Prof. Dr. Franz Porzsolt
Jürgen Geier

Vorteile und Limitationen von Registern und Klinischen Studien in der Versorgungsforschung

Zusammenfassung

Register werden zunehmend häufiger als Alternative zu Klinischen Studien für die systematische Sammlung klinischer Daten verwendet. Da jede Methode zur systematischen Sammlung klinischer Daten spezifische Vorteile und Limitationen aufweist, werden sieben Aspekte von drei Methoden einander gegenübergestellt. Diese drei Methoden sind das nicht spezifizierte Register, das spezifizierte Register (die pragmatische kontrollierte Studie) und die explanatorische Studie (randomisierte kontrollierte Studie). Die sieben Aspekte sind die Einschlusskriterien, die vorab definierte Fragestellung, die von der Fragestellung abhängige Erfassung der Ausgangsrisiken, die Bedingungen, unter welchen die Studie durchgeführt wird, die Ausschlusskriterien, die externe und letztlich die interne Validität. Wenn diese Aspekte der drei Methoden mit den Zielen der Datensamlung abgeglichen werden, wird sich für jede Fragestellung die am besten geeignete Methode finden lassen.

Schlüsselwörter

Register spezifiziert und nicht spezifiziert, pragmatische Studie, randomisierte Studie

Klinische Register werden zunehmend für den Erkenntnisgewinn in der Gesundheitsversorgung genutzt (Bertoletti 2013, Lesko 2013, Majek 2013, Veiby 2013, Weiss 2013). Die Befürworter der Register sehen den Vorteil, aus einem gut geführten Register eine Vielzahl von Fragen beantworten zu können, die zum Zeitpunkt, zu dem das Register initiiert wurde, noch nicht bekannt waren. Da regelmäßig und weitgehend vollständige Datensätze zur Erfassung der erbrachten Leistungen und zur Ermittlung der Leistungsvergütung erhoben werden, ist naheliegend, diese Daten auch für die Beantwortung relevanter Fragen bei der Nutzenbewertung von Arzneimitteln und dem Versorgungsmanagement zu nutzen. Zudem können diese Daten mit Datensätzen aus verschiedenen Quellen zusammengeführt werden, z.B. klinische Daten mit den Daten von Versicherungen oder von Apotheken. Die Skeptiker von Registern befürchten, dass mit Registerdaten nicht immer belastbare Antworten gegeben werden können. Dieses Argument ist nicht von der Hand zu weisen, weil es zwar möglich erscheint, dass ein Register z.B. alle Daten erfassen kann, die beim Betrieb eines technischen Gerätes entstehen. Wenn es sich nicht um ein isoliertes technisches Gerät sondern ein Gerät in einem bestimmten Umfeld, z.B. ein Auto im Straßenverkehr handelt, das komplexen Prozessen, wie dem Verhalten aller Verkehrsteilnehmer, ausgesetzt ist, wird es erheblich schwieriger sein, alle denkbaren Variablen a priori in einem Register abzubilden, die letztlich den Outcome eines Prozesses, beim Auto die Erreichung des Zielorts, beeinflussen. Wenn Register zur Beschreibung von patienten-relevanten Problemen verwendet werden sollen, lässt sich vorhersagen, dass die Komplexität dieser bio-psycho-sozialen Prozesse so hoch sein wird, dass möglicherweise die Grenze der Machbarkeit erreicht wird. Diese beiden Sichtweisen lassen die Vorhersage zu, dass Register in der Versorgungsforschung zweifellos nützlich sein werden, aber auch gewissen Einschränkungen – wie auch die Klinischen Studien – unterliegen. Deshalb möchten wir theoretische Überlegungen wie auch praktische Erfahrungen zu den Chancen und Risiken von Registern und Klinischen Studien vorstellen und Empfehlungen zu deren optimierter Nutzung diskutieren.

