



Routinedaten – das ungenutzte Potenzial in der Versorgungsforschung

Djalali, Sima ; Markun, Stefan ; Rosemann, Thomas

Abstract: Routinedaten entstehen als Nebenprodukt des normalen Betriebsalltags. Diese ohne Zusatzaufwand massenhaft generierten Daten lassen sich mit ökonomischen oder im Falle der Medizin auch mit gesundheitlichen Fragestellungen auswerten. Durch die nicht-kontrollierte Art der Sammlung von Routinedaten sowie aufgrund der uneinheitlichen Handhabung von elektronischen Krankengeschichten ergeben sich aber je nach Fragestellung Grenzen der Aussagekraft von Resultaten. Dieser Artikel widmet sich den Fragen rund um Herkunft, Verarbeitung, Interpretation und Nutzen von Routinedaten. Dabei werden auch Machine Learning und Big Data kritisch in den Kontext gebracht, sowie die Aspekte Datenschutz und Ethik. Hinsichtlich des Schweizer Gesundheitssystems zeigt der Artikel die notwendigen Voraussetzungen, damit das Potential von Routinedaten auch hierzulande zugunsten einer besseren medizinischen Versorgung genutzt werden kann. Today, a huge amount of data is created when medical care is provided, either in a hospital setting or in ambulatory care. Even if those data are not collected with the purpose to answer scientific questions, they can be used to this effect. The fact that data collection differs in quality and extent from setting to setting and no standardization regarding the documentation exists, means that the use of these data is often limited. Switzerland is one of the countries with the highest variety regarding Electronic Medical Records (EMR), nevertheless, routine data can be used if some standardization is used in aggregating and summarizing these data. They can provide important information about the extent as well as the quality of care and contribute to improve the health care system.

DOI: <https://doi.org/10.1024/1661-8157/a002630>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-136611>

Journal Article

Accepted Version

Originally published at:

Djalali, Sima; Markun, Stefan; Rosemann, Thomas (2017). Routinedaten – das ungenutzte Potenzial in der Versorgungsforschung. *Praxis*, 106(7):365-372.

DOI: <https://doi.org/10.1024/1661-8157/a002630>

Routinedaten – das ungenutzte Potenzial in der Versorgungsforschung

Sima Djalali, Stefan Markun, Thomas Rosemann
Institut für Hausarztmedizin, Universität Zürich

Zusammenfassung

Routinedaten entstehen als Nebenprodukt des normalen Betriebsalltags. Diese ohne Zusatzaufwand massenhaft generierten Daten lassen sich mit ökonomischen oder im Falle der Medizin auch mit gesundheitlichen Fragestellungen auswerten. Durch die nicht kontrollierte Art der Sammlung von Routinedaten, sowie aufgrund der uneinheitlichen Handhabung von elektronischen Krankengeschichten ergeben sich aber je nach Fragestellung Grenzen der Aussagekraft von Resultaten. Dieser Artikel widmet sich den Fragen rund um Herkunft, Verarbeitung, Interpretation und Nutzen von Routinedaten. Dabei werden auch Machine Learning und Big Data kritisch in den Kontext gebracht, sowie auch die Aspekte Datenschutz und Ethik. Hinsichtlich des Schweizer Gesundheitssystems zeigt der Artikel die notwendigen Voraussetzungen, damit das Potential von Routinedaten auch hierzulande zugunsten einer besseren medizinischen Versorgung genutzt werden kann.

Summary

Today, a huge amount of data is created, when medical care is provided, regardless if in a hospital setting or in ambulatory care. Even those data are not collected with the purpose to answer scientific questions, they can be used to do so. The fact that the data collection differs in quality and extent from setting to setting and no standardisation regarding the documentation exists, the use of these data is often limited. Switzerland is one of the countries with the highest variety regarding Electronic Medical Records (EMR), but nevertheless, routine data can be used if some standardisation is used in aggregating and summarising these data. They can provide important information about the extent as well as the quality of care and contribute to improve the health care system.

Die Auswertung von Routinedaten, die im Betriebsalltag anfallen, gehört in der Industrie zum Standard. In der medizinischen Forschung wird dieses Feld erst langsam erschlossen. Dabei sind Routinedaten eine der bedeutendsten Datenquellen für die Versorgungsforschung. Diese Arbeit gibt einen Überblick über den Stand der Forschung mit Routinedaten im Schweizer Gesundheitswesen.

Im Unterschied zu Studien- und Registerdaten werden Routinedaten nicht primär zur Beantwortung einer spezifischen Forschungsfrage erhoben. Vielmehr entstehen Routinedaten automatisch im Betriebsalltag einer jeden Arbeitsstätte – daher Routine. Sie werden entweder als Bestandteil der Wertschöpfungskette generiert (funktionale Daten; Beispiel: Labordaten, generiert im Zuge einer ärztlichen Untersuchung) oder durch die Administration der Wertschöpfungskette (administrative Daten; Beispiel: Krankenkassenabrechnung der Untersuchung). Primär dienen Routinedaten dem operativen Geschäftsbetrieb. Gesammelt und statistisch aufbereitet, können Routinedaten jedoch einen sekundären Nutzen entfalten: Sie ermöglichen es, die Betriebsleistung zu monitorieren und daraus Erkenntnisse zur Performanceoptimierung abzuleiten. Industrieunternehmen machen sich dies seit längerem zu Nutze. Mit Hilfe sogenannter Business-Intelligence-Verfahren, d.h. Methoden zur systematischen Sammlung, Darstellung und Analyse, werden vor allem finanzbezogene Routinedaten ausgewertet, um Entscheidungsgrundlagen für die strategische Unternehmensführung zu schaffen.

Informationsgehalt von Routinedaten

Auch Leistungserbringer im Gesundheitswesen nutzen auf Organisationsebene heute oft Business-Intelligence, um ihre Finanzen zu kontrollieren. Routinedaten im Gesundheitswesen können jedoch für weit mehr als ökonomische Individualzwecke sekundär genutzt werden, nämlich für die medizinische Forschung, speziell die Versorgungsforschung.

Die funktionalen Routinedaten im Gesundheitswesen (klinische Parameter, Anamnesen, Befunde, Diagnosen, Behandlungsprotokolle etc.) widerspiegeln den Gesundheitszustand, Krankheitsverlauf, Behandlungsprozess und -outcome individueller Patienten, Gruppen und Bevölkerungen; je nachdem auf welcher Ebene des Gesundheitssystems die Daten aggregiert werden (Mikro-, Meso- oder Makroebene). Sie enthalten Variablen, die eine Untersuchung ätiologischer, pathogenetischer, epidemiologischer, diagnostischer und therapeutischer Fragestellungen erlauben.

