Artikel online veröffentlicht: 01.12.2021

Was verstehen wir unter Sekundärdaten? – Ein Grundsatzbeitrag zur terminologischen Einordnung und Definition

What do We Mean by Secondary Data? – A Keynote Contribution to the Terminological Classification and Definition

Autoren
Holger Gothe[1, 2, 3], Peter Ihle[4], Enno Swart[5]

Institute
1 Lehrstuhl für Gesundheitswissenschaften / Public Health, Medizinische Fakultät Carl Gustav Carus, Dresden, Deutschland
2 Department für Public Health, Versorgungsforschung und HTA, Institut für Public Health, Medical Decision Making und HTA, UMIT, Hall, Austria
3 Versorgungsforschung, IGES Institut GmbH, Berlin, Deutschland
4 PMV forschungsgruppe an der Medizinischen Fakultät und Uniklinik Köln, Universität zu Köln, Medizinische Fakultät, Köln, Deutschland
5 Institut für Sozialmedizin und Gesundheitssystemforschung, Otto-von-Guericke-Universität Magdeburg Medizinische Fakultät, Magdeburg, Deutschland

Schlüsselwörter
Sekundärdaten, Routinedaten, Sozialdaten, Krankenkassendaten, Definition

Key words
Secondary Data, Routine Data, Claims Data, Sickness Fund Data, Definition

Bibliografie
Gesundheitswesen 2021; 83 (Suppl. 2): S64–S68
DOI 10.1055/a-1686-8936
ISSN 0949-7013
© 2021. Thieme. All rights reserved.
Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Korrespondenzadresse
Dr. Holger Gothe
Affiliated Senior Scientist
Lehrstuhl für Gesundheitswissenschaften / Public Health
Medizinische Fakultät Carl Gustav Carus
Technische Universität Dresden
Fetscherstraße 74
01307 Dresden
Deutschland
holger.gothe@mailbox.tu-dresden.de

ZUSAMMENFASSUNG

Daten sind für die Versorgungsforschung Material und Produkt zugleich. Als „Substrat“ oder Ausgangsbasis von gesundheitsbezogenen Analysen haben sie in den vergangenen Dekaden immens an Bedeutung gewonnen. Für die Beurteilung von Leistungen des Gesundheitssystems, für dessen Steuerung und Entwicklung sind Daten ein essenzieller Rohstoff. Im wissenschaftlichen Alltag und in Forschungsarbeiten, die diesen Rohstoff nutzen, kommt der Bedeutungsgehalt des Begriffs „Daten“ – insbesondere „Sekundärdaten“ – oft zu kurz. Der vorliegende Beitrag möchte den Bedeutungskontext der beiden Begriffe vertiefend beleuchten und den Versuch unternehmen, den Begriff „Sekundärdaten“ terminologisch einzuordnen.

ABSTRACT

Data are both material and product for health services research. As a “substrate” or starting point for health-related analyses, they have gained immense importance in recent decades. Data are an essential raw material for the assessment of services in the health care system, for its control and development. In everyday scientific life and in research work that uses this raw material, the meaning of the term “data” – especially “secondary data” – is often neglected. This article aims to shed more light on the context of meaning of the two terms and to attempt to classify the term “secondary data” terminologically.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

Daten sind für die Versorgungsforschung Material und Produkt zugleich. Als „Substrat“, als Grundsubstanz oder Ausgangsbasis von gesundheitsbezogenen Analysen haben sie in den vergangenen Dekaden immens an Bedeutung gewonnen – so sehr, dass der Volksmund sie heutzutage auf eine Stufe stellt mit einem Rohstoff, der eine Schlüsselrolle in der Entwicklung der industrialisierten Gesellschaft gespielt hat: „Daten sind das Öl des 21. Jahrhunderts“ [1]. Für die Beurteilung von Leistungen des Gesundheitssystems, für dessen Steuerung und Entwicklung sind Daten eine Essenz, auf die nicht verzichtet werden kann. Doch was sind eigentlich „Daten“, und was verstehen wir insbesondere unter „Sekundärdaten“?