>> Bei beiden Methoden, beim Erkenntnisgewinn aus einem Register oder aus einer Klinischen Studie, werden jeweils ein induktiver und ein deduktiver Prozess vollzogen. Der induktive Prozess findet statt, wenn klinische Wissenschaftler Erkenntnisse aus Einzelbeobachtungen gewinnen, diese in einem Register oder in einer Klinischen Studie sammeln und daraus eine allgemein gültige Theorie ableiten. Der deduktive Prozess wird von Praktikern vollzogen, die aus einer erarbeiteten Theorie, die aus den Ergebnissen eines Registers oder einer klinischen Studie abgeleitet wurden, eine Schlussfolgerung für den Einzelfall, d.h. für die Versorgung eines individuellen Patienten ableiten. In beiden Fällen, beim Erkenntnisgewinn aus einem Register oder aus einer klinischen Studie, wird das Risiko eingegangen, dass die Ableitung einer allgemein gültigen Theorie, d.h. der induktive Prozess – an dessen Ergebnis sich der Praktiker später orientiert – nicht oder nur teilweise korrekt war. Dieses Risiko eines fehlerhaften induktiven Prozesses, d.h. die Akzeptanz einer unzutreffenden oder teilweise unzutreffenden Theorie, wird bei einer schulmedizinischen Methode, bei der Daten für die Ableitung einer Theorie gesammelt und interpretiert werden, geringer eingeschätzt als bei einer nicht-schulmedizinischen Methode, bei der die Theorie gebildet wird, ohne dass eine systematische Sammlung von Daten vorausgeht. Bei Anwendung strikter wissenschaftlicher Kriterien ist allerdings einzuräumen, dass wir Wissenschaftler keine Daten haben, mit welchen wir zeigen können, dass die schulmedizinische Bildung von Theorien der nicht-schulmedizinischen Bildung von Theorien überlegen ist. An diesem Punkt werden die Register zunehmend an Bedeutung gewinnen. Für die Entwicklung der Gesundheitsversorgung werden systematisch erhobene Outcomes absolut bedeutend sein. In Tabelle 1 sind sieben Eigenschaften und Bedingungen zusammengefasst, die notwendig sind, um anhand der verfügbaren Daten eine deskriptive und/oder bewertende Analyse durchführen zu können. Diese Tabelle vergleicht die Eigenschaften und Bedingungen von drei Methoden in sieben Punkten. Die drei Methoden sind Register, pragmatische Studien und explanatorische Studien.

1) Allen drei Methoden der Datengewinnung ist gemeinsam, dass die Einschlusskriterien zu definieren sind, um die Bedingungen festzulegen, die erfüllt sein müssen, damit ein Gesunder oder ein Patient in ein Register oder eine Studie aufgenommen werden kann.
2) Der grundsätzliche Unterschied zwischen einem Register und einer klinischen Studie besteht in der Sequenz der Fragestellung und der Datenerhebung. Bei einem Register wird unabhängig von der konkreten Fragestellung das Raster der zu erhebenden Daten definiert. Bei klinischen Studien besteht der erste Schritt in der exakten Definition der Fragestellung. Jede Unschärfe in der Fragestellung wird sich später in einer unpräzisen Antwort auf die Frage niederschlagen (Porzsolt 2011: Fragestellung in Grundlagen der Klinik).

Ursache der unpräzisen Antwort ist eine zu wenig präzisierte Auswahl der Daten, die erforderlich sind, um die Frage wirklich präzise beantworten zu können. Als trivial erscheinendes Beispiel sei erwähnt, dass der Effekt einer Therapie nur präzise beschrieben werden kann, wenn die Kriterien bekannt sind und auch tatsächlich erhoben werden, anhand welcher die Ausgangsrisiken der eingeschlossenen Probanden exakt vorhergesagt werden können. Jeder Störfaktor, der dazu führt, dass die verglichenen Patientengruppen unterschiedlichen Ausgangsrisiken ausgesetzt sind, beeinflusst die beobachteten Outcomes.