Die administrativen Routinedaten im Gesundheitswesen (Abrechnungen, Kostengutsprachen, Niederlassungsbewilligungen, Patienten- und Leistungserbringerstammdaten etc.) können indirekt ebenfalls Gesundheitszustände, Krankheitsverläufe, Behandlungsprozesse und -

outcomes widerspiegeln – allerdings meist unpräziser als funktionale Daten, weil medizinische Detailinformationen nicht oder nur grob zusammengefasst (z.B. Tarmed-Positionen, Diagnosis related Groups [DRG]) enthalten sind. In erster Linie widerspiegeln die Variablen administrativer Routinedaten die Infrastruktur, Ressourcen, Inanspruchnahme und Wertschöpfung im Gesundheitssystem, je nach Aggregationslevel auf Mikro-, Meso- oder Makroebene. Damit erlauben sie die Untersuchung ökonomischer und organisatorischer Fragestellungen.

Nutzen in der Forschung

Unter Zuhilfenahme des Drei-Säulen-Modells der Forschung in der Medizin (nach Pfaff und Schrappe 2011 [1]) kann der Nutzen von Routinedaten für die medizinische Wissenschaft leicht eingestuft werden. Die erste tragende Säule der medizinischen Forschung ist die Grundlagenforschung, welche heute vor allem auf molekularer Ebene stattfindet. Diese Forschung findet im Labor unter Experimentalbedingungen statt und untersucht die grundsätzliche Wirksamkeit (*Efficacy* unter Laborbedingungen). Routinedaten aus dem Gesundheitswesen haben hier keinen Stellenwert.

Die zweite Säule ist die klinische Forschung, deren Ziel es ist, die konkrete Anwendung der erarbeiteten Grundlagenkenntnisse am Menschen zu prüfen und die grundsätzliche Sicherheit und Wirksamkeit neu entwickelter Behandlungen zu beurteilen (*Efficacy* unter kontrollierten klinischen Bedingungen). Hier liegt der Nutzen der funktionalen Routinedaten auf der Hand. Sie enthalten medizinische Informationen, die andernfalls – im Rahmen klassischer prospektiver Interventions- und Observationsstudien – extra mit viel Aufwand gesammelt werden müssen. Auf die methodischen und statistischen Implikationen, die damit einhergehen, wird später eingegangen (siehe Abschnitt «Methodische und statistische Implikationen»). Fest steht: Funktionale Routinedaten aus dem Gesundheitswesen können genutzt werden, um den Ablauf klinischer Studien zu vereinfachen und können die doppelte Datenerfassung für Forschungsregister überflüssig machen, weil die relevanten Daten aus Routinedatensätzen entnommen werden können. Administrative Routinedaten eignen sich dafür weniger. Da sie, wie oben dargelegt, nicht selbst Träger medizinischer Informationen sind, können sie bestenfalls als Proxydaten (indirekte Anzeiger) dienen. Zur Beantwortung klinischer Fragestellungen auf hohem Evidenzniveau ist das nicht ausreichend.

Die dritte Säule der medizinischen Forschung ist die Versorgungsforschung, deren Aufgabe es ist, die Wirksamkeit (*Effectiveness*) und das Kosten-Nutzen-Verhältnis (*Efficiency*) der medizinischen Versorgung im Alltag zu überprüfen, die Rahmenbedingungen des Gesundheitssystems zu beschreiben, deren Einflüsse kausal zu erklären sowie auf Basis der Erkenntnisse effizientere Versorgungskonzepte zu entwickeln und zu implementieren. Hier

stellen funktionale wie administrative Routinedaten die essenzielle Datengrundlage für die Beschreibung der IST-Situation in einem real betriebenen Gesundheitssystem dar. Sie liefern Protokoll über die Funktionsweise des Gesundheitssystems, Versorgungsstrukturen, -prozesse und Outcomes unter Alltagsbedingungen. Zudem bieten sie die Möglichkeit, den Einfluss von Interventionen (auf Mikro-, Meso- und Makroebene) kontinuierlich zu monitorieren und so zu kontrollieren ob sich IST- und SOLL-Zustand einander annähern. Idealerweise lassen sich funktionale und administrative Routinedaten verknüpfen und ermöglichen es damit, konkrete Assoziationen zwischen Prozessen, medizinischen Outcomes und Kosten zu untersuchen.

Diskrepanz zwischen Theorie und Realität

Diese Einstufung des Nutzens von Routinedaten für die medizinische Forschung basiert freilich auf einem Modell, das eine Reihe idealer Eigenschaften von Routinedaten im Gesundheitswesen voraussetzt. Sie müssen:

- elektronisch vorliegen
- valide sein

und, idealerweise

- über verschiedene Datenquellen hinweg verknüpfbar sein.

In der Realität sind diese Voraussetzungen oft nicht gegeben. Dies limitiert zum einen die Verfügbarkeit der Daten für wissenschaftliche Zwecke. Nur ein Bruchteil der heute vorhandenen Routinedaten im Gesundheitswesen ist überhaupt für die Forschung zugänglich. Zum anderen limitiert es die Evidenz, die aus den verfügbaren Daten gewonnen werden kann: Nur ein Bruchteil der heute vorhandenen Routinedaten kann sinnvoll interpretiert werden. Diese Limiten begründen die Hauptkritik, die heute an Routinedaten im Allgemeinen, und im Besonderen an existierenden Forschungsprojekten auf Basis von Routinedaten, geübt wird.