Wortherkunft und Terminologie

Der Duden führt den deutschen Begriff *Daten* auf seinen lateinischen Ursprung zurück. Das Wort [Da · ten] ist der Plural von *Datum*, was wörtlich übersetzt so viel bedeutet wie *das Gegebene* (substantiviertes Partizip Perfekt Passiv (PPP) des lateinischen Verbs *dare* – geben). Die Verwurzelung des heute gebräuchlichen Daten-Begriffs in der lateinischen Sprache beschert dem unscheinbaren Wort seine internationale Verständlichkeit. Sie basiert letztlich auf seiner griechischen Herkunft: δεδόμενα – auch dieses flektierte Verb folgt in seiner Konjugation demselben Muster: Es handelt sich um ein PPP, seiner Morphe nach dem altgriechischen Augmenttempus entsprechend. Doch gleich, ob *data* [engl.], *dati* [ital.], *datos* [span.] oder *dados* [port.] – all diesen Wörtern ist ihre gemeinsame Herkunft anzusehen. Selbst das französische Wort *données* ist ohne Weiteres zu erkennen als etwas *Gegebenes*, darin stimmen auch die Niederländer ein, wenn sie von *gegevens* sprechen.

Im allgemeinen deutschen Sprachgebrauch hat sich der Bedeutungsgehalt der Singularform *Datum* mit der Zeit auf das *Kalenderdatum* reduziert. Für den Singular von *Daten* im Sinne einer Informationsangabe, eines gegebenen Messwertes oder einer Zeichenfolge wird im Deutschen nicht das Wort *Datum v*erwendet, sondern Begriffe wie *Angabe*, *Wert* oder *Variable*[2]. In sprachlicher Hinsicht handelt es sich bei dem Wort *Daten* somit um einen *pluralis tantum*, ein ausschließlich im Plural gebräuchliches Substantiv, das auch einen eigentlich singulären Bedeutungsgehalt annehmen kann: die *Daten* (Pl.) – gemeint ist aber ein *Datensatz* (Sg.), ein *Datenkörper* (Sg.), ein *Datenkranz* (Sg.) oder Ähnliches. Dies kommt auch darin zum Ausdruck, dass bei Verwendung des angloamerikanischen Begriffes *data* in englischen Texten Unsicherheit besteht, ob das Wort im weiteren Verlauf eine Satzkonstruktion im Singular oder Plural erfordert. Dabei ist kaum mehr gegenwärtig, dass *data*, aus dem Lateinischen abgeleitet, die Pluralform darstellt, wobei die wörtliche deutsche Übersetzung des substantivierten Neutrums „*das Gegebene*“ sich nicht ohne Zuhilfenahme eines weiteren Wortes in Pluralform konstruieren lässt (z. B. „*die gegebenen < Sachverhalte >* “).

Nach diesem Prolog könnten wir der Versuchung erliegen, die *Primärdaten* – zumindest aus sprachlicher Sicht – hintanzustellen, denn diese sind nicht bereits *gegeben* im Sinne von vorhanden, sie müssen im Forschungsprozess erst noch erzeugt werden, bevor sie für empirische Zwecke genutzt werden können. Anders verhält es sich bei den sog. *Sekundärdaten*: Sie sind im sprachlichen Sinne so etwas wie die „eigentlichen“ *Daten*, sind sie doch tatsächlich bereits *gegeben*, d. h. vorhanden. Diese Gegebenheit birgt jedoch ihre Tücken, wie erfahrene Sekundärdatenforscher/innen wissen. Da sie nicht für die Wissenschaft „gedacht und gemacht“ sind [3], sondern aus anderen Kontexten (z. B. Abrechnungsprozessen) bezogen werden und somit erst aufbereitet werden müssen, können Probleme wie *Missklassifikation* oder *Confounding* entstehen, denen neuere Ansätze der Sekundärdatenanalyse beizukommen versuchen [4–6].

