

Anforderung an die Daten für die Target-Trial-Emulation: Eine Diskussion unter Betrachtung von Patientenregistern

Data requirements for target-trial emulation to avoid bias: A discussion considering German patient registries

Abstract

The article first considers basic requirements (e.g. recording of relevant confounders) and frequent sources of bias (e.g. immortal time bias) when using registry and other real-world data as the basis for comparative effectiveness analyses of treatments. The target trial concept is described and estimands relevant to the benefit assessment are discussed. Furthermore, relevant requirements for the data in terms of practical feasibility, adequate emulation, and internal validity, in particular for the implementation of propensity score-based analyses, are reflected. The feasibility is considered against the background of the German register landscape. Finally, the experiences of projects that have attempted to emulate RCT effects using real-world data are presented.

Keywords: causal inference, target trial, propensity scores, patient registries, real-world data, routine practice data collection

Tim Mathes¹

1. Institut für Medizinische Statistik, Universitätsmedizin Göttingen, Göttingen, Deutschland

Zusammenfassung

In diesem Beitrag wird zunächst auf grundlegende Anforderungen (z.B. Erfassung von relevanten Confoundern) und häufige Biasquellen (z.B. Immortal-Time-Bias) bei der Nutzung von Register- und anderen versorgungsnahen Daten als Grundlage für die vergleichende Analyse von Therapien eingegangen. Das Target-Trial-Konzept wird skizziert und die möglichen bzw. für die Nutzenbewertung relevanten Estimands werden diskutiert. Weiterhin werden relevante Anforderungen an die Daten hinsichtlich praktischer Machbarkeit, adäquater Emulation und interner Validität, insbesondere zur Durchführung von Propensity-Score-basierten Analysen, erläutert. Dabei wird die Durchführbarkeit vor dem Hintergrund der deutschen Registerlandschaft betrachtet. Abschließend werden die Erfahrungen aus Projekten, in denen versucht wurde, Effekte aus randomisierten Studien mittels versorgungsnaher Daten zu emulieren, aufgezeigt.

Schlüsselwörter: Kausalschluss, Target-Trial, Propensity Scores, Patientenregister, versorgungsnaher Daten, anwendungsbegleitende Datenerhebung

Einleitung

Auf versorgungsnahen Daten beruhende Studien (NRSvDa) zum Vergleich der Effektivität von Interventionen müssen insbesondere belastbare Evidenz liefern, wenn diese die Ergebnisse aus randomisierten kontrollierten Studien (RCTs) ersetzen sollen, wie es z.B. bei der anwendungsbegleitenden Datenerhebung vorgesehen ist. Als Best-Practice-Ansatz für NRSvDa kann das Konzept der Target-Trial-Emulation angesehen werden [1].

Hierbei wird angestrebt, den Effekt einer hypothetischen RCT mittels Beobachtungsdaten nachzubilden (zu emulieren). Zu diesem Zweck werden Studienprotokoll (z.B. Einschlusskriterien, Behandlungsstrategien) und Analysemethoden dahingehend harmonisiert, dass möglichst der gleiche kausale Effekt geschätzt wird [2]. Eine Grundvoraussetzung für die Nutzenbewertung von medizinischen Maßnahmen mittels NRSvDa ist es, dass hierfür geeignete Daten vorhanden sind. Geeignet bedeutet hierbei insbesondere, dass die Daten es erlauben,

die angestrebte Analysemethode anzuwenden (z.B. Propensity-Score (PS)-Matching) und die Ergebnisse dieser intern valide sind. Die interne Validität bezieht sich auf die Charakteristika der versorgungsnahen Daten (vnDa), z.B. Datenqualität, und die anwendbaren Methoden, wobei sich diese Faktoren gegenseitig bedingen.

Der Beitrag fokussiert auf Register als Datenbasis für die Target-Trial-Emulation, da zum aktuellen Zeitpunkt andere Datenquellen (z.B. Routinedaten der Krankenkassen) nicht geeignet erscheinen, die Anforderung an die Daten im Rahmen der Nutzenbewertungen zu erfüllen [3]. Für diesen Artikel wird dabei ein Register weitgefasst als „eine Sammlung von standardisierten Informationen – für einen oder mehrere Zwecke – über eine Gruppe von Patienten, die eine Erkrankung oder ein Ereignis teilen“ definiert [4].