Elektronische Verfügbarkeit

Dass Routinedaten elektronisch vorliegen müssen, um in den beschriebenen Größenordnungen mit vertretbarem Aufwand extrahiert und verarbeitet werden zu können, dürfte unbestritten sein. Allerdings ist elektronische Datenhaltung weltweit in einigen Bereichen des Gesundheitswesens noch nicht die Regel. Länder wie England, die Niederlande, Schweden, Dänemark, Israel und Kanada – notabene alles Länder mit überwiegend staatlich finanziertem Gesundheitswesen – sind schon sehr weit in der Digitalisierung vorangeschritten und damit auch in der sekundären Nutzung ihrer Routinedaten. Doch sind auch dort längst nicht alle Bereiche des Gesundheitswesen

abgedeckt [2-9]. Hinzu kommt, dass «elektronische Datenerfassung» nicht gleichbedeutend ist mit «strukturierten» elektronischen Daten. Relevante Informationen sind sehr häufig als Freitext erfasst und können daher nicht ohne weiteres für Datenanalysen extrahiert und verarbeitet werden [10]. Die Schweiz liegt im internationalen Vergleich hinsichtlich der allgemeinen Digitalisierung und konsekutiven Nutzung der abfallenden strukturierte Routinedaten noch zurück [11]. Vor allem funktionale Routinedaten werden hierzulande von vielen Leistungserbringern gar nicht oder nur teilweise digital erfasst. Insbesondere im ambulanten Gesundheitswesen besteht ein Missverhältnis zwischen vielen elektronisch erfassten, *administrativen* Routinedaten gegenüber wenigen elektronisch erfassten *funktionalen* Routinedaten und einheitliche Standards für Datenstruktur sind kaum implementiert [12-14].

Der Zugang zu den existierenden elektronischen Daten wird zudem häufig durch Datenschutzgesetze und Hoheitsentscheide der datengenerierenden Institutionen beschränkt [5,15].

Einschränkungen der Validität

Dass Daten valide sein müssen, um gültige Forschungsergebnisse zu erlangen, sollte auch unbestritten sein, ist bei Routinedaten aber schwierig sicherzustellen. Das liegt am Wesen ihrer informellen Erhebung. Während bei der Durchführung prospektiver Studien ein Grossteil des Arbeitsaufwands darin besteht, die Datenerhebung zu standardisieren und Störeinflüsse (*Confounder*) zu minimieren, unterliegt die Routinedatengenerierung üblicherweise keiner Kontrolle. Ein Routinedatensatz ist daher immer das Produkt der mehr oder weniger unterschiedlichen Arbeitsroutinen der Werk tätigen, die zur Entstehung der Routinedaten beigetragen haben; hohe Inter- und Intraobserver-Variabilität inklusive. Objektivität, Reliabilität und damit schlussendlich Validität der in den Daten abgebildeten Messungen und Schätzungen sind daher fraglich.

Ein Beispiel: Extrahiert man Blutdruckwerte aus elektronischen Krankengeschichten, so können die eingetragenen Werte von den Untersuchern im Alltag auf höchst unterschiedliche Weise gemessen worden sein: Auskultatorisch, oszillatorisch, intravasal, im Sitzen, Stehen oder Liegen; basierend auf einer Messung, drei Messungen, einem 24-Stunden-Durchschnitt, etc. Jedes Messverfahren und jeder Untersucher haben unterschiedliche Fehlerwahrscheinlichkeiten. Obwohl alle Werte als «Blutdruckdaten» klassifiziert sind, sind sie aufgrund der inhomogenen Erhebungsweisen nur eingeschränkt miteinander vergleichbar. Einen auskultatorisch, im Sitzen, einmalig gemessenen Blutdruckwert als gleiche Entität anzusehen wie einen oszillatorisch, im Liegen gemessenen 24-Stunden-Durchschnittswert, mag bei gewissen Forschungsfragen («Blutdruckmessung vorgenommen ja/nein?») vertretbar

sein, stellt bei anderen («Wirkungsvergleich Antihypertensiva») jedoch eine relevante methodische Schwäche dar. Da die Erhebungsweisen meist unbekannt sind, ist keine Konvergenzvalidität (Messdaten von Testverfahren, die dasselbe Konstrukt abbilden, müssten hoch miteinander korrelieren [16]) garantiert. Dies schmälert insbesondere die Verwendbarkeit von funktionalen Routinedaten für klinische Forschungsfragen [17].

Im administrativen Bereich des Gesundheitswesens sind die Arbeitsroutinen der Werk tätigen tendenziell stärker standardisiert als in der Patientenversorgung, bedingt durch systemweit gültige Gesetze und Verordnungen, die den Rahmen der administrativen Tätigkeit vorgeben. Es ist daher prinzipiell mit weniger Variabilität bei der Datengenerierung zu rechnen als bei funktionalen Daten. Zum Beispiel geben Abrechnungskataloge (wie Tarmed/DRG) Kriterien vor, wie Leistungen abgerechnet werden. Dieselbe Leistung sollte prinzipiell also unabhängig von Vorlieben, Tagesform und Instruktionen der Leistungserbringer über dieselben Pauschalen abgerechnet werden, was die Zuordnung von Leistung zu Pauschale objektiv und reliabel macht. In der Realität besteht allerdings einiger Spielraum, weil die administrativen Regelwerke – um im Alltag für ihren primären Zweck ausführbar zu bleiben – relativ unspezifisch und generisch sein müssen. Um beim Abrechnungsbeispiel zu bleiben: Nicht jede Leistung ist durch eine spezifische Abrechnungsposition abgebildet, sondern geht in einer Sammelbezeichnung auf (Beispiel: Blutdruckmessung im Tarmed-Katalog unter «Kleine Untersuchung») oder wird durch Konjugation mehrerer unspezifischer Einzelleistungen (Leistungsblöcke) verrechnet. Das Auftauchen einer Sammelbezeichnung oder eines Blocks im Datensatz als Beweis für die Durchführung einer spezifischen Leistung zu werten, ist jedoch fehleranfällig. Für die Sekundärnutzung der administrativen Routinedaten bedeutet dies, dass die sogenannte Inhaltsvalidität der Daten – je nach Fragestellung, die beantwortet werden soll – zu gering ist. Inhaltsvalidität ist gegeben, wenn ein Verfahren zur Messung eines bestimmten Konstrukts (hier «Blutdruckmessung») oder Merkmals die bestmögliche Operationalisierung dieses Konstrukts ist [18]. Bei einer prospektiven Studie zum Vergleich, würde im Vorfeld der Datenerhebung, bei der Studienplanung, viel darauf verwendet, die zu untersuchenden Konstrukte zu operationalisieren und wenn möglich nach Goldstandard zu messen.

Mangelnde Konvergenz- und Inhaltsvalidität sind allerdings nur zwei der Probleme, die den wissenschaftlichen Wert von Routinedatensätzen beeinträchtigen können. Die häufigsten Validitätsprobleme, die aus der fehlenden Kontrolle auf die Routinedatenentstehung resultieren, sind fehlende und falsche Daten [19,20]. Dies liegt vor allem an der noch verbesserungswürdigen Implementierung vieler Informatiksysteme in den Arbeitsalltag des Gesundheitswesens. Viele Informatiksysteme sind suboptimal an die Erfordernisse des Arbeitsalltags im Gesundheitswesen angepasst. Vice versa behalten viele Nutzer ihre Papiererprobten Arbeitsabläufe bei und nutzen ihre Informatiksysteme dabei anders als vorgesehen. Daraus folgen Work-arounds, absichtliches Freilassen oder Fremdnutzen von Datenfeldern

und Medienbrüche bei der Datenübermittlung; Hinzu kommen schlichte menschliche Fehler bei der Datenerfassung, die durch benutzerunfreundliches Softwaredesign befördert werden und mangels Eingabekontrollen unbemerkt bleiben [21].