Definitionen

Daten in der Umgangssprache bezeichnen gemäß Wahrigs Deutschem Wörterbuch eine „Tatsache, Angabe“ und liefern „Informationen über Werte, Größen u. a.“ [7]. Daten in fachlicher Hinsicht sind das Ergebnis eines Erhebungsvorganges und „alle in Form von Zeichen darstellbaren Informationen“ [8]. Der Duden nennt als Bedeutung „(durch Beobachtungen, Messungen, statistische Erhebungen u. a. gewonnene) [Zahlen]werte, (auf Beobachtungen, Messungen, statistischen Erhebungen u. a. beruhende) Angaben, formulierbare Befunde“ [9]. Das Lehrbuch der Medizinischen Dokumentation [10] definiert Daten als eine „formalisierte, interpretierbare Repräsentation von Information, die geeignet ist für die Kommunikation, die Interpretation und die Weiterverarbeitung … D < aten > bilden den Input und das Ergebnis aller Schritte der Informationsverarbeitung“. In der *primären* Datenerfassung werden dabei „Eigenschaften eines Objektes der äußeren Welt beobachtet und … in einer ‚Urliste‘ protokolliert“. An diese Schritte „schließt sich gegebenenfalls die Phase der … sekundäre < n > Datenerfassung an“ [10].

An diesem Punkt nun wird der *Sekundärdaten*-Begriff relevant, den wir – in nur semantischer Hinsicht, jedoch nicht in inhaltlicher Abweichung – folgendermaßen definieren wollen:

> *Sekundärdaten sind abgeleitete, meist prozessproduzierte Daten, die aus Daten mit einem anderen primären Verwendungszweck durch Verarbeitungsschritte hervorgehen und die über ihre originäre, vorrangige Zweckbestimmung hinaus – z. B. im Rahmen einer wissenschaftlichen Untersuchung – einer zweiten – sekundären – Nutzung zugeführt werden, für die sie ursprünglich nicht erhoben worden sind. Maßgeblich für die Einstufung als Sekundärdaten sind somit Unterschiede zwischen dem primären Erhebungsanlass und der nachfolgenden Nutzung. In diesem Sinne stellen auch im Rahmen der Patientenversorgung generierte klinische Daten Sekundärdaten dar, sofern sie – die klinischen Daten – einer späteren wissenschaftlichen Nutzung zugeführt werden, beispielsweise nach der Extraktion aus einer Patientenakte.*
>
> *Neben den weiterverarbeiteten Primärdaten gehören zu den Sekundärdaten auch administrative Daten wie z. B. Daten der amtlichen Statistik, Daten aus pharmakoepidemiologischen Datenbanken oder Daten aus (krankheitsbezogenen) Registern. Sekundärdaten bezeichnen somit Daten, die bereits vorhanden sind und deren Datenfelder sich aus ihrem primären Zweck ergeben, die also nicht mehr infolge einer empirischen Heuristik oder einer forschenden Fragestellung de novo angelegt werden können. Damit stellen die vielfach in der Versorgungsforschung genutzten Abrechnungsdaten der GKV oder andere Sozialdaten eine Teilmenge der Sekundärdaten dar, die Verkürzung auf eine eineindeutige Gleichsetzung von Sekundärdaten = Sozialdaten ist der Sache nach falsch.*

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

„Für die Einstufung," so heißt es in der *Guten Praxis Sekundärdatenanalyse* (GPS) [11] weiter, „ist es unerheblich, ob die weitergehende Nutzung durch den Dateneigner selbst oder durch Dritte erfolgt. Demnach sind beispielsweise Routinedaten einer Krankenkasse nicht nur Sekundärdaten, wenn sie für wissenschaftliche Fragestellungen genutzt werden, sondern z. B. auch dann, wenn sie durch die Krankenkasse für Zwecke der Versorgungsplanung herangezogen werden."

