

Praktische Aspekte bei der Anwendung von Propensity Scores

Practical aspects in the application of propensity scores

Anika Großhennig¹

Ralf Bender²

Friedhelm Leverkus³

1 Institut für Biometrie,
Medizinische Hochschule
Hannover, Hannover,
Deutschland

2 Institut für Qualität und
Wirtschaftlichkeit im
Gesundheitswesen, Köln,
Deutschland

3 Pfizer, Berlin, Deutschland

Einleitung

Die Arbeitsgruppe „Therapeutische Forschung“ der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS) bietet eine Plattform zur Diskussion aktueller biometrischer Fragestellungen in der klinischen Forschung, insbesondere hinsichtlich der Nutzenbewertung [1]. Im Rahmen von regelmäßigen Workshops der Arbeitsgruppe kommen Expert:innen aus dem universitären Bereich, der Industrie und Institutionen des Gesundheitswesens zusammen, um aktuelle biometrische Herausforderungen und entsprechende Lösungsansätze aus ihren unterschiedlichen Perspektiven vorzustellen und zu diskutieren.

In dem Workshop im November 2022 haben wir uns mit dem Thema „Praktische Aspekte bei der Anwendung von Propensity Scores“ auseinandergesetzt, das für die aktuellen anwendungsbegleitenden Datenerhebungen (AbDs) bei neuen Arzneimitteln von großer Relevanz ist. Das starke Interesse am Thema, das sich für uns zum einen an der großen Anzahl an Workshopteilnehmenden (mehr als 160 Teilnehmende aus der pharmazeutischen Industrie, regulatorischen und universitären Einrichtungen) und an der angeregten Diskussion, gezeigt hat, haben wir zum Anlass genommen, die wichtigsten Inhalte des Workshops durch die Referent:innen für diese Themenausgabe zusammenzutragen. Alle vier Beiträge der Themenausgabe haben vor der Veröffentlichung ein unabhängiges Peer-Review-Verfahren durchlaufen.

Übersicht zur Themenausgabe

Die vorliegende Themenausgabe folgt der Struktur des Workshops. So fassen zunächst Kaiser & Vervölgyi [2] in ihrem Beitrag die wichtigsten Grundlagen zu AbDs für die Nutzenbewertung von neuen Arzneimitteln in Deutschland aus HTA-Sicht zusammen. Die Basis für AbDs ist das Gesetz für mehr Sicherheit in der Arzneimittelversorgung (GSAV) von 2019 [3]. Demnach kann der Gemeinsame Bundesausschuss (G-BA) vom pharmazeutischen Unternehmer verlangen, dass in der täglichen Praxis Daten gesammelt und ausgewertet werden, wenn für die Bewertung des Nutzens und Zusatznutzens neuer Arzneimittel keine ausreichend aussagekräftigen Daten vorliegen. Zu diesem Zweck sollen AbDs durchgeführt und – über die bereits bekannten Studiendaten aus der Zulassung hinaus – dazu beitragen, mehr Informationen über den Nutzen und Schaden des neuen Arzneimittels zu gewinnen, um eine Nutzenbewertung zu ermöglichen. Im Gegensatz zu den üblicherweise randomisierten Zulassungsstudien sind für AbDs insbesondere nicht-randomisierte Studien vorgesehen. Kaiser & Vervölgyi [2] weisen darauf hin, dass für die Durchführung von AbDs die Erstellung ausführlicher Studienprotokolle und statistischer Analysepläne (SAPs) unbedingt erforderlich sind. In diesen Dokumenten, die entsprechend auch für randomisierte Studien erforderlich sind, müssen zentrale methodische Aspekte wie die Adjustierung von Confoundern detailliert beschrieben werden.

Die zwingende Notwendigkeit einer adäquaten Adjustierung aller potenziellen Confounder bei AbDs macht es erforderlich, sich mit Methoden zur Confounderkontrolle, und zwar insbesondere mit Verfahren auf Basis von Propensity Scores, auseinanderzusetzen. Im Workshop wurde dann zunächst auf Basis der Arbeit von Kuß et al. [4] eine Einführung in die methodischen Grundlagen von Propensity Scores zur Verwendung in nicht randomisierten Studien gegeben.

Im zweiten Beitrag dieser Themenausgabe befasst sich Mathes [5] mit den Anforderungen an die Datenerfassung von versorgungsnahen Daten, um damit Nutzenbewertungen durchzuführen. Insbesondere diskutiert Mathes [5] die Anwendung von Propensity-Score-Verfahren in Registern und weist darauf hin, dass Propensity-Score-Verfahren deutlich höhere Fallzahlen benötigen, als i.d.R. in den verfügbaren Daten vorhanden sind. Dies lässt sich aus seiner Sicht nur schwer auflösen, denn auch die Zusammenführung mehrerer Datenquellen ist aufgrund vieler unterschiedlicher Strukturen regelhaft nicht praktikabel, so dass er hier als einen Ansatz die Weiterentwicklung von Propensity-Score-Verfahren für kleinere Fallzahlen vorschlägt. Zusammenfassend ist Mathes [5] skeptisch bezüglich der Anwendung von Registern für die Nutzenbewertung, weil zum einen in Deutschland die notwendige Datenqualität aktuell nicht erreicht werden kann und zum anderen mit diesen Verfahren nach wie vor die interessierenden Therapieeffekte im Vergleich zu randomisierten Studien nur sehr unzuverlässig geschätzt werden können. Als vielversprechende Alternative nennt Mathes [5] die Durchführung einer pragmatischen, registerbasierten Studie mit Randomisierung.

Abschließend befassen sich Basic [6] sowie Kuß & Strobel [7] mit methodischen Aspekten und Erweiterungen von Propensity-Score-Verfahren. Basic [6] diskutiert in seinem Beitrag unterschiedliche Gewichtungsansätze zur Schätzung verschiedener Behandlungseffekte in nicht randomisierten Studien auf Basis von Propensity Scores. Er weist darauf hin, dass insbesondere eine korrektive Spezifikation des Propensity-Score-Modells, eine ausreichende Überlappung zwischen den zu vergleichenden Gruppen sowie die Balanciertheit der einzelnen Confounder wichtige Kriterien für die Bewertung der Ergebnisse von Propensity-Score-Analysen sind.

Ein anschauliches und konkretes Anwendungsbeispiel, wie Überlappung und Balanciertheit als Maß für die Güte eines Propensity-Score-Modells, überprüft werden können, stellen Kuss & Strobel [7] in ihrem Beitrag für diese Themenausgabe vor. An einer Beispielstudie aus der Herzchirurgie leiten Sie pragmatische Empfehlungen für die konkrete Anwendung von Propensity-Score-Verfahren und die Überprüfung der Annahmen zur Balanciertheit der Confounder und dem Overlap des Propensity Score in den Behandlungsgruppen ab. Kuss & Strobel [7] weisen hierbei auch noch einmal darauf hin, dass die Behandlungseffekte mittels Propensity-Score-Analysen nur valide geschätzt werden können, wenn die Überprüfung der

zentralen Annahmen hinreichend zufriedenstellende Ergebnisse geliefert hat.

Fazit

Mit der vorliegenden Themenausgabe zum Workshop ist es gelungen, die unterschiedlichen Perspektiven zu einem bedeutenden Thema in der Nutzenbewertung zusammenzutragen und gegenüberzustellen. Die Autoren haben die methodischen Herausforderungen, die es unter anderem bezüglich der Datenerfassung und Datenqualität sowie bezüglich der Auswertung und Interpretation zu lösen gibt, aufgezeigt und diskutiert. Auch wenn der Trend zu beobachten ist, dass Standards und Anforderungen für die Nutzung von Propensity-Score-Verfahren nach und nach etabliert werden, muss sich noch zeigen, ob in den nächsten Jahren AbDs mit Anwendung von Propensity-Score-Methoden oder alternativen Methoden zur Confounderkontrolle für die Nutzenbewertung relevante Zusatzinformationen liefern können.

Anmerkungen

Danksagung

Wir möchten uns hiermit bei allen Autor:innen, Workshopteilnehmenden und den Reviewer:innen für ihre interessanten Beiträge zur Diskussion im Workshop und bei der Manuskripterstellung bedanken.

Interessenkonflikte

Die Autor:innen erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

Literatur

1. Arbeitsgruppe „Therapeutische Forschung“ der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. [Internet]. Verfügbar unter: <https://www.gmds.de/aktivitaeten/medizinische-biometrie/arbeitsgruppenseiten/therapeutische-forschung/>
2. Kaiser T, Vervölgyi V. Anwendungsbegleitende Datenerhebungen für die Nutzenbewertung von Arzneimitteln in Deutschland: Warum, wann und wie? GMS Med Inform Biom Epidemiol. 2024; 20:Doc04. DOI: 10.3205/mibe000260
3. Gesetz für mehr Sicherheit in der Arzneimittelversorgung. Bundesgesetzblatt Teil I. 2019;(30):1202-20.
4. Kuß O, Blettner M, Börgemann J. Propensity Score: An alternative method of analyzing treatment effects. Dtsch Arztebl Int. 2016 Sep 5;113(35-36):597-603. DOI: 10.3238/arztebl.2016.0597
5. Mathes T. Anforderung an die Daten für die Target-Trial-Emulation: Eine Diskussion unter Betrachtung von Patientenregistern. GMS Med Inform Biom Epidemiol. 2024;20:Doc03. DOI: 10.3205/mibe000259

6. Basic E. Alternative Ansätze für Confounder-Adjustierung durch Gewichtung auf der Grundlage des Propensity-Score. GMS Med Inform Biom Epidemiol. 2024;20:Doc02. DOI: 10.3205/mibe000258
7. Kuß O, Strobel A. Gütemaße und Kriterien bei der Anwendung von Propensity Scores. GMS Med Inform Biom Epidemiol. 2024;20:Doc01. DOI: 10.3205/mibe000257

Bitte zitieren als

Großhennig A, Bender R, Leverkus F. Praktische Aspekte bei der Anwendung von Propensity Scores. GMS Med Inform Biom Epidemiol. 2024;20:Doc05. DOI: 10.3205/mibe000261, URN: urn:nbn:de:0183-mibe0002618

Artikel online frei zugänglich unter

<https://doi.org/10.3205/mibe000261>

Veröffentlicht: 05.01.2024

Korrespondenzadresse:

Dr. rer. hum. biol. Anika Großhennig
Medizinische Hochschule Hannover, Institut für Biometrie,
Carl-Neuberg-Str. 1, 30625 Hannover, Deutschland
Grosshennig.Anika@mh-hannover.de

Copyright

©2024 Großhennig et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.