Verknüpfung von Datenquellen

Je nachdem auf welcher Ebene des Gesundheitssystems (Mikro-, Meso-, Makroebene) Erkenntnisse durch Versorgungsforschung mit Routinedaten gewonnen werden sollen, ist es sinnvoll bis notwendig, Daten aus verschiedenen operativen Informatiksystemen (Primärsystemen) miteinander zu verknüpfen; Einerseits, um innerhalb eines Sektors (z.B. ambulante ärztliche Versorgung) flächendeckend (z.B. in allen Arztpraxen einer Region statt nur in einer) Daten zu sammeln und damit eine möglichst hohe Repräsentativität zu erzielen. Man spricht von horizontaler Verknüpfung. Andererseits, um die real existierende Fragmentierung des Gesundheitssystems zu überwinden und Zusammenhänge in Daten aus mehreren Sektoren des Gesundheitswesens (z.B. Krankengeschichteinträge der ambulanten und stationären ärztlichen Versorgung sowie Reha und Krankenkassenrechnungen) zu betrachten. Man spricht von vertikaler Verknüpfung.

Nebst an technischen Schwierigkeiten (verschiedene Datenstrukturen, fehlende semantische Standards), scheitert die Verknüpfung von Routinedaten heute oft an Datenschutzregelungen resp. am Fehlen einer anonymen Identifikationsnummer für Personen, die das Gesundheitssystem in Anspruch nehmen.

Die technischen Schwierigkeiten erschweren vor allem die horizontale und vertikale Verknüpfung von Routinedaten, bei der es darum geht, dieselbe Art von Informationen aus möglichst vielen lokalen Informatiksystemen zu extrahieren, z.B. Blutdruckwerte von Patienten aus zehn verschiedenen Arztpraxen. Dies gelingt nur, wenn alle zehn Praxisinformationssysteme die Blutdruckwerte – vereinfacht gesagt – mit derselben Struktur und Semantik versehen. In Ländern mit hohem Entwicklungsgrad der Digitalisierung im Gesundheitswesen bestehen Standards, die eine Interoperabilität der Informatiksysteme gewährleisten. In der Schweiz bestehen solche Standards wie bereits weiter oben ausgeführt, bislang nicht. Alleine im Bereich der ambulanten ärztlichen Grundversorgung sind heute 85 verschiedene Praxisinformationssysteme am Markt vertreten [22], jedes davon mit proprietären Datenhaltungsmodellen. Die Interoperabilität dieser Systeme ist bereits untereinander äusserst eingeschränkt, von der Interoperabilität mit den zahlreichen Informatiksystemen anderer Sektoren des Gesundheitswesens (z.B. Spitexdienste) ganz zu schweigen.

Datenschutzregelungen erschweren vor allem die horizontale und vertikale Verknüpfung von Routinedaten, bei der es darum geht, Daten entlang einer Behandlungskette resp. eines

Bearbeitungsfalls zu sammeln, d.h. Daten auf Patienten-/Fallebene zu aggregieren und Assoziationen zwischen Variablen aus verschiedenen Datenquellen zu untersuchen. Anonymisierte Identifikationsnummern sind das Mittel der Wahl, um die eindeutige Zuordnungen von Datensätzen unter Aufrechterhaltung des Persönlichkeitsschutzes zu gewährleisten. «Anonymisierung» bedeutet in diesem Kontext «Unkenntlichmachung ohne Re-Identifikationsmöglichkeit der Person», wohingegen eine «Unkenntlichmachung mit erhaltener Re-Identifikationsmöglichkeit (bekannter Verschlüsselungsalgorithmus)» lediglich eine «Pseudonymisierung» darstellt. Sozialversicherungsnummern wie die AHV-Nummer in der Schweiz, ermöglichen technisch gesehen auch eine einwandfreie Datenverknüpfung, solange sie personenspezifisch und einmalig sind. Sie stellen jedoch lediglich ein Pseudonym dar, da die Zuordnung von Person und Sozialversicherungsnummer innerhalb der primären Informatiksysteme offen liegt. In Ländern, in denen dies nicht im Widerspruch zu geltenden Datenschutzgesetzen steht (z.B. Dänemark), wird die Sozialversicherungsnummer tatsächlich regelmässig als Identifikator bei der Verknüpfung von Routinedaten benutzt. In anderen Ländern müssen dagegen spezielle Register (Master Patient Index) und politische Vorkehrungen geschaffen werden, um Identifikatoren zu generieren, die dem Anspruch der Anonymisierung standhalten [23]. In der Schweiz wird die Entwicklung von Master-Patient-Indices erst langsam im Rahmen der Umsetzung des Elektronischen Patientendossiers (EPD) in Angriff genommen und ist aktuell noch nicht so weit fortgeschritten, dass die Versorgungsforschung davon profitieren könnte [24,25].

Methodische und statistische Implikationen

Weil die idealen Voraussetzungen in der Realität nicht vollumfänglich gegeben sind, stellen die Routinedaten heute in den meisten Fällen eine nicht-kontrollierte, lückenhafte und selektierte Observation dar. Ergo besteht die Gefahr, dass darauf basierende Forschungsergebnisse einem systematischen Fehler (*bias*) unterliegen. Es wäre jedoch falsch, Routinedaten darum per se eine mangelnde wissenschaftliche Güte zu attestieren. Unter methodischen Vorkehrungen bei der Studienplanung und Wahl geeigneter statistischer Kontrollverfahren können Ergebnisse von hohem wissenschaftlichen Wert aus Routinedaten gewonnen werden.

Wesentlich ist die Kenntnis der Arbeitsumstände, in denen die Routinedaten primär entstehen, um die Validitätseinschränkungen der Daten abschätzen zu können. Statistische Verfahren ermöglichen bis zu einem gewissen Grad eine Kompensation der Einschränkungen, z.B. Ausgleich fehlender Werte (Missing value imputation), Korrektur von Zeitreihen- und Cluster-Effekten (Mixed Models) und Ersatz für fehlende Randomisierung (Propensity Score Matching).