Stellen wir die oben gegebene Definition einer Auswahl existierender Definitionen des „Sekundärdaten"-Begriffs gegenüber, so ist festzustellen, dass im deutschsprachigen Raum Definitionsbestrebungen von verschiedenen gesundheitsbezogenen Fachdisziplinen unternommen werden. Dabei stehen i.d.R. spezielle Aspekte im Vordergrund, die sich aus der fachlichen Provenienz des jeweiligen Definitionsversuchs ergeben. Zwei unterschiedliche Exponenten des breiten Spektrums sollen im Folgenden exemplarisch vorgestellt werden: die Sichtweise des Arbeitsgebietes Forschungsdateninfrastruktur und Forschungsdatenmanagement [12] sowie die Perspektive der translationalen Forschung, hier erläutert am Beispiel der translationalen medizinischen Ethik [13].

Die erstgenannte Perspektive nimmt ihren Ausgangspunkt vom Begriff der Forschungsdaten, die ihrerseits von Franke (2015) [12] definiert werden als „jene Fakten, auf denen wissenschaftliche Publikationen gründen". Sekundärdaten werden sodann als Teilmenge der Forschungsdaten verstanden und anhand von klassifizierenden Merkmalen definiert. Zu den klassifizierenden Merkmalen gehören: der Charakter der Daten (quantitativ vs. qualitativ), das Format der Daten (Bilder, Multimedia, Messdaten, Born-Digitals), die Quelle der Daten bzw. die Art der Forschung (Beobachtungen vs. Experimente, Modelle oder Simulationen, abgeleitete oder kompilierte Daten, Referenzdaten oder kanonische Daten") sowie der Grad der Aggregation/Verdichtung (Rohdaten, Primärdaten, Sekundärdaten, Tertiärdaten). Im Kontext der letztgenannten Merkmalsausprägung werden Sekundärdaten schließlich definiert als „Daten, die in Prozessschritten aus Primärdaten gewonnen wurden."

Demgegenüber legt die andere exemplarisch angeführte Perspektive von Jungkunz et al. (2021) [13] den Fokus stärker auf die Risiken der sekundären Nutzung individueller klinischer Daten („secondary use of clinical data in data-gathering, non-interventional research or learning activities" – SeConts) und unternimmt den Versuch, die sich ergebenden Risiken für den Datenschutz und die informationelle Selbstbestimmung derjenigen Patienten, deren Daten verwendet werden, einem maßgeschneiderten Assessment zu unterziehen. Im Zuge dessen wird zunächst eine konzeptuelle Klärung und Definition des SeConts-Ansatzes vorgenommen, eine Liste von Forschungsarten und Lernaktivitäten erstellt sowie eine Klassifikation von SeCont-Aktivitäten vorgeschlagen. Als SeCont-Aktivitäten werden „non-interventional (observational) clinical research, quality control and improvement, or public health research" klassifiziert. Schließlich wird eine Liste mit Risikofaktoren erstellt, die die Wahrscheinlichkeit oder das Ausmaß von Schaden determinieren, der von SeCont-Aktivitäten potenziell ausgelöst wird.

Damit benennt die translationale medizinische Ethik auch ein sich vor dem Hintergrund erweiterter technischer Datennutzungsmöglichkeiten abzeichnendes Konfliktpotenzial bei der Sekundärdatennutzung, zum einen bezüglich des Datenschutzes aller Betroffenen (Patienten, Professionelle, Dritte), zum anderen bezüglich der IT-Sicherheit – dies insbesondere vor dem Hintergrund, dass sinnstiftende Sekundärdaten auf Patientenebene nach gegenwärtigem Verständnis allenfalls pseudonym, nicht jedoch anonym sein können.

