

Versorgungsforschung – Methoden und Ziele

Welche Studienarten sind relevant bei der Evaluation versorgungsbezogener Interventionen?

PROF. DR. VOLKER ULRICH, LEHRSTUHL FÜR VWL III, INS. FINANZWISSENSCHAFT, UNIVERSITÄT BAYREUTH



Bei der Evaluation versorgungsbezogener Interventionen ist die randomisiert kontrollierte Studie (RCT) der Goldstandard des wissenschaftlichen Belegs. In jüngster Zeit kommen vermehrt Arzneimittel für neuartige Therapien (ATMP = Advanced Therapy Medicinal Products) auf den Markt, für die zum Zeitpunkt der frühen Nutzenbewertung häufig nur Daten aus einarmigen und/oder offenen klinischen Studien vorliegen, auf deren Basis der Nachweis eines Zusatznutzens schwierig ist und vom G-BA auch kaum anerkannt wird. Mit dem Gesetz für mehr Sicherheit in der Arzneimittelversorgung (GSAV 2019) wird der G-BA ermächtigt, für Orphan Drugs und Arzneimittel mit bedingter oder unter besonderen Bedingungen erteilter Zulassung anwendungsbegleitende Datenerhebungen und Auswertungen zum Zweck der Nutzenbewertung zu fordern. Hierbei sind nach Intention des Gesetzgebers auch nicht-randomisierte Studien und damit niedrigere Ergebnissicherheiten zu akzeptieren. Der Beitrag diskutiert die Vor- und Nachteile der unterschiedlichen Studienarten und zeigt auf, welche Chancen bestehen, dass daraus Qualitäts- und Effizienzverbesserungen resultieren.

Stellenwert der Gesundheitsökonomie in der Evaluation versorgungsbezogener Interventionen

In vielen Ländern gibt es Vorschriften, welche die Durchführung gesundheitsökonomischer Studien bzw. Analysen bei versorgungsrelevanten Interventionen vorschreiben (vgl. Schöffski/von der Schulenburg 2012, S. 10). Im Bereich der Arzneimittel geht es beispielsweise nach der Zulassung einer Arzneimittel-Innovation (AMI), bei der die Gesundheitsökonomie keine Rolle spielt, um die Erstattungsfähigkeit der AMI, also um die Frage, ob das neue Medikament auf die Erstattungsliste einer gesetzlichen Krankenkasse kommt und wenn ja, zu welchem Preis. Da das Präparat unmittelbar nach der Zulassung verordnungsfähig ist, stellt die gesundheitsökonomische Evaluation in Deutschland gegenwärtig keine zusätzliche Hürde dar, das Präparat wird im ersten Jahr zum Herstellerabgabepreis vertrieben, bevor das AMNOG-Verfahren einen Erstattungsbetrag bestimmt hat, der ab dem ersten Jahr nach der Markteinführung gilt.¹

In dieser Postlizenzierungsphase, nach einer heute üblicherweise EU-weit erfolgten Zulassung, existieren in den Ländern der EU sehr unterschiedliche explizite Vorschriften darüber, ob und in welcher Form Studien durchgeführt werden müssen. Ohne die Erstattungsfähigkeit im Rahmen von Sozialversicherungssystemen oder anderer Gesundheitssysteme können neue Maßnahmen und Produkte im Gesundheitswesen nicht wirtschaftlich erfolgreich werden.

Zur Überprüfung, ob die zu Lasten der Krankenkassen erbrachten Leistungen medizinisch notwendig, wirksam und wirtschaftlich sind, sind qualifizierte Evaluationen notwendig. Aus diesem Grund kann der Gemeinsame

Bundesausschuss (G-BA) das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) beauftragen, entsprechende Untersuchungen durchzuführen, die dann vom G-BA zur Entscheidungsfindung herangezogen werden (vgl. G-BA 2020).

Die randomisiert kontrollierte Studie (RCT) ist der Goldstandard des wissenschaftlichen Belegs, um klinische Effekte im Sinne von Nutzen und Schaden kausal bestimmten medizinischen Maßnahmen zuschreiben zu können. RCT ist eine experimentelle Studie, bei der die Patienten nach einem Zufallsverfahren (mit verdeckter Zuordnung) auf die Verum- bzw. die Kontrollgruppe verteilt (Randomisierung) und auf das Auftreten der festgelegten Endpunkte in den einzelnen Gruppen nachbeobachtet werden (vgl. Cochrane Deutschland 2020). Zahlreiche Varianten des RCT-Studiendesigns wurden inzwischen entwickelt, um berechtigten kritischen Einwänden sinnvoll zu begegnen und sich an die Herausforderungen dynamischer klinischer Forschungskontexte besser anpassen zu können.

Seit längerem wird eine kontroverse Debatte zwischen der RCT und der Real World Evidence (RWE) geführt. RWE wird aus realen Daten (Real-World-Data, RWD) gewonnen, die außerhalb randomisierter klinischer Studien generiert werden. Zu den Quellen zählen beispielsweise Registerdaten, elektronische Patientenakten, digitale Anwendungen, Beobachtungsstudien oder auch Verwaltungsdaten. Nach Angaben von PharmaBoardroom machen RWD die überwiegende Mehrheit der Patientendaten aus: 95 Prozent im Vergleich zu nur 5 Prozent aus tatsächlichen klinischen Studien (vgl. Blasius 2019, S. 1).