Unter der Voraussetzung, dass die Datenerfassung in den Primärsystemen durch Interventionen (z.B. Schulung von Mitarbeitern, Programmierung von Kontrollfunktionen) kontrolliert und standardisiert werden und bedeutende umweltgegebene Selektionsfaktoren (z.B. Soziodemographie in der Umgebung eines Spitals, dessen Krankenakten ausgewertet werden) quantifiziert werden können, ist es sogar denkbar die Datensammlung im Rahmen prospektiver und randomisierter Studien zu vereinfachen, in dem auf Routinedatenquellen zurückgegriffen wird [26,27].

Letztendlich unterscheidet sich die wissenschaftliche Arbeit mit Routinedaten inhaltlich nicht stark von der Forschung mit klassisch erhobenen Studiendaten. Auch im klassischen Setting kommt der Forscher nicht umhin, die Datenerhebung zu planen, Störfaktoren zu antizipieren und ein entsprechend angepasstes statistisches Modell zu entwickeln. Der Unterschied ist, dass bei einer klassischen Studie nur eine begrenzte Anzahl spezifischer, auf das Studienziel genau zugeschnittener Variablen erhoben wird. Die Kunst des Studiendesigns besteht darin, die richtigen Variablen zu definieren. Bei der Forschung mit Routinedaten besteht die Kunst darin, aus einer bestehenden Flut von Variablen hypothesengeleitet die bestmögliche Wahl zu treffen.

Die Evaluation und Validierung von Routinedatenquellen und (Weiter-)Entwicklung der entsprechenden Forschungsmethoden ist selbst ein wichtiges thematisches Element der Versorgungsforschung – sozusagen die Grundlagenforschung in der Versorgungsforschung.

Big Data und Machine Learning

Je mehr Datensätze mit einander verknüpft werden können, desto interessanter werden theoretisch Big-Data-Ansätze für die Analyse von Routinedaten aus dem Gesundheitswesen. Das vielzitierte Schlagwort «Big Data» bezeichnet Methoden zur Speicherung und Analyse grosser und/oder komplexer inhomogener Datensätze. Welche Methoden exakt dazu gehören und welche nicht, ist nicht definiert [28]. Das Feld ist dynamisch und entwickelt sich getragen von den technologischen Entwicklungen der Informatik stetig. Man könnte auch sagen: Es steckt noch in den Kinderschuhen. So schrieb der Schweizerische Nationalfonds erst Anfang 2016 das Nationale Forschungsprogramm 75 zur «Erarbeitung wissenschaftlicher Grundlagen für die Nutzung von Big Data» aus.

Dahinter steht die Hoffnung, dass die Auswertung von Massendaten dereinst die Forschung revolutionieren werde, weil bisher unentdeckte Zusammenhänge erkannt und präzise Prädiktionsmodelle entwickelt werden können. Im Bereich der Medizin erhofft man sich nichts geringeres als Durchbrüche bei der Prävention, Diagnose und Behandlung von Krankheiten [29,30]. Dabei stellen Routinedaten aus dem Gesundheitssystem, wie wir sie heute kennen, nur eine Komponente des «Massendatenschatzes» dar, der die Evidenz der Zukunft in sich

bergen soll. Das grosse Potenzial wird in der Verknüpfung aller digitalen Lebenszeichen gesehen, die der moderne Mensch im Alltag hinterlässt: Bonuspunkte beim Einkaufen, elektronische Terminkalender, Fitnesstracker, Social-Media-Kommentare, etc. Die Liste der potenziell verwertbaren Datenspuren wächst täglich durch den Ausbau von IT-Services und des «Internet of Things» (Vernetzung und Ausstattung mit Sensoren von Alltagsgegenständen, von Stereoanlage bis Warenregal im Supermarkt).

Zu den bislang wichtigsten Methoden gehören «Natural Language Processing» und «Machine Learning». Beim Natural Language Processing geht es vereinfacht gesagt darum, Computer in die Lage zu versetzen, die menschliche Sprache über die reine Worterkennung hinaus, auch im semantischen Zusammenhang zu verstehen und trotz Mehrdeutigkeiten korrekt zu deuten. Dies kann die maschinelle Auswertung von Freitexten und Tonaufnahmen erlauben und damit die Extraktion von Informationen aus unstrukturierten Daten. Damit ist die Methode auch für die klassische Versorgungsforschung interessant.

Beim Machine Learning geht es darum, Computer durch massenhaft präsentierte Lerndaten im Erkennen von Mustern und Gesetzmässigkeiten zu schulen und das angeeignete Wissen im Sinne einer künstlichen Intelligenz auf unbekannte Daten zu übertragen. Auf diesem Gebiet eröffnen sich Möglichkeiten für die elektronische klinische Entscheidungsunterstützung, sogenannte Clinical Decision Support Systeme. Deren Entwicklung fällt unter das Thema «eHealth», ein wichtiges Teilgebiet der Versorgungsforschung.

Das gesellschaftspolitische und finanzielle Potenzial, das in Big Data gesehen wird, begründet einen Trend. Nicht nur Technik- und IT-Unternehmen wie IBM, auch Pharmafirmen – allen voran das Schweizer Unternehmen Novartis – haben längst damit begonnen, sich ein neues Standbein in der Sekundärverwertung von Routinedaten aufzubauen [31,32]. Seit 2008 hat die Anzahl der Publikationen zum Thema «Big-Data-Analysen von medizinischen Routinedaten» exponentiell zugenommen [33].

Die Kritik, die von wissenschaftlicher Seite an Big-Data-Ansätzen geübt wird, bezieht sich hauptsächlich darauf, dass die Methoden intransparent sind und die hypothesengeleitete Forschung untergraben [34]. Die Intransparenz rührt zum einen daher, dass das Funktionieren von Machine-Learning-Systemen wie z.B. neuronalen Netzwerken selbst für die daran Forschenden teils nicht nachvollziehbar ist (fehlender Einblick in die Lösungswege der Maschine); zum anderen, daher, dass bislang keine Reporting-Standards für Big-Data-Forschung existieren. Die Abkehr von der hypothesengeleiteten Forschung wird befürchtet, weil Assoziationen nicht mehr durch gezielte (experimentell reproduzierbare) Untersuchung entdeckt werden, sondern durch schrankenlose Mustererkennung in dynamischen Datensätzen. Es herrscht Skepsis, ob mehr Daten tatsächlich mehr Evidenz generieren.