Schließlich wollen wir unsere Aufmerksamkeit noch auf das internationale Schrifttum richten. Auf der Allgemeinbildungs-Website *ThoughtCo.com* findet sich folgende Definition von Crossman (2020) [14]:

> *„Bei der Sekundärdatenanalyse … werden Daten verwendet, die von einer anderen Person zu einem anderen Zweck erhoben wurden. In diesem Fall stellt der Forscher Fragen, die mit der Analyse eines Datensatzes adressiert werden, an dessen Erhebung er nicht beteiligt war. Die Daten wurden nicht erhoben, um die spezifischen Forschungsfragen des Forschers zu beantworten, sondern wurden zu einem anderen Zweck erhoben. Das bedeutet, dass ein und derselbe Datensatz für den einen Forscher ein Primärdatensatz und für einen anderen Forscher ein Sekundärdatensatz sein kann." [Übersetzung aus dem Englischen durch die Autoren]*

Diese Definition mag in erster Linie für die interessierte Allgemeinheit bestimmt sein, sie enthält jedoch auch zahlreiche relevante Hinweise für diejenigen Wissenschaftlerinnen und Wissenschaftler, die sich eingehender mit der Nutzung von Sekundärdaten beschäftigen möchten. Im weiterführenden Text auf der genannten Website wird auch ein informativer Überblick über die gängigsten Vor- und Nachteile der Sekundärdatenanalyse gegeben.

Terminologische Einordnung und Verhältnis zu verwandten Begriffen

Der Begriff Sekundärdaten wird bisweilen auch kritisch gesehen, weil der erste Wortteil *„Sekundär-"* einen minderen Wert suggerieren mag, indem er nahelegt, diese Daten seien zweitrangig, sozusagen „second best". Diese Skepsis könnte man in Anlehnung an den Claim einer erfolgreichen Werbekampagne[1] aus den USA der 1960er Jahre [15] ins Positive wenden nach dem Motto: „We try harder!" – und in der Tat: die Sekundärdatenforschung hat es bis heute aufgrund des Entstehungshintergrunds der Daten und durchaus auch wegen dieser pejorativen Konnotation des Begriffs nicht unbedingt leicht, ihre Erkenntnisse in hochrangigen Journals zu publizieren. Auch ist das pure Vorhandensein von Daten mitnichten ein Garant dafür, dass die damit vollzogenen Forschungsvorhaben sich methodisch weniger aufwändig gestalten würden, nur weil die Daten bereits „gegeben" seien. In aller Regel folgen auf die Bereitstellung der Datenbasis mannigfaltige, teils langwierige Schritte der Validierung, Plausibilisie-

[1] "When you're only No. 2, you try harder." war der brillante Claim einer Werbekampagne in den USA der 1960er Jahre, mit der ein Autovermieter, der seit seiner Gründung hinter dem Marktführer auf Platz zwei gelegen hatte, diese Position kommunikativ nutzte, indem die größeren Serviceanstrengungen, die mit der zweiten Position im Markt verbunden waren, besonders hervorgehoben wurden.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

rung, Bereinigung und Glättung, bevor aus einem Rohdatensatz ein operativer Datensatz entsteht, der für wissenschaftliche Auswertungen nutzbar ist [3, 11]. Im Zuge der Analysen ist dann weiteres methodisches Know-how gefragt, um den bereits angesprochenen typischen Problemen entgegenzuwirken und zu validen belastbaren Ergebnissen zu gelangen. Dennoch hat sich der Sekundärdatenbegriff quasi als Markenwort und Gattungsbegriff etabliert, wenngleich korrekter Weise von der *sekundären Nutzung vorhandener Daten* zu sprechen wäre, was auch die Verwendung vorhandener Forschungsdaten, etwa von scientific oder public use files abgeschlossener Forschungsprojekte (z. B. der Surveys des Robert Koch-Instituts), einschließen würde.

Konzentrieren wir uns im Folgenden auf Sozialdaten als einer wichtigen und großen Gruppe von Sekundärdaten, so resultiert das Vorhandensein dieser Daten in erster Linie daraus, dass sie im Zuge von Routineprozessen der medizinischen Leistungserbringung, -dokumentation, -abrechnung und/oder -vergütung gesammelt werden. Im deutschsprachigen Raum hat sich für die Sozialdaten der Versicherungsträger der Begriff *Routinedaten* etabliert. Derartige *Prozessdaten* (z. B. der Gesetzlichen Krankenversicherung) stellen zahlen- und volumenmäßig die größte Quelle für Sekundärdaten dar. Aber auch *Abrechnungs- bzw. Leistungsdaten* der Krankenhäuser oder der niedergelassenen Ärzteschaft, mit denen die erbrachten Leistungen „reklamiert" werden (engl. *claims data*), gehören zu den Sekundärdaten.