Schon aus diesem Grund ist es nicht sinnvoll, auf diese Daten in der gesundheitsökonomischen Evaluation zu ver-

zichten. RWD können für unterschiedliche Zwecke genutzt werden, so zum Beispiel für neue Indikationen mit wenigen oder gar keinen klinischen Studien und insbesondere auch für die Nachmarktbeobachtung. Eine späte Nutzenbewertung, die RWD heranzieht, hat aufgrund umfangreicherer Daten und längerer Nachbeobachtungszeiten eine höhere Wahrscheinlichkeit, den Nutzen der AMI zu erfassen. Die offene Frage lautet aber, ob sich RWD auch zur frühen Nutzenbewertung im AMNOG-Verfahren eignen. Es wird befürchtet, dass RWD sich von dem etablierten wissenschaftlichen Fundament der Evidence-Based-Medicine (EBM) entfernt (Antes 2016, S. 16).

Dennoch erscheint die Polarisierung zwischen RCT und RWD wenig zielführend. Auch bei der Bewertung eines RCT gilt es zu berücksichtigen, dass zahlreiche methodische Probleme auftreten können. Zum Beispiel können Patienten, die eine Behandlung benötigen, nicht eingeschlossen werden, falls es keine Kontrollgruppe gibt. Beispielhaft zeigt eine schottische Diabetes-Studie, dass nur ein Bruchteil der Patienten mit Typ-2-Diabetes in die sieben großen klinischen Studien eingeschlossen werden konnten (vgl. Braun 2020, S. 2). Auch wenn RWD noch keine RWE darstellt, besteht ein wachsendes internationales Interesse an RWD (vgl. George 2016). Grundsätzlich kann man zur Evidenz auch kommen, wenn man klar strukturierte Forschungsfragen, geeignete Studienarten, solide wissenschaftliche Datenanalyse und Daten aus der Versorgungsforschung miteinander kombiniert und dadurch einen aussagefähigen Vergleich ermöglicht. Bei der Frage RCT versus RWD geht es letztlich nicht um ein Entweder-oder, sondern um ein rationales Sowohl-als auch.

Randomisierte Klinische Studie (RCT)

RCT versuchen, kausale Effekte zu ermitteln. Bei einer Kausalanalyse werden erhobene Daten auf vermutete Ursache-Wirkungs-Beziehungen zwischen den Variablen überprüft. In der Kausalanalyse geht es darum, Daten anhand geeigneter Modelle und Methoden mit dem Ziel zu analysieren, Schlüsse über Interventionseffekte zu ziehen. Die Analyse kann sich auf eine hypothetische (hoher versus niedriger Blutdruck) oder auch eine ganz konkrete Intervention beziehen (Blutdruckbehandlung mit zwei unterschiedlichen Medikamenten A und B; vgl. Didelez 2019, S. 1).

Randomisierung bedeutet in diesem Zusammenhang eine zufällige Zuordnung, etwa von einer Patientin in eine Stichprobengruppe, oder von einem Medikament in eine der zur Verfügung stehenden Stichprobengruppen (Dircks/Gnadt 2016, S. 1ff.). Ziel einer Randomisierung ist zum einen der Ausschluss einer möglichen Einflussnahme auf die Zuordnung und zum anderen die Sicherstellung einer gleichmäßigen Verteilung von bekannten und unbekanntem Einflussfaktoren. Um eine möglichst gleichartige Zusammensetzung der Gruppen zu erzielen, sind mögliche Verzerrungen bei der Einteilung zu vermeiden oder zu minimieren. Ein RCT basiert auf einem Prüfplan, welcher u.a. das Ziel der Studie, den medizinischen Hintergrund, das Studiendesign, die Methodik und die gesamte Planung, Durchführung und Auswertung beinhaltet.

Aufgrund der Zufallszuteilung können randomisierte kontrollierte Studien ausreichend sicherstellen, dass sich die bekannten und unbekanntem Patientenmerkmale, die einen fairen Vergleich von zwei oder mehreren medizinischen Interventionen stören beziehungsweise verzerren können, gleich verteilen (vgl. Friedman et al. 2010, S. 21ff.). Patienten der Kontrollgruppe erhalten entweder keine Therapie bzw. Intervention oder ein Placebo, die Verum- oder Prüfgruppe hingegen erhält eine Therapie bzw. Intervention. Damit lassen sich mögliche Effekte der neuen Therapie bzw. Intervention untersuchen.

Randomisierte Studien gehen in den meisten Fällen mit einer Verblindung einher, die insbesondere den Behandlungs- und Detektionsbias vermeiden soll (vgl. Dircks/Gnadt 2016, S. 1ff.). Bei einer Verblindung weiß der Patient nicht, welche Therapie er erhält; bei doppelter Verblindung wissen es weder der Patient noch der Arzt bzw. Leistungserbringer. Auf diese Weise kann sichergestellt werden, dass

im Studienverlauf auf die Verum- und Kontrollgruppe die gleichen Rahmenbedingungen zutreffen.

Auch wenn die Randomisierung einen wesentlichen Bestandteil eines RCT bildet, ist als Basiselement eines RCTs der Vergleich zwischen Experimental- und Kontrollgruppe entscheidend. Ein Vergleich kann grundsätzlich auch ohne eine Randomisierung stattfinden. In diesen Fällen kann man durchschnittliche kausale Effekt für die Untersuchung heranziehen, die sich durch einen Vergleich der Gruppenmittelwerte bestimmen lassen.

Aktuell bestehen zahlreiche weitere Möglichkeiten zur Variation des RCT-Designs, um eine Anpassung an spezielle Forschungsfragen und klinische Rahmenbedingungen zu erreichen. Dazu gehören unter anderem Cross-over-Studien, n = 1-Studien, faktorielle RCT-Designs oder Cluster-randomisierte Studien (vgl. Lange et al. 2017, S. 637). Ferner stehen adaptive Designs wie moderne Plattformstudien und pragmatische RCT mit vereinfachten klinischen Fragestellungen und weniger stark eingeschränkten Patientengruppen zur Verfügung, die eine breite Rekrutierung von Patienten auch im klinischen Alltag erlauben.