Ethische Überlegungen

Die Kritik, die von sozialphilosophischer Seite an Big-Data-Ansätzen geübt wird, bezieht sich dagegen auf die gesellschaftspolitischen Folgen der umfänglichen Datenverknüpfung. Befürchtet wird der Verlust der Privatsphäre und der informationellen Selbstbestimmung des Individuums. Insbesondere wo es um die Verwertung gesundheitsbezogener Daten geht, wird eine Bedrohung für das Solidarprinzip des Gesundheitssystems gesehen. So sind beispielsweise Szenarien denkbar, in denen Menschen mit erhöhtem Krankheitsrisiko basierend auf Big-Data-Analysen Versicherungen oder Behandlungen vorenthalten werden [35].

Zweifelsohne besteht bei Routinedatensammlungen ein Missbrauchspotenzial, das mit Grösse und Informationsumfang der Sammlung zunimmt, besonders wenn die Daten nicht anonymisiert sind. Handkehrum dient gerade eine systematisch betriebene Versorgungsforschung mit Routinedaten der Aufrechterhaltung des Solidarprinzips, weil die Forschungsergebnisse die Grundlage für eine evidenzbasierte, utilitaristische Allokation der Ressourcen im Gesundheitswesen liefern.

Aus ethischer Sicht stellen weder die Verfügbarkeit noch die Verknüpfbarkeit von Routinedaten aus dem Gesundheitswesen ein Problem dar. Im Gegenteil, die möglichst umfassende Nutzbarmachung der vorhandenen Routinedaten zur Vermeidung von Über-, Unter- und Fehlversorgung ist sogar ein ethischer Imperativ. Gesellschaft und Politik stehen daher vor der Herausforderung, ethische und juristische Regeln für die Sekundärverwendung von Routinedaten aus dem Gesundheitswesen zu erlassen, die einerseits das Missbrauchspotenzial eindämmen, resp. Missbrauch unter Strafe stellen und andererseits die Freiheit der Wissenschaft und Forschung bewahren und fördern. Dazu gehört insbesondere in der Schweiz eine Datenschutzregelung, die Privatsphäre erhält, aber Anonymisierung und Verknüpfung ermöglicht, wie im Kapitel «Verknüpfung von Datenquellen» dargelegt.

Die Wissenschaft hat zur Aufgabe, methodische Standards zu entwickeln, die es ermöglichen, Routinedatenanalysen hinsichtlich ihrer Qualität und damit ihres Evidenzgrades zu differenzieren. In Anlehnung an die Standards der «Good Clinical Practice» und «Good Epidemiological Practice» für die klinische und epidemiologische Forschung haben die Deutschen Gesellschaften für Sozialmedizin und Prävention sowie Epidemiologie bereits Empfehlungen zur «Guten Praxis Sekundärdatenanalyse» und Reporting-Guidelines entwickelt, welche jedoch noch um Big-Data-Aspekte erweitert werden müssten, um zeitgemäss zu sein [36,37].

Forschung in der Schweiz

In der Schweiz steht die Forschung mit Routinedaten aus dem Gesundheitswesen erst am Anfang. Gründe dafür sind, wie oben dargelegt, vor allem die bruchstückhafte Digitalisierung der Leistungserbringer und fehlende Möglichkeiten der horizontalen und vertikalen Datenverknüpfung. Bisherige Forschungsprojekte basieren daher vornehmlich auf einer einzigen Datenquelle und sind stark abhängig von deren Limitationen. So findet die meiste Forschung aktuell mit administrativen Routinedaten statt: mit Rechnungsdaten einzelner Versicherer, Versicherungsverbände oder Rechnungsdaten und Betriebskennzahlen der Spitäler (über das Bundesamt für Statistik) [38-40]. Forschung mit funktionalen Routinedaten sind selten und meist auf Querschnitterhebungen beschränkt, da keine kontinuierliche Datensammlung betrieben wird oder werden kann [41].

Vorreiter in der Hausarztmedizin

Eine Ausnahme stellt das FIRE-Projekt (Family medicine ICPC Research using Electronic medical records) dar, das seit 2009 kontinuierlich funktionale Routinedaten aus Schweizer Hausarztpraxen sammelt und wissenschaftlich auswertet [42]. Inzwischen beteiligen sich über 200 Hausärztinnen und Hausärzte aus der Deutschschweiz an dem Projekt, das vom Institut für Hausarztmedizin der Universität Zürich geleitet wird. Sechs marktführende Hersteller von elektronischer Krankengeschichtssoftware (eKG) unterstützen das Projekt, in dem sie ihren Kunden eine Datenexportmöglichkeit bieten. Auf diese Weise können anonymisierte Daten zu jeder routinemässig dokumentierten Konsultation aus der eKG extrahiert und der Versorgungsforschung zur Verfügung gestellt werden. Das Datenset umfasst Alter und Geschlecht des Patienten, Vitalwerte (Grösse, Gewicht, Body Mass Index, Blutdruck, Puls), Laborwerte (Hämoglobin, Leukozytenzahl, C-reaktives Protein, Blutsenkungsgeschwindigkeit, Kreatinin, Gesamtcholesterin, HDL-Cholesterin, LDL-Cholesterin, Triglyzeride, Transaminasen, Glucose nüchtern, HbA1C, Prostata-spezifisches Antigen, International Normalized Ratio der Blutgerinnung), Medikamente und Vorstellungsgründe/Diagnosen (codiert nach der International Classification of Primary Care 2). Bislang liegen Daten von über 350'000 Patienten und zwei Millionen Konsultation vor. Daraus sind neun Publikationen entstanden, die einerseits die Machbarkeit der Datenextraktion belegen und Validität, aber auch Schwächen dieser Routinedaten darlegen. Andererseits werden die Versorgungswirklichkeit und Qualität der hausärztlichen Versorgung in der Schweiz dokumentiert [43-51]. Weitere Publikationen sind in Planung.

Die Ergebnisse sind vielversprechend, denn sie zeigen, dass – trotz technischer Hindernisse und mangelnder Implementierung elektronischer Arbeitsweisen im Praxisalltag – eine horizontale Verknüpfung von Routinedaten möglich ist. In umschriebenen Bereichen, in denen annähernd vollständige Daten extrahiert werden können (z.B. Gerinnungswerte [50]), sind gar Messungen der medizinischen Qualität möglich und zeigen, dass die Qualität der

hausärztlichen Versorgung im Vergleich zu Benchmarks aus der internationalen Literatur gut ist. Dennoch bedarf es einer weiteren Ausweitung und Homogenisierung der Datensammlung, um generalisierbare Ergebnisse zu erhalten. Sonst bleibt auch FIRE ein Inselprojekt.