Diese Aspekte greift die Definition von Hoffmann & Glaeske (2011, 2017) auf, die Routinedaten charakterisiert als „prozessproduzierte, umfangreiche Informationssammlungen, die im Rahmen der Verwaltung, Leistungserbringung bzw. Kostenerstattung (z. B. bei der gesetzlichen Kranken-, Renten- oder Unfallversicherung sowie in Arzt- oder Krankenhausinformationssystemen) anfallen und elektronisch erfasst sind" [16, 17].

Für solche elektronisch erfassten Daten aus Krankenakten (engl. *charts*) haben sich im englischsprachigen Schrifttum die Termini "Electronic Medical Record" (EMR) oder "Electronic Health Record" (EHR) etabliert; sie werden in den Medical Subject Headings (MeSH) unter dem Schlüsselbegriff "Medical Records Systems, Computerized" subsumiert. Der MeSH-Term EHR wird definiert als „Medien, die den Transport von relevanten Informationen über die Krankheit des Patienten über verschiedene Anbieter und geografische Standorte hinweg erleichtern. Einige Versionen enthalten direkte Verknüpfungen zu Gesundheitsinformationen für Verbraucher, die für den Gesundheitszustand und die Behandlung eines bestimmten Patienten relevant sind" (Übersetzung durch die Autoren) [18].

Routinedaten können in der Tat vergleichsweise aktuelle Gesundheitsinformationen für Verbraucher, aber auch für andere Stakeholder bereitstellen, was darauf beruht, dass sie fortlaufend erhoben werden, und wenn dies kontinuierlich bei einer Vielzahl von Leistungserbringern geschieht, entstehen rasch große Datenkörper, so dass Routinedatensätze in aller Regel sehr umfangreich sind. Sie haben zudem den Vorteil, dass sie nicht durch systematische Fehler verzerrt sind, wie sie bei Primärdatenerhebungen durch mehr oder weniger restriktive Ein- und Ausschlusskriterien bedingt sein können. Da sie zum Zeitpunkt ihrer sekundären Gewinnung für wissenschaftliche Fragestellungen bereits für die Vergangenheit abgeschlossen vorliegen, unterliegen Routinedaten nicht einem analysebedingten Selektionsbias. Zu geeigneten Datenquellen, ihren Vor- und Nachteilen, den Besonderheiten und Grenzen der praktischen Nutzung von Sekundär- und Routinedaten finden sich zahlreiche Hinweise in der einschlägigen Literatur [3, 19–22]. Die noch immer anzutreffende dichotomisierende Abgrenzung von „Real World Data" und Daten aus Randomized Controlled Trials (RCT) sollte in Zukunft überwunden werden – einerseits dadurch, dass beide Studienformen miteinander verknüpft werden, indem bspw. RCT auf Basis von GKV-Daten oder Registerdaten (sog. rRCT) durchgeführt werden, andererseits dadurch, dass neue Begriffe wie „versorgungsnahe Daten" Verwendung finden [23, 24].

