patienten weniger lange unter Behandlung und somit auch unter Risiko stehen, ein unerwünschtes Ereignis zu berichten. In solchen Konstellationen hat sich im Nutzenbewertungsverfahren die Ereigniszeitanalyse etabliert und wurde vom G-BA in die neue Formatvorlage für das Nutzendossier aufgenommen (G-BA 2019a).

Gebräuchlich ist dabei die Darstellung mittels Kaplan-Meier Abbildung, aus der die mediane Zeit bis zum Ereignis direkt ablesbar ist (Zwiener et al. 2011). Zudem wird in der Regel die Hazard Ratio für den Vergleich der Ereigniszeitenkurven geschätzt (Zwiener et al. 2011). Patienten ohne Ereignis werden dabei zum letzten Zeitpunkt, zu dem Daten für den relevanten Endpunkt vorliegen, zensiert. Die Zensierung kann zwei Gründe haben, entweder beendet der Patient die Studie vorzeitig oder der weiterhin beobachtete Patient hatte bis zum Datenschnitt kein Ereignis. Bei späteren Datenschnitten ändert sich damit die Datenlage, da nur für vorzeitig abbrechende Patienten keine weiteren Daten erhoben werden.

Spätere Datenschnitte führen damit in der Regel zu einer höheren Aussagesicherheit, solange es zu keinem Cross-over, d.h. keinem Übergang der Patienten von der Kontroll- zur Verumtherapie gekommen ist. Als Beispiel kann hier das Verfahren zu Abemaciclib dienen. In einem ersten Verfahren in 2019 (G-BA 2019b) wurde in einer der bewerteten Teilpopulationen mit Ereignisraten von 20% in beiden Armen kein signifikantes Ergebnis für das Gesamtüberleben erreicht. In der Neubewertung nach Befristung im Jahr 2020 wurde für den späteren Datenschnitt für die Verum- und Kontrollgruppe Ereignisraten von 46% bzw. 67% und eine signifikante Hazard Ratio berichtet mit der Konsequenz eines Zusatznutzens (G-BA 2020b).

Für Ereigniszeitanalysen wird vom G-BA neben der Hazard Ratio zudem die Differenz der medianen Zeit bis zum Ereignis in der Bewertung herangezogen. Auch in der medizinischen Fachwelt ist diese Differenz wichtig, da diese auch dem Laien verständlich erklärt werden kann, wohingegen die Hazard Ratio als relatives Effektmaß ein methodisches Verständnis der statistischen Methode voraussetzt. Kann der Median noch nicht geschätzt werden, so können alternativ das 25%-Quantil oder Ereignisraten zu bestimmten Zeitpunkten wie z.B. zu Jahr 1 oder 2 herangezogen werden.

Wie Prof. Hecken im Arzneimittelreport 2017 ausführt, sind wie im Beispiel Afatinib als Erstlinientherapie frühe Datenschnitte problematisch, spätere Datenschnitte deutlich aussagekräftiger und führten im Beispiel Afatinib in der Teilpopulation der Patienten mit EGFR-Mutation Del19 zu einem erheblichen Zusatznutzen (Hecken 2017; G-BA 2015).

Zensierung und der Einfluss auf die Ableitung des Zusatznutzens

Bei Endpunkten mit unterschiedlichen Beobachtungsdauern in

den Behandlungsarmen leitet das Institut für Qualität im Gesundheitswesen (IQWiG) den Zusatznutzen in der Regel auf Basis des relativen Effektmaßes Hazard Ratio ab (IQWiG 2017). Der G-BA hingegen verwendet zusätzlich den absoluten Behandlungseffekt in Form der Differenz der medianen Zeiten bis zum Ereignis. Bei letzterer spielt der Zeitpunkt der Zensierung eine entscheidende Rolle für die Präzision der Schätzung der Mediane. Sowohl bei der Hazard Ratio als auch bei der Schätzung der Mediane führt die frühe Zensierung zu einer Erhöhung der Unschärfe. Der Median wird in diesem Fall bestimmt durch Ereignisse der wenigen noch unter Beobachtung stehenden Patienten. Aufgrund der erwartbaren Variabilität kann dieses Ereignis zu einem kurzen oder längeren Median führen, je nachdem, ob in den verbliebenen Patienten das nächste Ereignis zufällig früh oder später eintritt. Dieser Effekt ist insbesondere deshalb relevant, da kein expliziter statistischer Test für die Differenz der Mediane durchgeführt wird, sondern die Schätzer der Mediane ohne ein Maß der Variabilität berücksichtigt werden.

Wird der Median in mindestens einer Behandlungsgruppe nicht erreicht, so kann keine Aussage zur Differenz der Mediane getroffen werden. Sind in beiden Gruppen die Mediane erreicht und liegen sehr viele Zensierungen vor dem Zeitpunkt des Medians vor, so liegen zur Bestimmung des Medians Daten von wenigen Patienten unter Risiko vor. Das führt zu einer großen Unschärfe und kann einen Unterschied von mehreren Monaten ausmachen.

Im Fall der Nutzenbewertung von Dacomitinib im Jahr 2019 hatte der G-BA im Beschluss dargelegt, dass zur Bewertung des Zusatznutzens für das Gesamtüberleben zwar eine signifikante Hazard Ratio erreicht wurde und auch die mediane Zeit bis zum Ereignis im Interventionsarm um 7,3 Monate verlängert sei, aber eine hohe Anzahl an Zensierungen im Bereich des Medians auffielen, die keine Quantifizierung des Zusatznutzens erlaubten (G-BA 2019c). Zu berücksichtigen ist dabei, dass zum nachgereichten Datenschnitt vom 19.05.2019 der überwiegende Anteil Patienten mit Zensierung protokollgemäß den finalen Beobachtungszeitpunkt Woche 48 erreicht hatten und die Studie protokollgemäß beendet hatten. Patienten, die die Woche 48 erreicht hatten, wurden demgemäß zum Datum der Beendigung der Studie zensiert. Der Grund der Zensierung war demnach ein administrativer, der einen informativen Effekt der Zensierung auf die Schätzung des Behandlungseffekts unwahrscheinlich erscheinen lässt. Spätere Datenschnitte lösen hier den Punkt der Zensierung nicht.

Eine ähnliche Konstellation findet sich in der schon genannten Wiedervorlage von Ribociclib mit Fulvestrant (G-BA 2020a). Hier wurde das mediane Gesamtüberleben bisher nur in der Kontrollgruppe erreicht und es zeigen sich sehr viele Zensierungen rund um den Median. In der Interventionsgruppe wurde der Median noch nicht erreicht bei ebenfalls vielen Zensierungen. Hier wurde zum einen der signifikante Vorteil in der Hazard Ratio für die Herleitung des Zusatznutzens herangezogen als auch das Verzerrungspotenzial als niedrig eingestuft, obwohl auch hier in der Kontrollgruppe bei Erreichen des Medians weniger als 10% der Patienten noch unter Risiko standen. Im schon genannten Fall von Apalutamid im Prostatakarzinom wurde der im Dossier gezeigte Datenschnitt zum Gesamtüberleben für die Herleitung des Zusatznutzens herangezogen, obwohl der Median ebenfalls nur im Kontrollarm bereits erreicht war, bei vielen Zensierungen vor dem Median in beiden Behandlungsgruppen (G-BA 2019f). Da bei Erreichen des Medians nur noch 2 von 401 Patienten unter Risiko standen, wurde die Bewertung des Medians als fragwürdig eingestuft.

Diskussion

Zwei wichtige Aspekte zur Beurteilung des Zusatznutzens eines Arzneimittels in der Nutzenbewertung nach §35a SGB V werfen immer wieder Fragen auf, zum einen die Begrifflichkeiten prospektiv vs. post-hoc, zum anderen die Fragen des Einflusses von Zensierung auf die Bewertbarkeit von Ereigniszeitanalysen.

Der Konflikt bezüglich prospektiv versus post-hoc ist insbesondere Thema im Rahmen der Prüfung auf Relevanz neuer Datenschnitte in laufenden Studien. Laut Verfahrensordnung sollen prospektiv geplante Datenschnitte im Nutzendossier präsentiert werden und es gebe Limitationen, wenn ein nicht geplanter Datenschnitt präsentiert wird. Hier sei im Einzelfall zu begründen, warum ein Datenschnitt herangezogen oder eben nicht herangezogen wird. Für die Einschätzung der Relevanz von Datenschnitten ist von außerordentlichem Interesse, wie die Begrifflichkeiten prospektiv und post-hoc definiert sind. Der Begriff „post-hoc“ sollte in der Nutzenbewertung nicht auf Datenschnitte angewendet werden, wenn der einzige Unterschied zu früheren Datenschnitten die Länge der Beobachtungszeit ist, sich aber an den Analysemethoden, den Endpunkten und der Darstellung der Ergebnisse nichts ändert.

Generell ist es sinnvoll, das Thema Datenschnitte im Rahmen der (frühen) Beratung als Frage an den G-BA zu adressieren, um Klarheit über die Relevanz insbesondere ungeplanter Datenschnitte zu erlangen, soweit dieses zum Zeitpunkt der Beratung schon absehbar ist.

Je länger man wartet, um so sicherer werden im Regelfall die Aussagen, da die Anzahl der Ereignisse, die in die Analyse eingehen, größer wird und die Anzahl der Zensierungen andererseits geringer.