Aufgrund der beschriebenen Möglichkeiten, systematische Verzerrungen zu minimieren, wird die randomisierte klinische Studie nach wie vor als Goldstandard für den Nachweis der Wirksamkeit und Unbedenklichkeit von Interventionen angesehen (Lange et al. 2017, S. 635). Natürlich stimmt die künstliche Studiensituation eines RCT nicht mit den Alltagsbedingungen der Patienten überein, denn die Adhärenz der Patienten und die ärztliche Betreuung sind sicherlich höher, als es im Praxisalltag der Fall wäre (vgl. Stang 2011, S. 661).

Es sollte betont werden, dass auch ein RCT nicht die Gültigkeit von kausalen Schlüssen garantieren kann, da auch die Kausalanalyse von nicht überprüfbar Annahmen ausgeht.² Evidenz-Probleme bei RCT ergeben sich beispielsweise aufgrund einer zunehmenden Zahl von Arzneimitteln, die beschleunigt zugelassen werden und auf einer noch unzureichenden Datenbasis aufbauen. Hier sind insbesondere Arzneimittel für neuartige Therapien (ATMP = Advanced Therapy Medicinal Products) zu nennen, d. h. Gentherapeutika, Zelltherapeutika und biotechnologisch bearbeitete Gewebeprodukte. Im Mai 2020 hatten in der EU und damit auch in Deutschland zehn ATMP eine zentrale Zulassung. Mit der Neuartigkeit gehen Besonderheiten bei ihrer Entwicklung, Produktion, Zulassung, dem Marktzugang und der Erstattung einher.

Einige der Herausforderungen resultieren direkt aus den genterapeutischen Charakteristika:

- relativ kleine Studiengröße,
- häufig kein Vergleichsarm,
- Schwierigkeiten, eine zweckmäßige Vergleichstherapie zu definieren,
- schwer abschätzbare Wirkdauer, die erst nach und nach gezeigt werden kann (vgl. Lücke/Bädeker/Hildinger 2020, S. 51).

Aufgrund des vorhandenen Bedarfs in vielen Anwendungsgebieten werden diese Arzneimittel mitunter bedingt (conditional) oder unter besonderen Umständen (exceptional circumstances) zugelassen. In diesen Fällen liegen zum Zeitpunkt der frühen Nutzenbewertung häufig nur Daten aus einarmigen und/oder offenen klinischen Studien vor, auf deren Basis der Nachweis eines Zusatznutzens schwierig ist und vom G-BA auch kaum anerkannt wird.

Zudem existieren für etliche Anwendungsgebiete, in denen ATMP entwickelt werden, noch gar keine Vergleichstherapien, da es noch keinen anerkannten Behandlungsansatz für die Erkrankungen gibt. Bei ATMP, die nach Einmalgabe über längere Zeit wirksam sein können, lassen sich zudem die Jahrestherapiekosten kaum berechnen oder auch nur abschätzen, sodass alternative Methoden der Kostenrechnung angemessener erscheinen.

In diesem Kontext werden daher unterschiedliche Erstattungskonzepte diskutiert und in Einzelfällen auch schon praktiziert, wie z.B. eine Ratenzahlung über mehrere Jahre hinweg – auch in Abhängigkeit vom Ergebnis der Behandlung oder eine Rückerstattung bei mangelndem Therapieerfolg (Pay for Performance/Pay for Outcome). Der mit dem GKV-FKG (GKV-Fairer-Kassenwettbewerb-Gesetz) im Jahr 2020 beschlossene Risikopool im morbiditätsbedingten Risikostrukturausgleich (Morbi-RSA) hilft den Krankenkassen, insbesondere die finanziellen Belastungen teurer Therapien gegen chronische und akute Erkrankungen besser zu verteilen. Hochkostenfälle werden dadurch finanziell abgedeckt, dass die Krankenkassen für jeden Patienten aus dem Pool 80 Prozent der Leistungsausgaben erstattet bekommen, die über 100.000 Euro pro Jahr hinausgehen.

Das wiederum lässt aber die Pay for Performance-Verträge unattraktiver werden, die mit Ratenzahlungen für die Krankenkassen verbunden sind. Insbesondere mit Blick auf Einmaltherapien oder die erwähnten ATMP besteht sicherlich weiterer Handlungsbedarf, um neue Vertragsformen,

die sowohl bei den pharmazeutischen Herstellern als auch den Krankenkassen Unterstützung finden, zu ermöglichen.

Real-World-Data (RWD)

Durch die Zunahme der therapeutischen Präzision ergeben sich neue Möglichkeiten für die Evidenzgewinnung in der Medizin außerhalb eines RCT-Designs (vgl. Staeck 2017, S. 58). RWD können den Akteuren des Gesundheitswesens (pharmazeutische Hersteller, Hersteller von Medizinprodukten, Behörden, Krankenkassen und Kliniken) wertvolle Zusatzinformationen liefern. Die Anwendungsfelder reichen von der Erkennung von Verhaltensmustern über die Bewertung der Sicherheit und Effektivität von Behandlungen, der Einschätzung der Kosten bestimmter Krankheiten, Risiko-Nutzen-Abwägungen bis hin zu komplexen Diagnosen und Aussagen zur Prävalenz bestimmter Krankheiten (vgl. George 2016, S. 24). Die interessante und kontrovers diskutierte Frage im Zusammenhang mit RWD lautet, ob Versorgungsdaten auch zu einer tragfähigen Evidenz beitragen, um auch im Rahmen der Nutzenbewertung eine zunehmende Rolle spielen zu können (vgl. IQWiG 2020b, S. 2). Eine Gegenüberstellung zentraler Merkmale von RCT und RWD enthält Abbildung 1.