Ausblick

Die wissenschaftliche Arbeit mit Sekundärdaten blickt in Deutschland auf eine jahrzehntelange Tradition zurück [25, 26], die sich als tragfähig für zukünftige Entwicklungen erweisen dürfte. Der bereits eingeleitete Transformationsprozess führt in vielen aktuellen Projekten schon heute dazu, dass sich ein veränderter Umgang mit den verschiedenen verfügbaren Datenquellen abzeichnet. Dabei wird die schubladenhafte Gegenüberstellung von Primär- und Sekundärdaten durch eine Reihe teils parallellaufender Entwicklungen aufgebrochen, sodass die Grenze zwischen Primär- und Sekundärdaten – und viel wichtiger: der zugehörigen ‚Lager' der klinischen und (versorgungsepidemiologischen) Forscher – zunehmend permeabel wird. Zum einen werden, angestoßen auch durch den Innovationsfonds, immer mehr Projekte umgesetzt, in denen über Datenlinkage-Prozesse Primärdaten und Sekundärdaten miteinander verknüpft werden, zum anderen öffnet sich mit dem Digitale-Versorgung-Gesetz eine Eintrittspforte für gesundheitsbezogene digitale Applikationen, die nicht nur die Nutzung elektronischer Datenformate von und für Anwender/inne/n vorantreiben, sondern auch kontinuierlich routinemäßig Individualdaten erheben, dokumentieren, archivieren und einer Auswertung zugänglich machen. An diesem Punkt die bisher geltende Abgrenzung von Primär- und Sekundärdaten aufrecht zu erhalten, dürfte schwierig werden. Für Routine- oder Sekundärdatenforscher/innen geht davon der Impuls aus, neue Möglichkeiten der Analyse kombinierter Datensätze methodisch zu erschließen. Die Versorgungsforschung – und damit auch die Versorgung per se – dürfte von der sich abzeichnenden Allianz der Disziplinen Primär- und Sekundärdatenforschung jedenfalls profitieren [27].

Interessenkonflikt

Die Autorinnen/Autoren geben an, dass kein Interessenkonflikt besteht.

Literatur

[1] Spitz M. Daten – das Öl des 21. Jahrhunderts? Nachhaltigkeit im digitalen Zeitalter. Hamburg: Hoffman und Campe; 2017

[2] Kluge F. Etymologisches Wörterbuch der deutschen Sprache. 25., durchgesehene und erweiterte Auflage, bearbeitet von E. Seebold. Berlin und New York: Verlag Walter de Gruyter; 2011

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

[3] Swart E, Ihle P, Gothe H et al. Hrsg. Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. 2., vollständig überarbeitete und erweiterte Auflage. Bern: Verlag Hans Huber; 2014

[4] Hernán MA, Alonso A, Logan R et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology 2008; 19: 766–779

[5] Hernán MA, Robins JM. Per-Protocol Analyses of Pragmatic Trials. N Engl J Med 2017; 377: 1391–1398

[6] Matteucci Gothe R, Buchberger B. Bias und Confounding. In: Swart E, Ihle P, Gothe H et al. Hrsg. Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. 2., vollständig überarbeitete und erweiterte Auflage. Bern: Verlag Hans Huber; 2014: 424–434

[7] Wahrig G. Deutsches Wörterbuch. Völlig überarbeitete Neuausgabe, bearbeitet von U. Hermann. Gütersloh: Bertelsmann Lexikon Verlag; 1993: 329

[8] Roche Lexikon Medizin. München: Urban & Schwarzenberg; 1987

[9] Duden-Redaktion. Berlin: Bibliographisches Institut GmbH, Dudenverlag; 2021. URL: https://www.duden.de/rechtschreibung/Daten [Zugriff am 13.05.2021]

[10] Leiner F, Gaus W, Haux R et al. Medizinische Dokumentation – Grundlagen einer qualitätsgesicherten integrierten Krankenversorgung. Lehrbuch und Leitfaden. Kap. 12: Thesaurus der Grundbegriffe der Medizinischen Dokumentation. 6., überarbeitete Auflage. Stuttgart: Schattauer; 2012: 186: 191

[11] Swart E, Gothe H, Ihle P et al. Gute Praxis Sekundärdatenanalyse (GPS): 2. Revision vorgelegt [Version 2012/2014]. Das Gesundheitswesen 2015; 77: 120–126

[12] Franke M.Forschungsdatenmanagement. Foliensatz zu einer Lehrveranstaltung an der Fachhochschule für öffentliche Verwaltung und Rechtspflege in Bayern, Sommersemester 2015. URL: https://docplayer.org/5069283-Forschungsdatenmanagement.html [Zugriff am 19.09.2021]