3) Zudem ist zu bedenken, dass bei mehreren Fragestellungen die Zugehörigkeit jedes Patienten zu Risikogruppe für jeden Endpunkt gesondert festzulegen ist. So kann ein Patient zur Hochrisiko-Gruppe gehören, wenn das Risiko eines Reinfarkts für den Endpunkt relevant ist; derselbe Patient kann aber zur Niedrigrisiko-Gruppe gehören, wenn als Endpunkt die Verträglichkeit eines Arzneimittels zu prüfen ist. Diese Notwendigkeit der „multiplen Allokation zu unterschiedlichen Risikogruppen" wurde evident, als wir die Methode der Pragmatic Controlled Trials (PCTs) entwickelt haben (Porzsolt 2013). In klinischen Registern werden die Kriterien zur Abschätzung der Risikogruppen nicht immer erfasst. In den Pragmatic Controlled Trials (PCT) ist die Erfassung dieser Kriterien obligat. In Randomized Controlled Trials (RCT) wird angenommen, dass die bekannten und unbekannten Risikofaktoren durch den Prozess der Randomisation annähernd gleich verteilt werden. Diese Annahme lässt sich nur für die bekannten, nicht für die unbekannten Risikofaktoren prüfen. Deshalb ist ein häufig vorgetragenes Argument nur als Annahme zu werten: Es wird behauptet, das RCT sei anderen Studientypen überlegen, weil es auch die unbekannten Risikofaktoren berücksichtigte. Dazu sollten auch die folgenden Überlegungen angestellt werden: Bei jedem Randomized Controlled Trial kann de facto nur eine einzige Fragestellung beantwortet werden, weil die Berechnung der power der Studie von den Eintrittswahrscheinlichkeiten des Endpunkts in der experimentellen Gruppe und der Kontrollgruppe abhängig ist. Wenn am selben Datenmaterial multiple Endpunkte gemessen werden sollen, besteht die Notwendigkeit, die Signifikanzschranken anzupassen. Unklar ist allerdings, ob bei multiplen Fragestellungen ebenfalls davon ausgegangen werden kann, dass die Verteilung der Risikofaktoren, die für jede Fragestellung spezifisch und damit nicht identisch sind, alle Risikofaktoren auf die zufallsgenerierten Studiengruppen gleich verteilt sind. Wir nehmen an, dass mit wachsender Zahl der zu berücksichtigenden Risikofaktoren auch die Wahrscheinlichkeit steigt, dass einer oder mehrere dieser Risikofaktoren nicht gleichmäßig auf alle Gruppen verteilt sind. Deshalb halten wir das Argument der „Gleichverteilung unbekannter Risikofaktoren" durch die Randomisation nur bedingt für korrekt.

4) Ein bedeutender Unterschied zwischen diesen drei Methoden besteht in den Bedingungen, unter welchen die Daten erhoben werden. Wenn ein RCT durchgeführt werden soll, sind die Alltagsbedingungen durch idealisierte Bedingungen zu ersetzen. Diese idealisierten Bedingungen werden durch die Geheimhaltung des Randomisationsplans, durch die Randomisation und ggf. durch die Verblindung gewährleistet. klinische Register und PCTs zeichnen sich durch den Vorteil aus, dass die Ereignisse mit allen Einflüssen abgebildet werden können, wie sie unter Alltagbedingungen tatsächlich auftreten. Deshalb sind Register und PCTs besser geeignet als RCT um die Outcomes unter Alltagsbedingungen unverfälscht abzubilden.

5) Wenn Register und PCTs keine Ausschlusskriterien beinhalten, d.h. alle Individuen aufnehmen, welche die Einschlusskriterien erfüllen, wird ein sampling bias, d.h. der Fehler einer nicht repräsentativen Stichprobe reduziert. In RCT sind dagegen nahezu immer Ausschlusskriterien definiert, um eine möglichst homogene Studienpopulation zu erhalten. Diese erwünschte Homogenität der Studienpopulation lässt sich herstellen, indem auf die Herstellung einer Stichprobe verzichtet wird, die für die Gesamtpopulation repräsentativ ist. An diesem Beispiel lässt sich zeigen, dass durch die Einhaltung bestimmter Bedingungen entweder die interne oder die externe Validität einer Studie, aber nicht beide Formen der Validität gleichzeitig gewährleistet werden können. PCTs nehmen insofern eine Sonderstellung ein, weil bei diesen Studien vor deren Start die Gesamtpopulation präzise zu definieren ist (z.B. alle eligiblen Patienten, die in den ausgewählten Kliniken und/oder Praxen versorgt werden) und bei Durchführung der Studie darauf zu achten ist, dass tatsächlich alle eligiblen Patienten, unabhängig von allen weiteren Variablen, bei der Auswertung der Studie berücksichtigt werden. Bei Beachtung dieser Regel wird die externe Validität der Studie gewährleistet. Die interne Validität eines PCTs kann gewährleistet werden, indem die vier häufigsten Formen systematischer Fehler (selection-, performance-, attrition- und detection-bias) vermieden werden (Angemessene Evidenz für Therapieentscheidungen).

Praktische Erfahrungen

Nicht spezifizierte Register

In einer Studie, welche die Rate der intrauterinen Todgeburten verschiedener Länder anhand von Registerdaten verglich, wurden Unterschiede in der Größenordnung von 2 Promille beobachtet (Weiss 2013). Die Unterschiede in den Zusammensetzungen der untersuchten Populationen betrugen aber in den berichteten Zielgruppen bis etwa 7 Prozent. Das besagt, dass die Ausgangsrisiken der untersuchten Populationen um ein Vielfaches stärker variieren als die beobachteten Effekte und deshalb deskriptive Aussagen auf der Basis von Registerdaten durchaus möglich sind, vergleichende Analysen wegen der Unsicherheit der vergleichbaren Ausgangsrisiken aber kaum möglich sind.

Spezifizierte Register (Pragmatic Controlled Trials; PCTs)

Verlässliche Ergebnisse wurden mittels eines Registers in einer nationalen Dänischen Studie zur Verbesserung der Versorgungsqualität bei Patienten mit Blutungen bei peptischen Ulcus (Rosenstock 2013) erhoben. Bei diesem Projekt war die Fragestellung zu Beginn der Datensammlung bereits definiert und es war bekannt, welche Daten zur Beantwortung der Studienfrage zu beantworten waren. Damit vergleichbar ist ein Tschechisches Register, in welchem entscheidungsrelevante Daten nahezu vollständig vorliegen. Hier lassen sich durch eine retrospektive Analyse gut dokumentierter Patientendaten zur Behandlung Akuter Myeloischer Leukämien spezifische Aussagen machen. So konnten Jindra et al (2013) zeigen, dass das Krankheitsstadium zum Zeitpunkt der Knochenmarkstransplantation und die zytogenetische Risikoklasse den Outcome verlässlicher vorhersagten als der Unterschied zwischen verwandten und nichtverwandten Spendern.

Vergleich von Registern und Randomisierten Kontrollierten Studien (RCT)

Wenn die Aussagen von Register und Randomisierten Kontrollierten Studien miteinander verglichen werden (Morrison 2002, Wenger 1999) lassen sich Aussagen zur Validität nur machen, wenn die Validität der Register und der Studien hinreichend detailliert und sorgfältig untersucht wurde. Die Randomisation einer Studie alleine reicht definitiv nicht aus, um valide Ergebnisse zu generieren. Um die Validität der Aussagen bestätigen zu können, ist es notwendig, zumindest die fünf erwähnten, häufig auftretenden Formen systematischer Fehler auszuschließen (Porzsolt 2012). Im Rahmen der ‚Thrombolysis In Myocardial Infarction (TIMI)' Studie (Bahit 2003) wurde gezeigt, dass in klinischen Studien und in Registern nicht notwendigerweise identische Risikogruppen untersucht werden und Vergleiche zwischen den Ergebnissen von Registern und Studien deshalb problematisch sein können. Bedeutende Aspekte der Versorgungsforschung wurde von den Gruppen um Franciosa (2004), Kalata (2009) und Peelen (2007) angesprochen. Die erste dieser Gruppen (Franciosa 2004) hat hinterfragt, ob die klinischen Erfolge der evidenz-basierten klinischen Forschung, die unter idealisierten Bedingungen erzielt wurden, auch unter Alltagsbedingungen beobachtet werden können. Im konkreten Fall wurde durch den Vergleich eines Registers und einer randomisierten Studie bestätigt, dass die Mortalität und die Hospitalisierung bei kardiologischen Patienten unter Therapie mit Betablockern in einer randomisierten Studie und in einem Register nicht unterschiedlich waren. Wenn die geforderte Validität der beiden Datenquellen vorliegt, bestätigt dieser Befund die Alltagstauglichkeit der Ergebnisse randomisierter kontrollierter Studien. Im Gegensatz dazu konnten Kalata und Kollegen (2009) bei Patienten mit Stadium II/III eines Rektumkarzinoms die Vergleichbarkeit von Register- und Studiendaten nicht bestätigen. Peelen und Kollegen (2007) haben darauf hingewiesen, dass viele Untersuchungen das angestrebte Ziel nicht erreichen, weil zwar die Gesamtzahl aller eingeschlossenen

Literatur

Bahit MC, Cannon CP, Antman EM, Murphy SA, Gibson CM, McCabe CH, Braunwald E; TIMI 9 Trial and TIMI 9 Registry. Thrombolysis In Myocardial Infarction. Direct comparison of characteristics, treatment, and outcomes of patients enrolled versus patients not enrolled in a clinical trial at centers participating in the TIMI 9 Trial and TIMI 9 Registry. Am Heart J. 2003;145:109-117.

Bertoletti L, Quenet S, Laporte S, Sahuquillo JC, Conget F, Pedrajas JM, Martin M, Casado I, Riera-Mestre A, Monreal M. Pulmonary embolism and 3-month outcomes in 4036 patients with venous thromboembolism and chronic obstructive pulmonary disease: data from the RIETE registry. Respir Res. 2013;14:75. [Epub ahead of print]

Dignam J, Huang L, Ries L, Reichman M, Mariotto A, Feuer E. Estimating Breast Cancer-Specific and Other Cause Mortality in Clinical Trial and Population-Based Cancer Registry Cohorts. Cancer. 2009; 115: 5272–5283. doi:10.1002/cncr.24617.

Franciosa JA, Massie BM, Lukas MA, Nelson JJ, Lottes S, Abraham WT, Fowler M, Gilbert EM, Greenberg B; COHERE Participant Physicians. Beta-blocker therapy for heart failure outside the clinical trial setting: findings of a community-based registry. Am Heart J. 2004;148:718-726.

Jindra P, Muzik J, Indrak K, Zak P, Sabty FA, Kozak T, Cetkovsky P, Karas VK, Raida L, Szotkowski T. The outcome of allogeneic HSCT in older AML patients is determined by disease biology and not by the donor type: An analysis of 96 allografted AML patients ≥50 years from the Czech acute leukaemia clinical register (alert). Neoplasma. 2013;60:576-583. doi: 10.4149/neo_2013_075.

Kalata P, Martus P, Zettl H, Rödel C, Hohenberger W, Raab R, Becker H, Liersch T, Wittekind C, Sauer R, Fietkau R; German Rectal Cancer Study Group. Differences between clinical trial participants and patients in a population-based registry: the German Rectal Cancer Study vs. the Rostock Cancer Registry. Dis Colon Rectum. 2009;52:425-437. doi: 10.1007/DCR.0b013e318197d13c.

Kim SY. Need for Clinical Trial Registry and Institutional Review Board Approval in Human Subjects Research. Korean J Fam Med. 2013;34:1 http://dx.doi.org/10.4082/kjfm.2013.34.1.1

Lesko MM, Jenks T, O'Brien S, Perel P, Childs C, Bouamra O, Lecky F. Models of Mortality Probability in Severe Traumatic Brain Injury: Results of the Modelling by the UK Trauma Registry (TARN). J Neurotrauma. 2013. [Epub ahead of print]

Majek O, Gondos A, Jansen L, Emrich K, Holleczek B, Katalinic A, Nennecke A, Eberle A, Brenner H; GEKID Cancer Survival Working Group. Sex differences in colorectal cancer survival: population-based analysis of 164,996 colorectal cancer patients in Germany. PLoS One. 2013;8:e68077. doi: 10.1371/journal.pone.0068077. Print 2013.

Morrison DA, Sethi G, Sacks J, Henderson W, Grover F, Sedlis S, Esposito R, Ramanathan KB, Weiman D, Talley JD, Saucedo J, Antakli T, Paramesh V, Pett S, Vernon S, Birjiniuk V, Welt F, Krucoff M, Wolfe W, Lucke JC, Mediratta S, Booth D, Barbiere C, Lewis D; VA AWESOME (Angina With Extremely Serious Operative Mortality Evaluation) Multicenter Registry. Percutaneous coronary intervention versus coronary bypass graft surgery for patients with medically refractory myocardial ischemia and risk factors for adverse outcomes with bypass: The VA AWESOME multicenter registry: comparison with the randomized clinical trial. J Am Coll Cardiol. 2002;39:266-273.

Peelen L, Peek N, de Jonge E, Scheffer GJ, de Keizer NF. The use of a registry database in clinical trial design: assessing the influence of entry criteria on statistical power and number of eligible patients. Int J Med Inform. 2007;76:176-183. Epub 2006 Jul 7.

Porzsolt F. Bedeutung einer zielgerichteten Versorgungsstrategie. In: Porzsolt F (Hrsg). Grundlagen der Klinischen Ökonomik. Schriftenreihe PVS Verband Band 11, 1.Auflage. Berlin 2011, pp 25-33.

Porzsolt F, Braubach P, Flurschütz PI, Göller A, Sailer MB, Weiss M, Wyer P. Medical Students Help Avoid the Expert Bias in Medicine. Creative Education 2012;3:1115-1121 doi:10.4236/ce.2012.326167

Porzsolt F, Eisemann M, Habs M, Wyer P. Form Follows Function: Pragmatic Controlled Trials (PCTs) have to answer different questions and require different designs than Randomized Controlled Trials (RCTs). J Publ Health 2013;21:307-313. DOI 10.1007/s10389-012-0544-5

Rosenstock SJ, Møller MH, Larsson H, Johnsen SP, Madsen AH, Bendix J, Adamsen S, Jensen AG, Zimmermann-Nielsen E, Nielsen AS, Kallehave F, Oxholm D, Skarbye M, Jølving LR, Jørgensen HS, de Muckadell OB, Thomsen RW. Improving Quality of Care in Peptic Ulcer Bleeding: Nationwide Cohort Study of 13,498 Consecutive Patients in the Danish Clinical Register of Emergency Surgery. Am J Gastroenterol. 2013. doi: 10.1038/ajg.2013.162. [Epub ahead of print]

Thomas DC, Siemiatycki J, Dewar R, Robins J, Goldberg M, Armstrong BG. The problem of multiple inference in studies designed to generate hypotheses. Am J Epidemiol. 1985;122:1080-1095.

Veiby G, Daltveit AK, Schjølberg S, Stoltenberg C, Oyen AS, Vollset SE, Engelsen BA, Gilhus NE. Exposure to antiepileptic drugs in utero and child development-A prospective population-based study. Epilepsia. 2013. doi: 10.1111/epi.12226. [Epub ahead of print]

Weiss E, Krombholz K, Eichner M. Fetal mortality at and beyond term in singleton pregnancies in Baden-Wuerttemberg/Germany 2004-2009. Arch Gynecol Obstet. 2013. [Epub ahead of print]

Weng C, Bigger JT, Busacca L, Wilox A, Getaneh A. Comparing the Effectiveness of a Clinical Registry and a Clinical Data Warehouse for Supporting Clinical Trial Recruitment: A Case Study. AMIA Annu Symp Proc 2010. 2010:867-871.

Wenger NK. Women, myocardial infarction, and coronary revascularization: concordant and discordant clinical trial and registry data. Cardiol Rev. 1999;7:117-120.

Vergleich der Eigenschaften und Bedingungen				
Nr	Eigenschaft/Bedingung	Nicht spezifiziertes Register	Spezifiziertes Register (Pragmatic Controlled Trial)	Randomized Controlled Trial (RCT)
1	Einschlusskriterien	Definiert	Definiert	Definiert
2	Fragestellung vorab definiert	Nein	Ja	Ja
3	Ausgangsrisiko, abh. v. Fragestellung	In nicht spezif. Registern unberücksichtigt	Spezifisch Risikogruppen f. jede einzelne Fragestellung	Risiken durch Randomisation ausgeglichen
4	Bedingungen	Alltagsbedingungen	Alltagsbedingungen	Idealisierte Bedingungen
5	Ausschlusskriterien	Keine	Keine	Definiert
6	Externe Validität	Gewährleistet	Gewährleistet	Nicht gewährleistet
7	Interne Validität	Nicht gewährleistet	Meist gewährleistet	Gewährleistet

Tab. 1: Vergleich der Eigenschaften und Bedingungen nicht spezifizierter Register, spezifizierter Register (Pragmatic Controlled Trial; PCT) und von randomisierten Studien (Randomized Controlled Trial; RCT). Gleiche Farben in einer Zeile weisen auf gleiche Eigenschaften/Bedingungen hin (eigene Darstellung).

Probanden in ein Register oder in eine Studie hinreichend groß erscheint, aber die Subgruppen, bei welchen bestehende Risiken gesenkt werden sollen, für belastbare Aussagen zu klein sind. Andere Fehlerquellen beim Vergleich von Register- und Studiendaten betreffen die Quellen der erhobenen Daten. So beinhalten Totenscheine und klinisch schwer zu eruierende Todesursachen wie die ‚disease specific mortality' selten validen Daten, um kausale Zusammenhänge bestätigen zu können (Dignam 2009). Die Messung der krankheitsspezifischen Mortalität ist von erheblichen Unsicherheiten begleitet besonders wenn es sich um chronisch verlaufende Erkrankungen handelt, weil mit zunehmender Krankheitsdauer die Wahrscheinlichkeit steigt, dass kein kausaler Zusammenhang zwischen einem eingetretenen Todesfall mit der bekannten chronischen Erkrankung besteht. Dennoch darf nicht unerwähnt bleiben, dass Registerdaten für die Rekrutierung von Probanden in klinischen Studien äußerst wertvoll sein können (Wenig).