Randomisiert kontrollierte Studien (RCT) versus Real-World-Evidence (RWE)

Variable	RCT	RWE
Zweck	Efficacy	Effektivität
Setting	Experimentelles Setting	Reales Setting
Follow-up	Konzept/Studienplan	Praxisorientiert
Behandlung	Festes Muster	Variables Muster
Studiengruppe	Homogen	Heterogen
Leitung	Teilnehmende Forschungsinstitution	Teilnahme verschiedener Ärzte
Komparator	Placebo/ZVT	Alternative Vergleiche möglich
Patientenüberwachung	Kontinuierlich nach Protokoll	Variabel

Quelle: Eigene Darstellung nach Kim/Lee/Kim 2018.



Abbildung 1: Ein RCT misst die Wirksamkeit unter klinischen Bedingungen, RWE ermittelt die Zielwirksamkeit unter Alltagsbedingungen in der Versorgung.

RCT messen Effekte unter klinischen Bedingungen und zielen somit auf die so genannte Efficacy ab, die Wirksamkeit unter klinischen Studienbedingungen. RWE ermittelt im Unterschied dazu die Zielwirksamkeit (Effektivität) unter Alltagsbedingungen in der Versorgung. Das RCT baut auf einem experimentellen Setting mit homogenen Gruppen auf, die RWE basiert dagegen auf einem realen Setting. Das Follow-up erfolgt nach Studienplan in einem RCT, der RWE liegt eher ein Praxissetting mit heterogenen Studiengruppen zugrunde. Als Komparator dient im RCT-Kontext Placebo oder eine zweckmäßige Vergleichstherapie (ZVT), bei der RWE sind sehr unterschiedliche Interventionen denkbar. Auch die Patientenüberwachung erfolgt unterschiedlich, einmal kontinuierlich gemäß den Studienprotokollen (RCT), zum anderen diskontinuierlich mit mehr Freiheitsgraden (RWE).

Grundsätzlich sind Beobachtungsdaten, die teilweise für andere Zwecke erhoben werden, vergleichsweise stärker von Verzerrungen betroffen, wovon Confounding, also das Auftreten und der Einfluss von Störgrößen, der offensichtlichste und auch schwerwiegendste Bias ist (vgl. Nørgaard/Ehrenstein/Vandenbroucke 2017).

Für Krankheiten, bei denen es bisher keine oder nur eingeschränkte Therapiemöglichkeiten gibt, können RWD bei begrenzter Datenlage die frühere Zulassung von Arzneimitteln über adaptive Entwicklungspfade beschleunigen. Neben den RCT können pragmatische RWD-Ergebnisse möglicherweise auch Laufzeiten und Kosten reduzieren (Staeck 2017, 2020). Interessante Einsatzfelder sind auch die späte oder nachgelagerte Nutzenbewertung, die auf Daten der Versorgungsforschung aufbauen (Nachmarktbeobachtung).

In den USA werden RWD von der Federal Drug Administration (FDA) vor allem für die bereits erwähnte Nachmarktbeobachtung eingesetzt (vgl. FDA 2020a, Johnson 2020). Die Marktverfolgung kann mit den nun zur Verfügung stehenden Daten und Analysemethoden dazu beitragen, Informationsdefizite bei einer frühen Bewertung unmittelbar bei Markteinführung zu beseitigen. Daten aus der realen Welt macht sich die FDA schon seit Langem beispielsweise über die Sentinel-Initiative zunutze (FDA 2020).

Im Herbst 2007 verabschiedete der Kongress den FDA Amendments Act (FDAAA), der die FDA beauftragte, eine Active Postmarket Risk Identification and Analysis (ARIA) einzurichten. Das Gesetz beauftragt die FDA in

Zusammenarbeit mit anderen Institutionen, Methoden zu entwickeln, um einen Zugang zu unterschiedlichen Datenquellen zu erhalten und die Anwendung validierter Methoden zu ermöglichen. Das Sentinel-System ist ein EDV-gestütztes System zur Überwachung der Sicherheit medizinischer Produkte. Das vollständige Sentinel-System wurde im Jahr 2016 implementiert.

Ein Ziel des Sentinel-Systems besteht darin, insbesondere aus Routinedaten der Versicherungen qualitativ hochwertige Registerstudien aufzulegen. Sentinel ersetzt aber nicht die vorhandenen Bewertungsmethoden der FDA, sondern ergänzt diese, indem es den Beteiligten ermöglicht wird, stärker proaktiv bei Datenfragen tätig zu werden. Hier hinkt Deutschland aus einer internationalen Perspektive hinterher, insbesondere was die Möglichkeiten der Stakeholder zu proaktivem Verhalten betrifft (vgl. FDA 2020b).

Ob Daten, die beispielsweise in Registern gewonnen werden, im Kontext der frühen Nutzenbewertung RCT ergänzen oder sogar ersetzen können, wird intensiv diskutiert (vgl. IQWiG 2020a, Staeck 2017 und 2020). Die Frage lautet, ob RWD auch zu einer RWE führen. Viel wird davon abhängen, ob die Versorgungsforschung in der Lage ist, vergleichende Daten in hoher Qualität zu generieren. Die zunehmende Bedeutung von RWD wird dabei auch von den RCT-Verfechtern keineswegs bestritten, es geht im Wesentlichen um die Frage, ob RWD verlässliche Informationen zum (Zusatz-) Nutzen medizinischer Interventionen generieren können (vgl. IQWiG 2020b).

Die gegenwärtige Diskussion zeigt ohne Zweifel interessante Entwicklungsperspektiven für RWD auf, dennoch bleibt es eine Herausforderung, aus den neuen Möglichkeiten eine tragfähige Evidenz abzuleiten, solange die Daten aus der Versorgung wenig strukturiert und mit qualitativen Mängeln erhoben werden. Abbildung 2 fasst Vorteile und Einschränkungen von RWD im Vergleich zu RCT noch einmal synoptisch zusammen.

RWD können im Vergleich zu RCT mit einem geringeren Zeit- und Kostenaufwand einhergehen, zudem besteht eine Chance, die Dauer der klinischen Forschung zu verkürzen. Fristen für Patientenrekrutierung und -einschreibung bestehen nicht in gleichem Ausmaß wie bei einem RCT. RWD liefern zudem Möglichkeiten, Evaluationsforschung auch auf Gebiete auszudehnen, die mit dem klassischen RCT-Ansatz nicht ohne Schwierigkeiten durchgeführt werden können, etwa auf dem Gebiet der

Vorteile und Einschränkungen von Real-World-Data (RWD)

Vorteile RWD

Weniger Zeit- und Kostenaufwand im Vergleich zu RCT

Möglichkeit, die Dauer der klinischen Forschung zu verkürzen

Keine Frist für die Patientenrekrutierung/-einschreibung erforderlich

Forschung wird möglich, die mit RCT nicht durchgeführt werden kann

Seltene Nebenwirkungen erkennbar

Schneller Zugriff/einfacher Informations- und Datenabruf

RCT Prognosemodelle oder Gruppenauswahl mit Risiko behaftet

RWD Fundament für Künstliche Intelligenz-Ansätze (KI)

Nachteile RWD

RWD-Analyse erfordert große Datenmengen

Zeitintensives Datenqualitäts-Management

Erfahrene Experten für die RWD-Analyse erforderlich

Probleme bei der Vertraulichkeit der Daten

Standardisiertes Forschungsprotokoll muss zu Beginn erstellt werden

Voreingenommenheit möglich

Fehlinterpretationen denkbar

Quelle: Eigene Darstellung nach Kim/Lee/Kim 2018.



Abbildung 2: Aktuell ist es trotz der Entwicklungsperspektiven herausfordernd, aus RWD-Daten tragfähige Evidenz abzuleiten.

erwähnten Arzneimittel für neuartige Therapien. Das an der Versorgung orientierte Setting ermöglicht das Erkennen auch seltener Nebenwirkungen und der Zugriff auf die Daten dürfte schneller erfolgen.

Die klinische Welt mit Stratifizierungsstrategie und Prognosemodell ist auch nicht ohne Risiko, hier kann RWD das Fundament für stärker an der Künstlichen Intelligenz (KI) orientierten Versorgungsansätzen bilden. Allerdings erfordert RWD große Datenmengen und ein zeitintensives Datenqualitäts-Management. Da Störgrößen vorhanden sind, braucht es erfahrene Experten, die nicht voreingenommen und vorsichtig bei möglichen Fehlinterpretationen sind. Auch die Vertraulichkeit der Daten dürfte im Vergleich zum RCT schwieriger zu gewährleisten sein.

Rapid Report des IQWiG: Künftige Bedeutung von Registerdaten

Im Rahmen von RWD spielen Registerdaten eine besondere Rolle. Der G-BA hat im Mai 2019 das IQWiG nach Paragraph 139a Abs. 3 SGB V mit einer wissenschaftlichen Ausarbeitung von Konzepten zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach Paragraph 35a SGB V beauftragt. Der sogenannte Rapid Report wurde dem G-BA inzwischen zugestellt (Mai 2020) und enthält Vorschläge zur anwendungsbegleitenden Datenerhebung nach dem Gesetz für mehr Sicherheit in der Arzneimittelversorgung (GSAV), das im August 2019 in Kraft getreten ist (vgl. IQWiG 2020a).

Insbesondere bei beschleunigten Zulassungen und bei Arzneimitteln gegen seltene Erkrankungen (Orphan Drugs) ist die Datenlage zum Zeitpunkt des Marktzugangs für die frühe Nutzenbewertung von Arzneimitteln oft unzureichend. Zur Schließung dieser Evidenzlücken ist das IQWiG bereit, künftig unter bestimmten Bedingungen auch versorgungsnaher Daten für die frühe Nutzenbewertung zu akzeptieren.

In dem Rapid Report wird dazu festgehalten: „Ausführliche Analysen methodischer Fachliteratur sowie intensive Gespräche mit Registerbetreibern und externen medizinischen Biometrikern haben uns zu der Auffassung geführt, dass es bei qualitativ hochwertigen Patientenregistern möglich ist, Studien auf diese Register aufzusetzen und die erhobenen versorgungsnahen Daten für die erweiterte Nutzenbewertung von Arzneimitteln zu verwenden“ (IQWiG 2020b, S. 2).

Die versorgungsnahen Daten können sowohl in Studien mit als auch ohne Randomisierung erhoben werden, nicht aber ohne einen Vergleich zwischen dem neuen Arzneimittel und der vom G-BA bestimmten Vergleichstherapie. Dazu benötigt es entweder studienindividueller Datenerhebungen oder Datenerhebungen in Registern. Insbesondere die Bewertung krankheitsbezogener Patientenregister wird vom IQWiG positiv gesehen (vgl. IQWiG 2020b, S. 66ff.).

Einen umfassenden Überblick über Register in Deutschland gibt es bislang nicht (vgl. Staeck 2020, S. 80). Es ist davon auszugehen, dass bundesweit über 1000 strukturierte Datensammlungen existieren (vgl. Interdisziplinäre Plattform zur Nutzenbewertung 2020, S. 80). Die Heterogenität von Registern hängt auch mit unterschiedlichen Zielen und Be-

dürfnissen der Anwender zusammen, die mit den Daten sehr unterschiedliche Fragestellungen untersuchen. Auch gibt es gegenwärtig noch zu wenige krankheitsbezogene gegenüber produktbezogenen Register.

Eine wichtige Rolle in diesem Prozess kommt auch dem G-BA zu, der anlässlich von Befristungsaufgaben im Zuge der frühen Nutzenbewertung Hersteller verpflichten kann, ein Register zu führen. Auch bei den neuen biotechnologisch hergestellten Arzneimitteln dürften Regis-

ter eine wichtige Rolle spielen, um die Versorgungsprozesse zu analysieren und zu bewerten (vgl. Staeck 2020, S. 81). Das IQWiG gibt in seinem Bericht konkrete Empfehlungen, wie die versorgungsnahe Datenerhebung in Registern für die Nutzenbewertung von Arzneimitteln perspektivisch verwendbar gemacht werden kann (vgl. IQWiG 2020, S. 63ff.).

Sollen versorgungsnahe Daten aus Registern zum Zweck der Nutzenbewertung herangezogen werden,

Anforderungen an versorgungsnahe Daten zum Zwecke der Nutzenbewertung in Abhängigkeit vom Studientyp

Prozessschritt	Generelle Anforderungen (für alle Studientypen)	Spezielle Anforderungen an vergleichende Studien ohne Randomisierung	Spezielle Anforderungen an vergleichende Studien mit Randomisierung	Spezielle Anforderungen an adjustierte indirekte Vergleiche über einen gemeinsamen Brückenkompator
Formulierung, Fragestellung und Entscheidung für ein Studiendesign	<ul style="list-style-type: none"> • Identifizierung der Evidenzlücke • Formulierung der Fragestellung (PICO) gemäß Evidenzlücke • Berücksichtigung der Anforderungen an die Nutzenbewertung aus §35a 	<ul style="list-style-type: none"> • keine Faktoren, die unwahrscheinlich machen, dass mit diesem Studiendesign ausreichend valide Ergebnisse erreicht werden können 	<ul style="list-style-type: none"> • keine sehr großen (dramatischen) Effekte für entscheidungsleitende Endpunkte zu erwarten, interessierende Endpunkte auch unter Vergleichstherapie erreichbar 	<ul style="list-style-type: none"> • prinzipielle Verfügbarkeit von Studien für einen solchen Vergleich (Vorabrecherche)
Studienplanung	<ul style="list-style-type: none"> • explizites, vor Beginn der Datenerhebung finalisiertes Studienprotokoll • präspezifizierter Analyseplan 	<ul style="list-style-type: none"> • Nachbildung der Planung von vergleichenden Studien mit Randomisierung • Präspezifikation möglicher Confounder und deren Adjustierung in der Analyse 	<ul style="list-style-type: none"> • Anpassung der Studienplanung an den Behandlungsalltag (pragmatische randomisierte Studie: Ein- und Ausschlusskriterien, Interventionen, Endpunkte, Visiten) 	<ul style="list-style-type: none"> • Berücksichtigung pragmatischer randomisierter Studien in den Einschlusskriterien • Planung eines systematischen Reviews inkl. Festlegung der Anforderungen an indirekte Vergleiche
Datenerhebung	<ul style="list-style-type: none"> • gewähltes Instrument zur Datenerhebung muss Daten in der notwendigen Qualität zur Verfügung stellen können 	<ul style="list-style-type: none"> • Sicherstellung der Verfügbarkeit der Daten zur Confounderkontrolle 	<ul style="list-style-type: none"> • ggf. Nutzung bestehender Datenstrukturen, z.B. Register 	<ul style="list-style-type: none"> • ggf. Re-Analyse vorhandener Studien, um Anforderungen an indirekten Vergleich zu erfüllen
Analyse und Interpretation	<ul style="list-style-type: none"> • Berücksichtigung der Aussagekraft der unterschiedlichen Studiendesigns und der konkreten Datenqualität bei der Interpretation der Ergebnisse 	<ul style="list-style-type: none"> • Annäherung an Strukturgleichheit der Behandlungsgruppen durch Adjustierung 	<ul style="list-style-type: none"> • Analyse und Interpretation gemäß bestehenden Standards 	<ul style="list-style-type: none"> • Prüfung der Voraussetzungen für einen indirekten Vergleich (Ähnlichkeit, Homogenität, Konsistenz der Studien)

Quelle: IQWiG 2020a, S. xiii.



Abbildung 3: Einschätzung der Möglichkeiten zur Gewinnung versorgungsrelevanter Daten aus Sicht des IQWiG.

bleibt als zentrale Herausforderung, dass das IQWiG als Grundlage über Effekte von Interventionen nur einen Vergleich akzeptiert (vgl. Windeler 2016, S. S1, IQWiG 2020b, S. 3). Die alleinige Betrachtung einarmiger Studien oder einzelner Studienarme reicht nach dem Rapid Report für die Nutzenbewertung nicht aus (vgl. IQWiG 2020b, S. 81). Ob Registerdaten in der Lage sind, Evidenzlücken zu schließen, beispielsweise bei den erwähnten Orphan Drugs oder bei den ATMP, bleibt daher abzuwarten.

Für die versorgungsnaher Datenerhebung zum Zwecke der Nutzenbewertung kommen nach dem IQWiG insbesondere indikationsweite Patientenregister infrage. Denn von den nicht primär auf vergleichende Studien ausgerichteten Instrumenten zur Datenerhebung bieten solche Register am ehesten die Möglichkeit der Anpassung an die notwendige Datenerhebung für diese Studien.

Abbildung 3 gibt einen Überblick, wie das IQWiG die Möglichkeiten zur Gewinnung versorgungsrelevanter Daten zum Zweck der Nutzenbewertung einschätzt. Aus der genannten Notwendigkeit eines Vergleichs folgt, dass ausschließlich vergleichende Studiendesigns relevant sind.

Eine nachvollziehbare Studienplanung ist für die Aussagekraft der Ergebnisse einer Studie von entscheidender Bedeutung. Für die Planung von vergleichenden Studien ohne Randomisierung zum Zweck des Vergleichs von Therapieeffekten wird die explizite Nachbildung der Planung von vergleichenden Studien mit Randomisierung empfohlen (emulation of target trials, vgl. IQWiG 2020a, S. xiii).

Wesentlich für die Analyse von Studien ohne Randomisierung sind Anpassungen bzw. Korrekturen, welche den Einfluss der Strukturungleichheit der Behandlungsgruppen ausgleichen sollen. Die relevanten Störgrößen müssen systematisch in der Studienplanung spezifiziert werden. In den vergangenen Jahren wurden Zielsetzung und Dokumentationsumfang von Registern stark erweitert. Insbesondere die zunehmende Dokumentation klinischer Informationen in Registern ist in diesem Zusammenhang zu erwähnen.

Ausblick

In die Diskussion über Vor- und Nachteile von RCT und RWD ist neuer Schwung gekommen. Der G-BA kann künftig bei ausgewählten Arzneimitteln anwendungsbegleitende Datenerhebungen beauftragen, die der Quantifizierung des Zusatznutzens dienen sollen. Insbesondere bei beschleunigten Zulassungen und bei Arzneimitteln

gegen seltene Erkrankungen ist die Datenlage zum Zeitpunkt des Marktzugangs oft unzureichend.

Das IQWiG bewertet die Situation zwar zurückhaltend, kommt aber in seinem Rapid Report zu dem Ergebnis, dass es bei qualitativ hochwertigen Patientenregistern möglich ist, Studien auf diese Register aufzusetzen und die versorgungsnahen Daten für die erweiterte Nutzenbewertung von Arzneimitteln zu verwenden.

Dennoch wird im Rapid Report an vielen Stellen auch deutlich, dass das IQWiG seine bekannten methodischen Vorbehalte gegen die Nutzung von Versorgungsdaten noch nicht überwunden hat. Auf die Erschließung der Daten aus der Patientenversorgung richten sich aber viele Hoffnungen auf Qualitäts- und Evidenzverbesserungen im Gesundheitswesen.

Diese kontroversen Sichtweisen müssen weiter aneinander angenähert werden, will man Deutschland langfristig nicht von internationalen Entwicklungen mit Bezug auf RWD/RWE abkoppeln. Grundsätzlich besteht bei RWD sicherlich immer das Risiko des Bias, also einer auf die Datenerhebung zurückzuführenden Verzerrung der Ergebnisse. Aber auch RCT können aufgrund der Patientenselektion und der vorgeschriebenen Protokolle mit Verzerrungen verbunden sein. Mittel- bis langfristig sollten sich RCT und RWD ergänzen. Entscheidend wird die Qualität der Daten sein.

E-Mail-Kontakt: Volker.Ulrich@uni-bayreuth.de

Literatur

1. Antes G. (2016): Ist das Zeitalter der Kausalität vorbei? In: Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ) 112S, S. 16-22, URL: <https://www.sciencedirect.com/science/article/pii/S1865921716300708>.
2. Blasius 2019, Real World Evidence. Arzneimittelforschung: In Zukunft mehr Daten aus der realen Welt, DAZ online, 31.01.2019, URL: <https://www.deutsche-apotheker-zeitung.de/news/artikel/2019/01/31/arzneimittelforschung-in-zukunft-mehr-daten-aus-der-realen-welt>.
3. Braun, L. (2016): Nutzenbewertung mit Real World Data? Kontroverse Experten-Diskussion auf dem IQWiG-Symposium, in: Gerechte Gesundheit, Ausgabe 33, URL: www.gerechte-gesundheit.de.
4. Cassel, D. und Ulrich, V. (2020): AMNOG-Daten 2020. Funktionsweise und Ergebnisse der Preisregulierung für neue Arzneimittel in Deutschland, herausgegeben vom Bundesverband der Pharmazeutischen Industrie e. V., Berlin, erscheint demnächst.
5. Cochrane Deutschland (2020): Randomisierte kontrollierte Studie (RCT), URL: <http://www.cochrane.de/de/cochrane-glossar>.
6. Didelez, V. (2019): Die Prinzipien und Methoden der kausalen Inferenz, IQWiG im Dialog, Liefern Kausalmodelle Belege für kausale Zusammenhänge?, Abstracts der Präsentationen, S. 1, URL: <https://www.iqwig.de/de/veranstaltungen/iqwig-im-dialog/iqwig-im-dialog-2019.11793.html>.
7. Dircks, M. und Gnadt, M. (2016): Die Auswahl des richtigen Studientyps. Klinische Studien und ihre Bewertung, Teil 1, in: DAZ online, URL: <https://www.deutsche-apotheker-zeitung.de/daz-az/2016/daz-2-2016/die-auswahl-des-richtigen-studientyps>.