[13] Jungkunz M, Köngeter A, Mehlis K et al. Secondary Use of Clinical Data in Data-Gathering, Non-Interventional Research or Learning Activities: Definition, Types, and a Framework for Risk Assessment. J Med Internet Res 2021; 23: e26631

[14] Crossman A. Pros and Cons of Secondary Data Analysis. ThoughtCo, August 15, 2011. Updated June 27, 2020. URL: https://www.thoughtco.com/secondary-data-analysis-3026536 [Zugriff am 19.09.2021]

[15] Stevenson S. 'We try harder': The story of most brilliant ad slogan of the 20th century. New Haven Register. August 13, 2013, updated July 26, 2017. URL: https://www.nhregister.com/connecticut/article/We-try-harder-The-story-of-most-brilliant-ad-11418823.php [Zugriff am 13.05.2021]

[16] Hoffmann F, Glaeske G. Analyse von Routinedaten. In: Pfaff H, Neugebauer EAM, Glaeske G, Schrappe M. Lehrbuch Versorgungsforschung. Systematik – Methodik – Anwendung. 1. Auflage Stuttgart: Schattauer; 2011: 317–322

[17] Hoffmann F, Glaeske G. Analyse von Routinedaten. In: Pfaff H, Neugebauer EAM, Glaeske G, Schrappe M. Lehrbuch Versorgungsforschung. Systematik – Methodik – Anwendung. 2., vollständig überarbeitete Auflage. Stuttgart: Schattauer; 2017: 122–133

[18] NCBI – National Center for Biotechnology Information, U.S. National Library of Medicine Hrsg. MeSH, 2010. Rockville Pike, Bethesda MD, USA. URL: https://www.ncbi.nlm.nih.gov/mesh/?term = electronic + medical + record

[19] Schubert I, Köster I, Küpper-Nybelen J et al. Versorgungsforschung mit GKV-Routinedaten – Nutzungsmöglichkeiten versicherten-bezogener Krankenkassendaten für Fragestellungen der Versorgungsforschung. Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz 2008; 51: 1095–1105

[20] Gothe H. Pharmakoepidemiologie – Nutzung der Arzneimittelverordnungsdaten. Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz 2008; 51: 1145–1154

[21] Schöffski O.Quantitative Versorgungsforschung mit Sekundärdaten. Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Lehrstuhl für Gesundheitsmanagement, 18.06.2018. URL: https://www.lgl.bayern.de/gesundheit/gesundheitsversorgung/informationsplattform_versorgungsforschung/akteure/faugesman_schwpkt2.htm [Zugriff am 13.05.2021]

[22] Crossman A. Pros and Cons of Secondary Data Analysis – A Review of the Advantages and Disadvantages in Social Science Research. ThoughtCo, 27.08.2020. URL: https://www.thoughtco.com/secondary-data-analysis-3026536 [Zugriff am 13.05.2021]

[23] Klinkhammer-Schalke M, Kaiser T, Apfelbacher C et al. Manual für Methoden und Nutzung versorgungsnaher Daten zur Wissensgenerierung. Gesundheitswesen 2020; 82: 717–722

[24] Hoffmann F, Kaiser T, Apfelbacher C et al. Versorgungsnahe Daten zur Evaluation von Interventionseffekten: Teil 2 des Manuals. Gesundheitswesen 2021; 83: 470–480

[25] Schwefel D, John J, Potthoff P et al. Diagnosenstruktur in der ambulanten Versorgung. Explorative Auswertungen. Berlin, Heidelberg: Springer Verlag; 1987

[26] Ferber von L, Behrens J. Hrsg. Public Health Forschung mit Gesundheits- und Sozialdaten. Stand und Perspektiven. St. Augustin: Asgard Verlag; 1997

[27] Swart E, Gothe H, Hoffmann F et. al. Jetzt die Weichen stellen für ein leistungsfähiges Forschungsdatenzentrum Gesundheit. Gesundheitswesen 2021; 83 (Suppl. 2): S139–S141

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.