Fazit

Vergleicht man die Anstrengungen von Industrieunternehmen, mit der die Digitalisierung der Arbeitsroutine und damit die Business-Intelligence vorangetrieben wird, erscheinen einem die bisherige Entwicklung von Digitalisierung und Routinedatenverwertung im Gesundheitswesen in der Schweiz rückständig. Die Forschung mit Routinedaten bleibt weit hinter ihrem theoretischen Potenzial zurück. Jedoch ist die Arbeitsroutine von Beschäftigten im Gesundheitswesen weit weniger standardisiert und repetitiv wie in jenen Branchen, die auf Business-Intelligence setzen. Entsprechend komplex sind die Routinedaten in Struktur und Informationsgehalt und es ist schwierig vollständige Datensätze zu erlangen, die die notwendige Informationstiefe besitzen, um Forschungsfragen mit ausreichender Präzision beantworten zu können. Dennoch sind Routinedaten in vielen Fällen die einzige Möglichkeit, um die Vorgänge im Gesundheitssystem zu untersuchen und Hypothesen zu generieren, was wie verbessert werden könnte, damit das Schweizer Gesundheitssystem seinem Ruf als «bestes Gesundheitssystem der Welt» gerecht werden kann.

Korrespondenzadresse:

Dr. Sima Djalali

Institut für Hausarztmedizin

Universität Zürich

Pestalozzistrasse 24

8091 Zürich

Referenzen:

1. Pfaff H. Lehrbuch Versorgungsforschung: Systematik - Methodik - Anwendung ; mit 19 Tabellen. Schattauer, 2011. 9783794527977.
2. Adler-Milstein J, Ronchi E, Cohen GR et al. Benchmarking health IT among OECD countries: better data for better policy. Journal of the American Medical Informatics Association : JAMIA 2014; 21: 111-116 10.1136/amiajnl-2013-001710:.
3. De Rosis S, Seghieri C. Basic ICT adoption and use by general practitioners: an analysis of primary care systems in 31 European countries. BMC Med Inform Decis Mak 2015; 15: 70 10.1186/s12911-015-0185-z:.
4. The Deloitte Centre for Health Solutions. Current eHealth data landscape within the EU. 2012 European Summit on Trustworthy Reuse of Health Data. The Deloitte Centre for Health Solutions 2012; .
5. European Commission: Overview of the national laws on electronic health records in the EU Member States and their interaction with the provision of cross-border eHealth services - Final report and recommendations. 23 July 2014; http://ec.europa.eu/health/ehealth/docs/laws_report_recommendations_en.pdf ; Accessed 14 September 2016.
6. Gheorghiu B, Hagens S. Measuring interoperable EHR adoption and maturity: a Canadian example. BMC Med Inform Decis Mak 2016; 16: 8 10.1186/s12911-016-0247-x:.
7. Rosen B, Waitzberg R, Merkur S. Israel Health System Review, European Observatory on Health Systems and Policies. Health Care Systems in Transition 2015; 17: 2-212. http://www.euro.who.int/data/assets/pdf_file/0009/302967/Israel-HiT.pdf?ua=1 ; Accessed 19 September 2016:.
8. Weiner MG, Embi PJ. Toward reuse of clinical data for research and quality improvement: the end of the beginning? Ann Intern Med 2009; 151: 359-360.
9. Deloitte & Touche Oy, Group of Companies: International review - Secondary use of health and social care data and applicable legislation. 9.5.2016; https://www.sitra.fi/julkaisut/Muut/International_review_secondary_use_health_data.pdf ; Accessed 19 September 2016.
10. Griffon N, Charlet J, Darmoni SJ. Managing free text for secondary use of health data. Yearbook of medical informatics 2014; 9: 167-169 10.15265/iy-2014-0037:.
11. Schoen C, Osborn R, Squires D et al. A Survey of Primary Care Doctors in Ten Countries Shows Progress in Use of Health Information Technology, Less in Other Areas. Health Aff Web First 2012; Nov. 15.
12. Djalali S, Ursprung N, Rosemann T et al. Undirected health IT implementation in ambulatory care favors paper-based workarounds and limits health data exchange. Int J Med Inform 2015; 84: 920-932 10.1016/j.ijmedinf.2015.08.001:.
13. Rosemann T, Marty F, Bhend H et al. Utilisation of information technologies in ambulatory care in Switzerland. Swiss Medical Weekly 2010; 140: w13088.
14. eHealth Suisse, Koordinationsorgan Bund und Kantone: Ziele und Stand von "eHealth", Version 5. Stand Mai 2016; <http://www.e-health-suisse.ch/umsetzung/00184/index.html?lang=de> ; Accessed 24 September 2016.
15. McGraw Jd DC, Leiter Jd AB. eGEMs: Pathways to Success for Multisite Clinical Data Research. eGEMs 2013; 1: 1041 10.13063/2327-9214.1041:.
16. Wikipedia Article "Konvergenzvalidität". <https://de.wikipedia.org/wiki/Konvergenzvalidit%C3%A4t> ; Accessed 21 September 2016.