Diskussion

Die Literatur über klinische Register begründet die Annahme, dass ein weites Spektrum besteht, welches bei der Sammlung klinischer Basisdaten beginnt, die für viele Fragestellung benötigt werden und bis hin zu sehr spezifisch strukturierten klinischen Studien reicht, mit welchen sich nur eine einzige, sehr spezifische Frage beantworten lässt. Wir sehen deshalb – wie eingangs erwähnt – den wesentlichen Unterschied zwischen Registern und klinischen Studien in der Sequenz der Fragestellung. Bei klinischen Studien ist die Fragestellung vor der Entscheidung über die zu sammelnden Daten zu definieren, während bei Registern zuerst die Daten gesammelt werden und anschließend die zu beantwortenden Fragen gestellt werden. Dazwischen gibt es Variationen, die mehr einer klinischen Studie oder mehr einem Register ähneln.

Wir gehen davon aus, dass Register ebenso wie klinische Studien zum Erkenntnisgewinn beitragen und von klinischen Forschern, ebenso wie von Praktikern, von Krankenkassen und von Marketing-Abteilungen der Industrie genutzt werden. In Tab. 1 ist dargestellt, dass der wesentliche Unterschied der verschiedenen Methoden zum Gewinn neuer Erkenntnisse in der Spezifizierung der zu erhebenden Daten besteht.

Wenn nur eine einzelne konkrete Fragestellung zur Wirkung unter Idealbedingungen (efficacy) zu beatworten ist, ist das RCT die am besten geeignete Methode. Bei einem RCT kann nur das Hauptzielkriterium verlässlich geprüft werden; Ergebnisse, die sich bei der Prüfung von Nebenkriterien ergeben, sind als Hypothesen zu verstehen, weil sich die Power-Berechnung bei einem RCT (unabhängig, ob es sich um eine Überlegenheitsstudie oder ein „Non-Inferiority-Trial" handelt) lediglich auf jenen Unterschied hinsichtlich des Hauptzielkriteriums bezieht, der als klinisch relevant angesehen wird. Wenn die Nebenkriterien ebenfalls mit statistischen Methoden geprüft werden sollten, wäre für jedes einzelne Nebenkriterium ebenfalls der Unterschied zu benennen, der als klinisch relevant angenommen wird. Diese Überlegung zum Grenzwert, der noch als klinisch relevant angesehen wird, ist Grundlage jeder Berechnung der Fallzahl, die notwendig ist, um den von uns Praktikern erwünschten „statistisch signifikanten Unterschied" nachweisen zu können. Diese Überlegungen zeigen zudem, dass bei einer sorgfältig geplanten und sorgfältig durchgeführten klinischen Studie beide Aspekte, jener der statistischen Signifikanz und jener der klinischen Relevanz berücksichtigt werden. Daraus lässt sich im Umkehrschluss ableiten, dass die Power-Berechnung einer Studie offensichtlich von unrealistischen Annahmen ausgegangen war, wenn ein zwar statistisch signifikanter Unterschied nachgewiesen wird, der klinisch aber nicht relevant ist. Die klinische Relevanz einer Intervention lässt sich nur beurteilen, wenn die unter Alltagsbedingungen erreichten Versorgungsergebnisse abgebildet werden.

Da unter Alltagsbedingungen Unterschiede nicht nur hinsichtlich eines einzelnen Endpunkts, sondern hinsichtlich mehrerer Endpunkte zu prüfen sind, ist die bekannte Limitation des multiplen Testens zu berücksichtigen (Thomas 1985). Diese Limitation betrifft die für einen signifikanten Unterschied akzeptierten statistischen Grenzen, die umso enger zu fassen sind, je mehr Teste mit einem definierten Datensatz durchgeführt werden. Die Berücksichtigung der statistischen Grenzen alleine reicht aber nicht aus, um belastbare Ergebnisse aus erhobenen Daten abzuleiten. Es ist auch sicherzustellen, dass die Ausgangsrisiken in den verglichenen Gruppen identisch sind. Bei jedem RCT wird angenommen, dass die bekannten und unbekannten Risikofaktoren durch die zufällige Gruppenzuordnung der Probanden etwa gleich verteilt sind. Diese Annahme mag gelten, solange nur ein einzelnes Hauptzielkriterium untersucht wird. Wenn aber viele Kriterien in einer Studie untersucht werden, steigt das Risiko, dass die Verteilung der Ausgangsrisiken, die für jedes dieser Zielkriterien unterschiedlich sind (Porzsolt 2013), in den verglichenen

Gruppen nicht mehr identisch ist. In der Versorgungsforschung sind in der Regel mehrere Fragstellungen innerhalb einer einzigen Studie zu beantworten, weil z.B. neben Unterschieden in den klinischen Erfolgen auch Unterschiede der Verträglichkeit, der Nebenwirkungen und vor allem der Kosten zu untersuchen und gegebenenfalls zu bestätigen sind.