8. FDA – Federal Drug Administration (2020): Postmarketing Surveillance Programs, URL: <https://www.fda.gov/drugs/surveillance/postmarketing-surveillance-programs>.
9. FDA – Federal Drug Administration (2020b): FDA’s Sentinel Initiative – Background, URL: <https://www.fda.gov/safety/fdas-sentinel-initiative/fdas-sentinel-initiative-background>.
10. Friedman, L.M., Furberg, C.D. und DeMets, D.L. (2010): Fundamentals of Clinical Trials, Springer, 4. Aufl., Berlin et al.
11. G-BA - Gemeinsamer Bundesausschuss (2020): Einholung wissenschaftlicher Expertise, URL: <https://www.g-ba.de/ueber-den-gba/aufgabe-arbeitsweise/sach-verstaendige/>.
12. George, E. (2016): How real-world data compensate for scarce evidence in HTA, in: Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ) 112S, S. 23-26., URL: <https://www.sciencedirect.com/science/article/pii/S1865921716300708>.
13. IQWiG (2020a): Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V, Rapid Report, IQWiG-Berichte – Nr. 863, 13.5.2020, Berlin, URL: <https://www.iqwig.de/de/projekte-ergebnisse/publikationen/iqwig-berichte.1071.html>.
14. IQWiG (2020b): Registerdaten eignen sich – bei entsprechender Qualität – für die erweiterte Nutzenbewertung von Arzneimitteln, URL: [iqwig.de/de/presse/pressemitteilungen/2020/registerdaten-eignen-sich-bei-entsprechender-qualitaet-fuer-die-erweiterte-nutzenbewertung-von-arzneimitteln.12795.html](https://www.iqwig.de/de/presse/pressemitteilungen/2020/registerdaten-eignen-sich-bei-entsprechender-qualitaet-fuer-die-erweiterte-nutzenbewertung-von-arzneimitteln.12795.html).
15. Johnson, P. (2020): Maximizing post-market surveillance with real world data. White Paper, IQVIA, URL: https://www.iqvia.com/-/media/iqvia/pdfs/library/white-papers/maximizing-post-market-surveillance-with-real-world-data.pdf?_=1593874522447.
16. Kim, H.-S., Lee, S. und Kim, J.H. (2018): Real-world Evidence versus Randomized Controlled Trial: Clinical Research Based on Electronic Medical Records, in: J Korean Med Sci. 2018, Aug 20; 33(24): e213, S. 1-7, <https://doi.org/10.3346/jkms.2018.33.e213>.
17. Lange, S. et al. (2017): Vielfalt und wissenschaftlicher Wert randomisierter Studien. Teil 24 der Serie zur Bewertung wissenschaftlicher Publikationen, in: Dtsch Arztebl Int, 114: S. 635-640; DOI: 10.3238/arztebl.2017.0635, URL: <https://www.aerzteblatt.de/archiv/193429/Vielfalt-und-wissenschaftlicher-Wert-randomisierter-Studien>.
18. Leschke, M. und Pies, I. (Hrsg.) (2009): Karl Poppers kritischer Rationalismus. Konzepte der Gesellschaftstheorie, Band 5, Mohr Siebeck, Tübingen.
19. Lücke, J., Bädeker, M. und Hildinger, M. (2020): Biotech-Report. Medizinische Biotechnologie in Deutschland 2020, Boston Consulting Group, München.
20. Nørgaard, M., Ehrenstein, V. und Vandembroucke, J.P. (2017): Confounding in observational studies based on large health care databases: problems and potential solutions – a primer for the clinician, in: Clinical Epidemiology, 9, S. 185-193, URL: DOI <https://doi.org/10.2147/CLEP.S129879>.
21. Schöffski, O. und von der Schulenburg, J.-M. Graf (Hrsg.) (2012): Gesundheitsökonomische Evaluationen, 4. Aufl., Berlin et al.
22. Stang, A. (2011): Randomisierte kontrollierte Studien – unverzichtbar in der klinischen Forschung, in: Dtsch Arztebl Int 2011; 108, S. 661-662, DOI: 10.3238/arztebl.2011.0661, URL: <https://www.aerzteblatt.de/archiv/106947/Randomisierte-kontrollierte-Studien-unverzichtbar-in-der-klinischen-Forschung>.
23. Staeck, F. (2017): Das Potenzial von Registern ist groß, ihre langfristige Etablierung anspruchsvoll, in: Schriftenreihe Interdisziplinäre Plattform zur Nutzenbewertung, Lücken in der Evidenz – Was leisten Registerdaten, Heft 5, 2017, Springer Medizin, S. 58-62, URL: <https://www.aerztezeitung.de/Dateien/Luecken-in-der-Evidenz-Was-leisten-Registerdaten-d138.pdf>.
24. Staeck, F. (2020): Unklare regulatorische Anforderungen sind eine Hürde für Register in Deutschland, in: Schriftenreihe Interdisziplinäre Plattform zur Nutzenbewertung, Welchen (Zusatz-)Nutzen haben Registerdaten?, Heft 10, 2020, Springer Medizin, S. 80-84, URL: <https://www.aerztezeitung.de/Dateien/Welchen-Zusatz-Nutzen-haben-Registerdaten-d251.pdf>.
25. Windeler, J. (2016): Real World Data - Adaptive Pathways: Wohin führt der Weg?, in: Z. Evid.Fortbild.Qual.Gesundh.wesen (ZEFQ) 112, S1-S2.

Anmerkungen

¹ Mit dem Arzneimittelmarktneuordnungsgesetz (AMNOG) ist Anfang 2011 eine neue Preisregulierung für patentgeschützte Arzneimittel in Kraft getreten. Seitdem wird nach einer frühen Nutzenbewertung ein Erstattungsbetrag zwischen dem Hersteller und dem Spitzenverband Bund der Krankenkassen (GKV-SV) vereinbart. Dies ist der Preis, zu dem das neue Präparat mit Bindungswirkung für alle gesetzlichen und privaten Krankenversicherungsträger in Deutschland nach dem ersten Jahr erstattet wird (vgl. Cassel/Ulrich 2020).

² Zudem gilt Karl Poppers Einschränkung von der grundsätzlichen Fehlerbarkeit unseres Erkenntnisstrebens (vgl. Leschke/Pies 1999).

PROF. DR. VOLKER ULRICH



Ordinarius für Volkswirtschaftslehre, insbesondere Finanzwissenschaft an der Universität Bayreuth. Er ist Mitglied des Ausschusses für Gesundheitsökonomie der Gesellschaft für Wirtschafts- und Sozialwissenschaften – Verein für Socialpolitik –, Mitglied im wissenschaftlichen Beirat des Bundesversicherungsamts zur Weiterentwicklung des Risikostrukturausgleichs und Vorsitzender des wissenschaftlichen Beirats des Bundesverbands Managed Care (BMC). Im akademischen Turnus 2010/11 war er Vorsitzender der Deutschen Gesellschaft für Gesundheitsökonomie (DGGÖ). Seit 2015 ist der Präsident der Gesellschaft für Recht und Politik im Gesundheitswesen (GRPG).