Im ersten Teil dieser Arbeit werden die für die Nutzenbewertung relevanten Estimands aufgezeigt, da diese maßgeblich für die Anforderungen an die Analysemethoden und die Daten sind. Der zweite Teil der Arbeit widmet sich der Frage, was häufige Risiken für systematische Verzerrung (Bias) von NRSvnDa sowie andere potentielle Ursachen für abweichende Ergebnisse sind. Bei den Biasrisiken wird dabei insbesondere auf jene fokussiert, die für NRSvnDa besonders relevant sind [5]. Im dritten Teil werden geeignete PS-basierte Analysemethoden aufgezeigt. Der vierte Teil skizziert anhand der Deutschen Registerlandschaft mögliche praktische Probleme bei der Durchführung von NRSvnDa. Abschließend wird auf Ergebnisse aus metaepidemiologischen Studien eingegangen, in denen versucht wurde, Ergebnisse von RCTs mittels vnDa zu emulieren.

Target-Trial-Emulation und Implikationen für die Anforderungen an die Daten

Mögliche Estimands

Ein Estimand ist eine systematische Beschreibung des Behandlungseffektes, der zur Beantwortung des Studienziels in der Studie quantifiziert werden soll. Ein Estimand umfasst die Angabe von fünf Aspekten: Population, Behandlung, Endpunkt, Umgang mit Intercurrent-Events nach dem Start der Intervention und zusammenfassendes Effektmaß (z.B. Vergleich der Änderung zum Ausgangswert oder Vergleich der Endwerte) [6]. Auf Basis von vnDa können verschiedene Estimands geschätzt werden. In der Regel kann davon ausgegangen werden, dass im Rahmen der Nutzenbewertung der Average-Treatment-Effect in der Gesamtpopulation (ATE) unter einer Treatment-Policy-Strategie (TPS) bzw. der Intention-To-Treat-Effekt der relevante und somit zu emulierende Estimand ist, da dieser dem Effekt einer RCT unter Verwendung einer Intention-To-Treat-Analyse entspricht [7].

Häufige Verzerrungsquellen und Emulationsunterschiede

Bias aufgrund von Confoundern

Eine Verzerrungsquelle von nicht-randomisierten Studien (NRS) ist die Gefahr von Confounding. Dieses gilt gleichermaßen für NRSvnDa. Bei NRSvnDa können gleich zwei Aspekte dazu führen, dass relevante Confounder nicht berücksichtigt werden: Erstens können wichtige Confounder einfach vergessen werden, z.B. aufgrund von mangelndem Vorwissen. Zweitens, und dies gilt speziell für vorbestehende Datenquellen wie vnDa, können die notwendigen Confounder nicht oder nicht in ausreichender Qualität in der Datenbank/den Datenbanken (z.B. Register) vorliegen. Dementsprechend zeigen Studien, dass bei registerbasierten NRS häufig relevante Confounder nicht berücksichtigt werden, obgleich sich das Studienteam über das Fehlen von relevanten Confoundern bewusst ist [5]. Dies legt nahe, dass sich die Auswahl der Confounder oftmals mehr nach der Verfügbarkeit als nach objektiven inhaltlichen Gesichtspunkten richtet.

Bias aufgrund der Wahl des Studienstartpunkts und Follow-up-Zeitraums

Eine weitere Verzerrungsquelle ist Immortal-Time- oder Selection-Bias, da NRSvnDa im Gegensatz zu prospektiven Studien keinen festen Startzeitpunkt pro Patient in die Betrachtung der NRSvnDa haben. Diese Biasarten entstehen, wenn Studienstart (Zeitpunkt T0), Start der Nachbeobachtung, Beginn der Erfüllung der Einschlusskriterien (z.B. Erkrankungsbeginn) und Behandlungszuteilung nicht übereinstimmen. Hernán et al. unterscheiden hier vier verschiedene Arten der Follow-up-bezogenen Verzerrungsquellen [8]. Zeitpunktbedingte Verzerrungsrisiken – Verzerrungsquellen und mit diesen einhergehende Bias – sind demnach:

- T0 wird (für einige Beobachtungen) nach dem Zeitpunkt der Erfüllung der Einschlusskriterien und Behandlungszuteilung festgesetzt. → Selection-Bias, da nur Patient*innen in die Analyse einbezogen werden, die dauerhaft die Behandlung erhalten haben (Persistent Users).
- T0 wird auf den Zeitpunkt der Erfüllung der Einschlusskriterien, aber nach der Behandlungszuteilung festgesetzt. → Selektion-Bias, da nur Personen in die Analyse eingeschlossen werden, die genau zu T0 die Einschlusskriterien erfüllen.
- T0 wird auf einen Zeitpunkt vor Erfüllung der Einschlusskriterien und Behandlungszuteilung gesetzt. → Immortal-Time-Bias
- T0 wird auf den Zeitpunkt der Erfüllung der Einschlusskriterien, aber vor Behandlungszuteilung gesetzt. → Immortal-Time-Bias

Gleichermaßen wie beim Confounding-Bias können dieses Biasarten durch geeignete Analysen vermieden werden, vorausgesetzt die notwendigen Informationen sind in den Datenquellen vorhanden.

Bias aufgrund von Abweichung von der vorgesehenen Intervention

Aus der TPS ergibt sich zum einen die Anforderung, dass Informationen über die Behandlungsintention vorliegen müssen. In einer RCT ergibt sich dies aus der Zuteilungssequenz. Analog müssten für vnDa Informationen über die ursprüngliche Behandlungsintention vorliegen. Diese Information ist im Normalfall jedoch nicht vorhanden. Bekannt ist nur, ob die Patienten*innen auch wirklich die Behandlung gestartet haben [5]. Falls es zu relevanten Abweichungen zwischen Behandlungsintention und wirklichem Start der Intervention kommt, besteht hier ein Verzerrungsrisiko. Der TPS-Effekt erfordert zudem, dass Intercurrent Events ignoriert werden. Da es sich bei vnDa um Daten aus der Routineversorgung handelt, kann es zu einer Vielzahl von Intercurrent Events, insbesondere non-Adhärenz und Treatment-Switching kommen. In der Praxis finden sich häufig keine diesbezüglich klaren Analysen und somit sind die Estimands nicht eindeutig bestimmbar [5]. Zur Schätzung eines „reinen“ TPS-Effekts muss bei der Analyse streng darauf geachtet werden, dass sämtliche Intercurrent Events auch wirklich ignoriert werden, wofür wiederum die notwendigen Daten (z.B. Unterscheidung von Therapieabbrechern und Lost-To-Follow-ups) vorhanden sein müssen.

Bias aufgrund von fehlenden Werten

Der TPS-Effekt (Intention-To-Treat) erfordert, dass alle Patient*innen, die zugeteilt wurden, gemäß ihrer ursprünglichen Gruppenzuteilung ausgewertet werden. Dieses verlangt einen adäquaten Umgang mit fehlenden Werten. Voraussetzung hierfür (z.B. für die Wahl des Imputationsverfahrens) sind ausreichend Informationen zu fehlenden Werten.

In prospektiven Studien gibt es feste Erhebungszeitpunkte für die Endpunkte. Dies erlaubt für den jeweiligen Erhebungszeitpunkt eine Zuordnung als Ereignis/Nicht-Ereignis oder fehlenden Wert. Bei vnDa – die per Definition unter Routinebedingungen erfasst werden – gibt es oftmals keine festen Erhebungszeitpunkte, sondern die Ereignisse (z.B. Tod) werden nur registriert, wenn sie tatsächlich auftreten. Aus diesem Grund kann oftmals nicht beurteilt werden, ob ein Ereignis nicht aufgetreten ist oder ein fehlender Wert vorliegt. Ohne Informationen zu Häufigkeit und Grund für das Fehlen eines Werts sind die Auswahl eines geeigneten Verfahrens zum Umgang mit fehlenden Werten und die Anwendung eines solchen jedoch praktisch unmöglich.

Emulationsunterschiede

Neben diesen Verzerrungsquellen aufgrund von systematischer Verzerrung können auch Abweichungen bzgl. Patient*innen, Intervention, Kontrolle und Endpunkten zwischen der NRSvnDa und der Zielpopulation zu abweichenden Ergebnissen führen. Da diese nicht Bias im eigentlichen Sinne verursachen, sondern – ähnlich dem Konzept der Übertragbarkeit (Transferability) – die Vergleichbarkeit von Target-Trial und Emulation beeinflussen, werden diese auch als Emulationsunterschiede bezeichnet [9]. Um Emulationsunterschiede zu vermeiden, müssen die vnDa eine adäquate Emulation, d.h. ausreichende Übereinstimmung mit dem Target-Trial zulassen.

Anforderung an die Analyse

Verschiedene PS-basierte Verfahren gehen mitunter mit verschiedenen Zielpopulationen, über die eine Aussage getroffen werden kann, einher. PS-Matching, Inverse Probability Treatment Weights (IPTW), Fine Stratification Weights (FSW), Matching Weights (MW) und Overlap Weights (OW) erlauben das Schätzen eines ATEs, d.h. sie betrachten die Gesamtpopulation, die für eine Therapie infrage kommt. Studien und Simulationen deuten darauf hin, dass PS-Matching, OW und FSW bzgl. Sicherstellung der Überlappung und Balanciertheit gegenüber anderen Verfahren tendenziell besser abschneiden könnten [10], [11], [12]. Allerdings haben diese den Nachteil, dass ggf. nicht die gesamte potentielle Registerpopulation ausgenutzt wird, sondern nur der überlappende Teil. Dies verringert zum einen die Stichprobengröße und erfordert somit tendenziell größere Datenmengen (z.B. Register) um ein ausreichend großes Analysekollektiv sicherzustellen. Zum anderen werden somit nur „klinisch“ vergleichbare Patient*innen betrachtet, was ggf. die Generalisierbarkeit der Ergebnisse auf die Gesamtpopulation von Interesse negativ beeinflusst.

Implikationen für die Anforderungen an die Daten

Aus den oben aufgezeigten Quellen für Verzerrung und Emulationsunterschiede, ergeben sich die folgenden wesentlichen Anforderungen an die Daten in einem Register, welches für die Nutzenbewertung in Betracht gezogen wird:

- Das Register muss die relevante Indikation gemäß Target-Trial umfassen (korrekte Emulation der Population).
- Das Register muss geeignete Daten zur Emulation der Studienintervention gemäß Target-Trial enthalten (korrekte Emulation der Interventionsgruppe).
- Das Register muss geeignete Daten zur Emulation der Kontrollintervention gemäß Target-Trial enthalten (korrekte Emulation der Kontrollgruppe).

- Das Register muss Informationen zu allen für den Target-Trial relevanten Confoundern enthalten (Vermeidung von Confounding-Bias).
- Das Register muss Endpunkte enthalten, die dem Target-Trial entsprechen (korrekte Emulation der Endpunkte).
- Das Register muss genaue Angaben zu Studienstart, Start der Nachbeobachtung, Beginn der Erfüllung der Einschlusskriterien (z.B. Erkrankungsbeginn) und Behandlungszuteilung umfassen (Vermeidung von Selection- und Immortal-Time Bias).
- Das Register muss, falls relevant (d.h. ggf. relevante Abweichungen zwischen Behandlungsintention und Start der Behandlung), Daten zur Behandlungsintention enthalten (Vermeidung von Bias aufgrund von Abweichung vom TPS-Effekt).
- Das Register muss detaillierte Information über den Behandlungsverlauf enthalten (Vermeidung von Bias aufgrund von Abweichung vom TPS-Effekt).
- Das Register sollte detaillierte Informationen zu fehlenden Werten liefern (Vermeidung von Bias aufgrund von fehlenden Werten).
- Die für die NRSvnda in Frage kommende Population muss ausreichend groß sein, um eine adäquate Power zu gewährleisten (Anwendbarkeit geeigneter Analysemethoden).

In der praktischen Anwendung, d.h. für die Beurteilung der Eignung eines oder mehrerer Register, müssen diese Aspekte zudem hinreichend dokumentiert sein.

Neben diesen wesentlichen, zur Vermeidung von Bias erforderlichen Kriterien gibt es eine Vielzahl von Kriterien der Registerqualität [13]. Diese haben allerdings keinen direkten Einfluss auf die interne Validität. Somit kommt ihnen für die Nutzenbewertung nur eine untergeordnete Rolle zu.

Eignung von Registern in Deutschland

Für die Durchführung von NRSvnda kommt potentiell eine Vielzahl an verschiedenen Datenquellen in Frage. Hierzu zählen u.a. Routinedaten der gesetzlichen Krankenversicherung, Daten der externen stationären Qualitätssicherung und Register. Aktuell scheinen für NRSvnda Register am ehesten geeignet, da für andere Datenquellen davon ausgegangen werden kann, dass insbesondere Einschränkungen aufgrund des Mangels an Informationen zu Confoundern (z.B. klinische Daten) und relevanten Endpunkten bestehen und zudem nicht immer die notwendige Datenqualität (z.B. Information über fehlende Werte) gegeben ist [3].

Mittlerweile gibt es in Deutschland eine große Anzahl an Registern zu verschiedenen Indikationen. So wurden im Rahmen des „Gutachten zur Weiterentwicklung medizinischer Register zur Verbesserung der Dateneinspeisung und -anschlussfähigkeit“ 356 Register identifiziert [14]. Jedoch sind von diesen aktuell nur 29,4% frei zugänglich und nur 36 Register erstellen regelmäßige Qualitätsberichte. Neben diesen generellen Einschränkungen muss für

die jeweilige Fragestellung ein Register vorhanden sein, das eine adäquate Emulation erlaubt, und die Daten müssen geeignet sein, die oben genannten Verzerrungsquellen zu kontrollieren oder zumindest einschätzen zu können (vgl. Anforderung an die Daten). Neben diesen Einschränkungen der praktischen und/oder aufgrund von potentiellen Verzerrungsquellen sinnhaften Durchführbarkeit kommt hinzu, dass der überwiegende Teil der Register als zu klein für die Durchführung von PS-basierten NRSvnda angesehen werden kann (ca. 50% mit weniger als 2.000 Teilnehmern) [14]. Hinzu kommt, dass mitunter nur ein Teil der gesamten Registerpopulation die Einschlusskriterien der Trial-Emulation erfüllt, was sich zusätzlich negativ auf die Power auswirken würde. In Anbetracht der eingeschränkten Zugänglichkeit und der genannten Anforderungen an die Daten wird deutlich, dass aktuell ein hohes Risiko besteht, dass kein geeignetes deutsches Register für eine bestimmte Fragestellung identifiziert werden kann. Wenn ein Register nicht sämtliche notwendigen Informationen enthält, können diese theoretisch zusätzlich erfasst werden. Dies impliziert zum einen, dass die retrospektive Nutzung und somit schnelle Datenverfügbarkeit nicht gegeben ist. Zum anderen kann dies ggf. mit erheblichen Zusatzaufwand verbunden sein. Gleichmaßen wie zusätzliche Datenerfassung ist ein weiterführendes Monitoring der Datenqualität nur prospektiv möglich und mit zusätzlichem Aufwand verbunden.

Ergebnisse von Target-Trial Emulationen

Der Vergleich von Effekten aus NRS und RCT erscheint aufgrund der Vielzahl an potentiellen Ursachen für Unterschiede ohne weitere Kontrolle und Exploration von Einflussfaktoren (Biasquellen und Emulationsunterschiede) nicht zielführend. Erwartungsgemäß zeigen ältere Studien, die naive Vergleiche zwischen den Studientypen anstellen, sehr heterogene Ergebnisse [15]. Neuere Studien versuchen, diese Limitation zu überwinden, indem sie Effekte aus gemäß dem Target-Trial-Konzept durchgeführten NRSvnda und Effekten aus RCTs vergleichen und zudem den Einfluss von Emulationsunterschieden bewerten [16], [17]. Trotz aller methodischen Bemühungen, möglichst viele Verzerrungsquellen, insbesondere Bias aufgrund von Confounding, zu minimieren, zeigen sich auch in diesen Studien abweichende Ergebnisse. Analysen zur Exploration der Ursachen für Abweichungen deuten darauf hin, dass insbesondere datenbezogene Emulationsunterschiede, d.h. abweichende Population, Intervention/Kontrolle, Endpunktdefinition, hierfür verantwortlich sind [17], [18]. Unter den analysebezogenen Faktoren scheinen auf Studienstart/Follow-up-Periode bezogene Verzerrungsquellen den größten Einfluss zu haben [5], [16], [17], [18].

Diskussion und Schlussfolgerung

Es besteht eine Vielzahl an Anforderungen an die Daten, um PS-basierte NRSvnDa zum Zweck der Nutzenbewertung durchzuführen. Anhand der Registerlandschaft in Deutschland konnte aufgezeigt werden, dass die adäquate Emulation eines RCT-Effekts mittels PS-basierten NRSvnDa vermutlich zumeist nicht ohne Weiteres (z.B. zusätzliche Datenerfassung und Maßnahmen zur Sicherung der Datenqualität) möglich ist. Dieses ist mitunter damit zu begründen, dass viele der Register vermutlich nicht vornehmlich mit dem Ziel der Durchführung von NRSvnDa auf deren Basis geplant wurden. In Anbetracht der kleinen Registerpopulation vieler Register muss zudem davon ausgegangen werden, dass Analysen mittels der oben beschriebenen PS-Verfahren ggf. eine geringe Power aufweisen. Linkage verschiedener Register sowie Register und anderer Datenquellen ist aktuell in der Regel nicht (ohne Weiteres) möglich, sodass auch eine Verknüpfung von verschiedenen Registern aktuell keine generell praktikable Lösung darzustellen scheint. Eine geringe Fallzahl bedarf mitunter einer Adaption der bevorzugten kausalen Inferenzmethode bzw. Default-Settings von Software [12]. Da die meisten PS-Verfahren für große Datensätze entwickelt und dementsprechend deren Eigenschaften vorwiegend für solche Situationen untersucht wurden, aktuell jedoch eher kleine Datenbestände zur Verfügung stehen, scheint weiterer Forschungsbedarf zu Eigenschaften von PS-basierten Methoden bei kleinen Fallzahlen gegeben.

In Anbetracht der vielen potentiellen Ursachen für Abweichungen bei den Effekten ist es nicht verwunderlich, dass Studien, die Effekte aus NRS und RCT vergleichen, sehr heterogene Ergebnisse zeigen und dass es bisher nicht gelungen ist, die Effekte aus RCTs auch mit modernen Methoden der kausalen Inferenz zuverlässig zu reproduzieren. Grund hierfür ist offenbar, dass die vorhandenen Daten mitunter keine optimale Emulation des Target-Trials erlauben [16], [17]. Bedenklich erscheint zudem, dass häufig keine eindeutigen Ursachen für Unterschiede von Effekten aus NRSvnDa und RCTs identifiziert werden können [16], [17], [18].

Falls ein Register aufgrund fehlender Daten und/oder Informationen zur Datenqualität nicht direkt zur Durchführung einer PS-basierten NRSvnDa genutzt werden kann, entfällt der Vorteil der schnellen Verfügbarkeit von Ergebnissen und es ist zusätzlicher Ressourceneinsatz notwendig. Dieser sollte gegenüber einer pragmatischen, randomisierten, registerbasierten Studie abgewogen werden, insbesondere da hier keine Gefahr von unbekanntem Confounding und/oder Gefahr von abweichenden Ergebnissen aufgrund von Emulationsunterschieden besteht [19].

Anmerkung

Interessenkonflikte

Der Autor erklärt, dass er keine Interessenkonflikte in Zusammenhang mit diesem Artikel hat.

Literatur

1. Matthews AA, Danaei G, Islam N, Kurth T. Target trial emulation: applying principles of randomised trials to observational studies. *BMJ*. 2022 Aug;378:e071108. DOI: 10.1136/bmj-2022-071108
2. Lodi S, Phillips A, Lundgren J, Logan R, Sharma S, Cole SR, Babiker A, Law M, Chu H, Byrne D, Horban A, Sterne JAC, Porter K, Sabin C, Costagliola D, Abgrall S, Gill J, Touloumi G, Pacheco AG, van Sighem A, Reiss P, Bucher HC, Montoliu Giménez A, Jarrin I, Wittkop L, Meyer L, Perez-Hoyos S, Justice A, Neaton JD, Hernán MA; INSIGHT START Study Group; HIV-CAUSAL Collaboration. Effect Estimates in Randomized Trials and Observational Studies: Comparing Apples With Apples. *Am J Epidemiol*. 2019 Aug;188(8):1569-77. DOI: 10.1093/aje/kwz100
3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V. 2020.
4. Workman TA. Engaging Patients in Information Sharing and Data Collection: The Role of Patient-Powered Registries and Research Networks. Report No. AHRQ 13-EHC124-EF. Rockville (MD): Agency for Healthcare Research and Quality; 2013. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK164513/>
5. Mathes T, Rombey T, Kuss O, Pieper D. No inexplicable disagreements between real-world data-based nonrandomized controlled studies and randomized controlled trials were found. *J Clin Epidemiol*. 2021 May;133:1-13. DOI: 10.1016/j.jclinepi.2020.12.019
6. Pohl M, Baumann L, Behnisch R, Kirchner M, Krisam J, Sander A. Estimands - A Basic Element for Clinical Trials. *Dtsch Arztebl Int*. 2021 Dec 27;118(51-52):883-8. DOI: 10.3238/arztebl.m2021.0373
7. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016 Apr;183(8):758-64. DOI: 10.1093/aje/kwv254
8. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016 Nov;79:70-5. DOI: 10.1016/j.jclinepi.2016.04.014
9. Franklin JM, Glynn RJ, Suissa S, Schneeweiss S. Emulation Differences vs. Biases When Calibrating Real-World Evidence Findings Against Randomized Controlled Trials. *Clin Pharmacol Ther*. 2020 Apr;107(4):735-7. DOI: 10.1002/cpt.1793
10. Cheng C, Li F, Thomas LE, Li FF. Addressing Extreme Propensity Scores in Estimating Counterfactual Survival Functions via the Overlap Weights. *Am J Epidemiol*. 2022 May;191(6):1140-51. DOI: 10.1093/aje/kwac043
11. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ*. 2019 Oct;367:l5657. DOI: 10.1136/bmj.l5657
12. Friedrich S, Friede T. Causal inference methods for small non-randomized studies: Methods and an application in COVID-19. *Contemp Clin Trials*. 2020 Dec;99:106213. DOI: 10.1016/j.cct.2020.106213

13. Stausberg J, Maier B, Bestehorn K, Gothe H, Groene O, Jacke C, Jänicke M, Kostuj T, Mathes T, Niemeyer A, Olbrich K, Schmitt J, Neugebauer E. Memorandum Register für die Versorgungsforschung: Update 2019 [Memorandum Registry for Health Services Research: Update 2019]. Gesundheitswesen. 2020 Mar;82(3):e39-e66. DOI: 10.1055/a-1083-6417
14. Niemeyer A, Semler S, Veit C, Hoffmann W, van den Berg N, Röhrig R, Gurisch C, Schlünder I, Beckedorf I. Gutachten zur Weiterentwicklung medizinischer Register zur Verbesserung der Dateneinspeisung und -anschlussfähigkeit. Erstellt für das Bundesministerium für Gesundheit. 2021.
15. Anglemeyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. Cochrane Database Syst Rev. 2014 Apr 29;2014(4):MR000034. DOI: 10.1002/14651858.MR000034.pub2
16. Matthews AA, Szummer K, Dahabreh IJ, Lindahl B, Erlinge D, Feychting M, Jernberg T, Berglund A, Hernán MA. Comparing Effect Estimates in Randomized Trials and Observational Studies From the Same Population: An Application to Percutaneous Coronary Intervention. J Am Heart Assoc. 2021 Jun;10(11):e020357. DOI: 10.1161/JAHA.120.020357
17. Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, Pawar A, Bessette LG, Lee H, Garry EM, Gautam N, Schneeweiss S. Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT DUPLICATE Initiative. Circulation. 2021 Mar;143(10):1002-13. DOI: 10.1161/CIRCULATIONAHA.120.051718
18. Wang SV, Sreedhara SK, Bessette LG, Schneeweiss S. Understanding variation in the results of real-world evidence studies that seem to address the same question. J Clin Epidemiol. 2022 Nov;151:161-70. DOI: 10.1016/j.jclinepi.2022.08.012
19. Mathes T, Klößen P, Pieper D. No differences were found between effect estimates from conventional and registry-based randomized controlled trials. J Clin Epidemiol. 2019 Jan;105:80-91. DOI: 10.1016/j.jclinepi.2018.09.011

Korrespondenzadresse:

Prof. Dr. Tim Mathes
 Institut für Medizinische Statistik, Universitätsmedizin
 Göttingen, Humboldtallee 32, 37073 Göttingen,
 Deutschland
 tim.mathes@med.uni-goettingen.de

Bitte zitieren als

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, URN: urn:nbn:de:0183-mibe0002593

Artikel online frei zugänglich unter

<https://doi.org/10.3205/mibe000259>

Veröffentlicht: 05.01.2024

Copyright

©2024 Mathes. 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/>.