Ein RCT ist deshalb aus zwei Gründen für Studien zur Versorgungforschung wenig geeignet. Zum einen kann in einem RCT nur ein Zielkriterium geprüft werden, und zudem setzt ein RCT idealisierte Studienbedingungen voraus, die in der Versorgungsforschung gerade nicht von Interesse sind. Dort will man untersuchen, welche Effekte unter Alltagsbedingungen zu beobachten sind. Deshalb halten wir ein Pragmatic Controlled Trial (PCT), das auch als spezifiziertes Register bezeichnet werden kann (Tab. 1), für besser geeignet als ein RCT, um Fragen der Versorgungsforschung zu beantworten.

Als nicht spezifiziertes Register bezeichnen wir eine Sammlung von Daten zur Beantwortung von Fragestellungen, die erst zu einem späteren Zeitpunkt formuliert werden (Tab.1). Da aus den genannten Gründen die notwendigen Informationen (z.B. die Risikofaktoren) in den nicht spezifizierten Registern nicht regelmäßig vorliegen werden, eigen sich diese Register weniger für vergleichende als für deskriptive Analysen. Wenn in einer deskriptiven Analyse Informationen fehlen, entsteht zwar eine Informationslücke, aber die Validität der Analyse wird dadurch nicht beeinträchtigt. Wenn auf der Basis nicht spezifizierter Register, d.h. ohne prospektiv formulierte Fragestellungen gestartete Register, vergleichende Analysen angestellt werden sollen, ist vorab zu klären, ob das Register auch alle Informationen z.B. zu den Risikofaktoren enthält, die für den geplanten Vergleich erforderlich sind.

Diese Unterschiede zwischen den drei Arten von Datensammlungen, den nicht spezifizierten Registern, den spezifizierten Registern (oder pragmatischen Studien) und den randomisierten Studien möchten wir als Beitrag zur Versorgungsforschung zur Diskussion stellen. <<

Advantages and limitations of registries and clinical trials in health services research

Registries are increasingly used as alternatives to clinical trials for the systematic collection of clinical data. As any method for the systematic collection of clinical data has specific advantages and limitations seven aspects of three methods are compared. The three methods are the not specified registry, the specified registry (pragmatic controlled trial) and the explanatory trial (randomized controlled trial). The seven aspects are the inclusion criteria, the pre-defined study question, the baseline risks which depend on the study question, the study conditions, the exclusion criteria, the external and finally the internal validity. When these aspects of the three methods fit the goals of the data collection the optimal method to answer any study question can be found.

Keywords

Registry specified and not specified, pragmatic controlled trial, randomized controlled trial

Autorenerklärung

Jürgen Geier ist Mitarbeiter der Servier Deutschland GmbH. Franz Porzsolt ist Berater der Servier Deutschland GmbH. Beide haben zu gleichen Teilen zum vorliegenden Manuskript beigetragen.

Prof. Dr. Franz Porzsolt
ist Hämatologe und internistischer Onkologe mit langjähriger klinischer Erfahrung. Er hat das Konzept der Klinischen Ökonomik entwickelt, das eine Bewertung des Nutzens von Gesundheitsleistungen primär aus Sicht des Patienten fordert und für Patienten mehr Rechte aber auch mehr Pflichten (subjektive Bewertung des Erfolges von Gesundheitsleitungen) für unabdingbar hält.
Kontakt: franz.porzsolt@uniklinik-ulm.de



Jürgen Geier
Jürgen Geier hat eine wirtschaftswissenschaftliche Ausbildung und übte als Krankenkassenfachwirt verschiedene Geschäftsführungs -und Vorstandspositionen in der gesetzlichen Krankenversicherung aus, bis er 2001 in die pharmazeutische Industrie wechselte. Hier leitet er die Abteilung Gesundheitsökonomie und Politik der Servier Deutschland GmbH. Die Erfahrungen aus den Tätigkeiten in den Leitungsbereichen beider Arbeitsgebiete haben seine nicht nur einseitige Bewertung gesundheitspolitischer Fragestellungen geprägt. Kontakt: juergen.geier@de.netgrs.com