Verknüpft mit dem Thema Datenschnitte ist die Frage der Reife der Daten für die Analysen und damit einhergehend die Frage des Anteils Patienten in Ereigniszeitanalysen, die noch kein Ereignis erfahren haben. Auch bei einer geringeren Anzahl Ereignisse kann sich bereits ein deutlicher Behandlungseffekt zeigen, auch wenn der Median für die Zeit bis zum Ereignis noch nicht erreicht ist. So sind viele Zensierungen rund um den Median kein Grund, eine Bewertung nicht durchzuführen oder aber den Zusatznutzen nicht zu quantifizieren. In solchen Fällen kann der relative Behandlungseffekt in Form der Hazard Ratio herangezogen werden. Offen bleibt, was in

Abwesenheit einer Differenz der medianen Zeiten bis zum Ereignis herangezogen werden kann, um eine Bewertung des absoluten Behandlungseffekts zu erzielen. Alternativen sind das 25% Quantil oder aber die Ereignisraten zu sinnvollen Zeitpunkten, sinnvoll definiert als Zeitpunkte, zu denen noch ausreichend viele Patienten nicht zensiert sind.

Auch auf europäischer Ebene gibt es bei der Nutzenbewertung viele Fragezeichen, auch wie in Zukunft mit den Themen Datenschnitte und Zensierungen umgegangen wird (Juhnke 2019).

Zusammenfassend sollten sich alle Parteien auf eine gemeinsame Begrifflichkeit einigen, anhand derer die Erfordernisse zur Erstellung eines Nutzendossiers standardisiert und die Einheitlichkeit der Nutzenbewertung von Arzneimitteln gewährleistet ist. <<

Zitationshinweis

Schwenke, C.: „Implikation von Datenschnitten und Zensierung in klinischen Studien auf die Bewertbarkeit von Arzneimitteln“ in: „Monitor Versorgungsforschung“ (03/21), S. 30-33. <http://doi.org/10.24945/MVF.03.21.1866-0533.2315>

Korrespondenzadresse

Dr. Carsten Schwenke
SC0:SSiS
Karmeliterweg 42
13465 Berlin
Tel. +49 30 - 60409712
carsten.schwenke@scossis.de

Anmerkungen

Dr. Carsten Schwenke ist für pharmazeutische, biotechnologische und medizintechnische Unternehmen sowie Universitäten als freiberuflicher Statistiker in der klinischen Entwicklung neuer Therapien und Impfstoffe sowie in Nutzenbewertungsverfahren auf deutscher und europäischer Ebene tätig. Die Analyse erfolgte mit finanzieller Unterstützung von Pfizer Pharma GmbH.

Evidentes Vorbild für Regionen

>> „Gesundheit und sozialer Status sind eng miteinander verknüpft. Gesundheitsversorgung ist daher nicht nur eine medizinische Frage, sondern muss umfassender gedacht werden“, erklärt Dr. Dirk Heinrich, Bundesvorsitzender des Virchowbundes anlässlich der Publikation von Forschungsergebnissen zum innovativen Versorgungsprojekt „INVEST Billstedt/Horn“. Der Verband der niedergelassenen Ärzte (Virchowbund) will das Modellprojekt nun so rasch wie möglich in die Regelversorgung überführen. „In Hamburg-Billstedt konnten wir zeigen, dass eine populationsorientierte, niedrigschwellige Versorgung möglich ist, wenn die relevanten Akteure sich engagieren und kooperieren“, so Heinrich weiter. Das sei eine Form der Vernetzung, die Patienten hilft und Ärzte entlastet. Aus dem Prototyp sollte nach Heinrichs Meinung „jetzt ein Vorbild für andere Städte und Regionen“ werden.

Die Autoren der nun veröffentlichten Evaluation des Hamburg Center for Health Economics (HCHE) wiesen nach, dass INVEST u. a. den Zugang zur Versorgung verbessert und die Zufriedenheit der Patienten, Leistungs- und Kooperationspartner erhöht hat.

Der Virchowbund hatte 2016 die „Gesundheit für Billstedt/Horn UG“ gemeinsam mit u. a. dem Ärztenetz Billstedt-Horn e.V. und der SKH Stadtteilklinik Hamburg GmbH gegründet. Das daraus entstandene Modellprojekt „INVEST“ wurde vom Innovationsfonds gefördert und vom HCHE an der Universität Hamburg wissenschaftlich begleitet. Konsortialpartner sind OptiMedis AG, AOK Rheinland/Hamburg, BARMER, DAK-Gesundheit, Universitätsklinikum Hamburg-Eppendorf (Universität Hamburg) sowie das Hamburg Center for Health Economics (Universität Hamburg) und die connected-health.eu GmbH. <<

Wichtige Aufträge an das IQWiG

>> Das IQWiG wurde vom G-BA beauftragt, ein Konzept zu entwickeln, um das noch relativ neue Instrument der anwendungsbegleitenden Datenerhebung (abd) zielgerichtet auch für Arzneimittel einer Wirkstoffklasse einsetzen zu können. Zum anderen wird das IQWiG mit der Nutzenbewertung für die drei Wirkstoffe Clopidogrel (laut G-BA „inhaltliches „Neuland“) sowie Prasugrel und Ticagrelor (jeweils „klassische“ Nutzenbewertungen) beauftragt. Angesichts neuer Studieninformationen soll das IQWiG jedoch auch das Binnenverhältnis dieser drei Wirkstoffe klären. <<