17. Hersh WR, Weiner MG, Embi PJ et al. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med Care* 2013; 51: S30-37 10.1097/MLR.0b013e31829b1dbd:.
18. Wikipedia Article "Inhaltsvalidität". <https://de.wikipedia.org/wiki/Inhaltsvalidit%C3%A4t> ; Accessed 21 September 2016.
19. Haneuse S, Daniels M. A General Framework for Considering Selection Bias in EHR-Based Studies: What Data Are Observed and Why? *EGEMS (Wash DC)* 2016; 4: 1203 10.13063/2327-9214.1203:.
20. Bayley KB, Belnap T, Savitz L et al. Challenges in using electronic health record data for CER: experience of 4 learning organizations and solutions applied. *Med Care* 2013; 51: S80-86 10.1097/MLR.0b013e31829b1d48:.
21. Price M, Singer A, Kim J. Adopting electronic medical records: are they just electronic paper records? *Canadian family physician Medecin de famille canadien* 2013; 59: e322-329.
22. Amherd P. Oral presentation; Info Society Days Bern; 07/03/2013. .
23. OECD Health Policy Studies Health Data Governance Privacy, Monitoring and Research: Privacy, Monitoring and Research. 2015; http://www.keepeek.com/Digital-Asset-Management/oecd/social-issues-migration-health/health-data-governance_9789264244566-en#.WBMSazXNHwA#page3 ; Accessed 28 October 2016.
24. eHealth Suisse: Standards und Architektur Empfehlungen III - Personenidentifikation und Berechtigungssystem. 2011; <http://www.e-health-suisse.ch/umsetzung/00146/00148/> ; Accessed 28 October 2016.
25. Public Health Schweiz: Manifest - Bessere Gesundheitsdaten für ein effizienteres Gesundheitswesen. 2013; http://www.public-health.ch/logicio/client/publichealth/file/konferenz/d_Manifest_Gesundheitsdaten_Druck.pdf ; Accessed 07 December 2015.
26. Beresniak A, Schmidt A, Proeve J et al. Cost-benefit assessment of using electronic health records data for clinical research versus current practices: Contribution of the Electronic Health Records for Clinical Research (EHR4CR) European Project. *Contemporary clinical trials* 2016; 46: 85-91 10.1016/j.cct.2015.11.011:.
27. Coorevits P, Sundgren M, Klein GO et al. Electronic health records: new opportunities for clinical research. *Journal of internal medicine* 2013; 274: 547-560 10.1111/joim.12119:.
28. Ward JS, Barker A. Undefined By Data: A Survey of Big Data Definitions. *Cornell University Library* 2013; <https://arxiv.org/abs/1309.5821>; Accessed 28 October 2016.
29. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015; 372:10.1056/NEJMp1500523:.
30. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Information Science and Systems* 2014; 2: 1-10 10.1186/2047-2501-2-3:.
31. Chen Y, Elenee Argentinis JD, Weber G. IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. *Clinical therapeutics* 2016; 38: 688-701 10.1016/j.clinthera.2015.12.001:.
32. Miller J. Big Pharma's bet on Big Data creates opportunities and risks. *Reuters* 2016; <http://www.reuters.com/article/us-pharmaceuticals-data-idUSKCN0V41LY> ; Accessed 02 November 2016.
33. Andreu-Perez J, Poon CC, Merrifield RD et al. Big data for health. *IEEE journal of biomedical and health informatics* 2015; 19: 1193-1208 10.1109/jbhi.2015.2450362:.

34. Antes G. Big Data und Personalisierte Medizin: Goldene Zukunft oder leere Versprechungen? Dtsch Arztebl International 2016; 113: 712-.
35. Marconi K, Lehmann H. Big Data and Health Analytics. CRC Press, 2014. 9781482229257.
36. Swart E, Gothe H, Geyer S et al. Gute Praxis Sekundärdatenanalyse (GPS): Leitlinien und Empfehlungen. Gesundheitswesen 2015; 77: 120-126 10.1055/s-0034-1396815:.
37. Swart E, Schmitt J. [STandardized Reporting Of Secondary data Analyses (STROSA)-a recommendation]. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen 2014; 108: 511-516 10.1016/j.zefq.2014.08.022:.
38. Huber CA, Brandle M, Rapold R et al. A set of four simple performance measures reflecting adherence to guidelines predicts hospitalization: a claims-based cohort study of patients with diabetes. Patient preference and adherence 2016; 10: 223-231 10.2147/ppa.s99895:.
39. Hemkens LG, Saccilotto R, Reyes SL et al. Personalized prescription feedback to reduce antibiotic overuse in primary care: rationale and design of a nationwide pragmatic randomized trial. BMC Infectious Diseases 2016; 16: 421 10.1186/s12879-016-1739-0:.
40. Berlin C, Busato A, Rosemann T et al. Avoidable hospitalizations in Switzerland: a small area analysis on regional variation, density of physicians, hospital supply and rurality. BMC health services research 2014; 14: 289 10.1186/1472-6963-14-289:.
41. Bovier PA, Sebo P, Abetel G et al. Adherence to recommended standards of diabetes care by Swiss primary care physicians. Swiss Med Wkly 2007; 137: 173-181 2007/11/smw-11592:.
42. Homepage des Instituts für Hausarztmedizin der Universität und des Universitätsspitals Zürich: Das FIRE-Projekt. <http://www.hausarztmedizin.uzh.ch/de/fire2.html> ; Accessed 08 November 2016.
43. Chmiel C, Bhend H, Senn O et al. The FIRE project: a milestone for research in primary care in Switzerland. Swiss Med Wkly 2011; 140: w13142 10.4414/smw.2011.13142:.
44. Busato A, Bhend H, Chmiel C et al. Improving the quality of morbidity indicators in electronic health records in Swiss primary care. Swiss Med Wkly 2012; 142: w13611 10.4414/smw.2012.13611:.
45. Rizza A, Kaplan V, Senn O et al. Age- and gender-related prevalence of multimorbidity in primary care: the Swiss FIRE project. BMC Fam Pract 2012; 13: 113 10.1186/1471-2296-13-113:.
46. Djalali S, Frei A, Tandjung R et al. Swiss quality and outcomes framework: quality indicators for diabetes management in Swiss primary care based on electronic medical records. Gerontology 2014; 60: 263-273 10.1159/000357370:.
47. Zellweger U, Bopp M, Holzer BM et al. Prevalence of chronic medical conditions in Switzerland: exploring estimates validity by comparing complementary data sources. BMC public health 2014; 14: 1157 10.1186/1471-2458-14-1157:.
48. Hasse B, Tarr PE, Marques-Vidal P et al. Strong Impact of Smoking on Multimorbidity and Cardiovascular Risk Among Human Immunodeficiency Virus-Infected Individuals in Comparison With the General Population. Open Forum Infect Dis 2015; 2: ofv108 10.1093/ofid/ofv108:.
49. Streit S, Kaplan V, Busato A et al. General Practitioners' vitamin K antagonist monitoring is associated with better blood pressure control in patients with hypertension--a cross-sectional database study. BMC cardiovascular disorders 2015; 15: 47 10.1186/s12872-015-0053-x:.

50. Djalali S, Valeri F, Gerber B et al. Anticoagulation Control in Swiss Primary Care: Time in Therapeutic Range Percentages Exceed Benchmarks of Phase III Trials. *Clin Appl Thromb Hemost* 2016; 10.1177/1076029616642514:.
51. Scherz N, Valeri F, Rosemann T et al. Quality of secondary prevention of coronary heart disease in Swiss primary care: Lessons learned from a 6-year observational study. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 2016; <http://dx.doi.org/10.1016/j.zefq.2016.06.018>: