

Dimitra Panteli | Julia Röttger | Ulrike Nimptsch | Reinhard Busse

Internationale Datengrundlagen für die Versorgungsforschung – Impulse für Deutschland



Dimitra Panteli | Julia Röttger | Ulrike Nimptsch | Reinhard Busse
Internationale Datengrundlagen für die Versorgungsforschung – Impulse für Deutschland

Die Schriftenreihe *Working papers in health policy and management* wird herausgegeben von:
Prof. Dr. med. Reinhard Busse, Leiter des Fachgebietes Management im Gesundheitswesen der
Technischen Universität Berlin.

Internationale Datengrundlagen für die Versorgungsforschung – Impulse für Deutschland

März 2021 (mit Stand Dezember 2019)

Berichtsautoren: Dimitra Panteli, Julia Röttger, Ulrike Nimptsch, Reinhard Busse

Konzeption: Reinhard Busse, Dimitra Panteli

Recherche und Mitgestaltung (in alphabetischer Reihenfolge): Ulrike Nimptsch, Dimitra Panteli, Katharine Polin, Christoph Reichebner, Julia Röttger, Giada Scarpetti, Melissa Spoden, Maximilian Teitscheid, Charlotte von Rosenstiel, Erin Webb

Wissenschaftliche Beratung: Tobias Dreischulte, Jochen Gensichen, Falk Hoffmann, Lars Kroll, Bernd Mühlbauer, Georgios Raptis, Michel Wensing, Dominik von Stillfried

Internationale Fachexpertise: Sara Allin, Andrew Barnes, Alexander Degelsegger-Márquez, Tobias Dreischulte, Adina Feldman, Anna Gref, Yoon Kim, Allie Peckham, Trine Aagren Rasmussen, Yeun Sook Rho, David Rudoler, Juliet Rumball-Smith, Lene Seidler, Nanna Skovgaard, Merran Smith, Morgan Snell, Robert Verheij, PHARMO Team, Stats NZ Microdata Customer Liaison Team

Internationale Datengrundlagen für die Versorgungsforschung – Impulse für Deutschland

Autor*innen:
Dimitra Panteli
Julia Röttger
Ulrike Nimptsch
Reinhard Busse

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Die vorliegende Studie wurde vom Zentralinstitut für die kassenärztliche Versorgung (Zi) gefördert.

Universitätsverlag der TU Berlin, 2022

<https://verlag.tu-berlin.de/>

Fasanenstr. 88, 10623 Berlin
Tel.: +49 (0)30 314 76131 / Fax: -76133
E-Mail: publikationen@ub.tu-berlin.de

Diese Veröffentlichung – ausgenommen Zitate und anderweitig gekennzeichnete Teile – ist unter der CC-Lizenz CC BY 4.0 International lizenziert.
Lizenzvertrag: Creative Commons International 4.0
<https://creativecommons.org/licenses/by/4.0/>

Umschlagbild: geralt (Gerd Altman) | <https://pixabay.com/de/illustrations/netz-netzwerk-programmierung-3706551/> | Pixabay License | <https://pixabay.com/de/service/license/>

Satz/Layout: Yulia Litvinova, Dimitra Panteli, Katherine Polin, Erin Webb
Druck: docupoint GmbH

ISBN 978-3-7983-3217-1 (print)
ISBN 978-3-7983-3218-8 (online)

ISSN 2197-8123 (print)
ISSN 1869-6287 (online)

Zugleich online veröffentlicht auf dem institutionellen Repositorium der Technischen Universität Berlin:
DOI 10.14279/depositonce-10939
<http://dx.doi.org/10.14279/depositonce-10939>

Abkürzungsverzeichnis

ASCO	American Society of Clinical Oncology
BBR	Bundesamt für Bauwesen und Raumordnung
Bev-StatG	Bevölkerungstatistikgesetz
BKRG	Bundeskrebsregistergesetz
BMBF	Bundesministerium für Bildung und Forschung
BMG	Bundesministerium für Gesundheit
BVA	Bundesversicherungsamt
CCGs	Clinical Commissioning Groups (NHS England)
CDM	Common Data Model (PCORNet USA)
CMS	Centers for Medicare and Medicaid Services (USA)
CPES	Cancer Patient Experience Survey (UK)
CPR	Centrale Personregister (Dänemark)
CPRD	Clinical Practice Research Datalink (UK)
DaTraV	Datentransparenzverordnung
DALY	Disability-Adjusted Life Years
DAV	Datenannahme- und Verteilstelle
DFG	Deutsche Forschungsgemeinschaft
DIMDI	Das Deutsche Institut für Medizinische Dokumentation und Information
DLU	Data Linkage Unit (Australien)
DRD	Derived Record Depository (Kanada)
DRG	Diagnosis Related Groups
DSGVO	Datenschutz-Grundverordnung
DVG	Digitale-Versorgung-Gesetz
eDRIS	Electronic Data Research and Innovation Service (NHS Schottland)
ELGA	Elektronische Gesundheitsakte (Österreich)
ePA	elektronische Patientenakte
FDZ	Forschungsdatenzentrum
GIS	Geografisches Informationssystem (Österreich)
GKV	Gesetzliche Krankenversicherung
GP	General Practitioner
HES	Hospital Episode Statistics
HIRA	Health Insurance Review and Assessment Service (Südkorea)
ICD	International Statistical Classification of Diseases and Related Health Problems
IDI	Integrated Data Infrastructure (Neuseeland)
IGeL	Individuelle Gesundheitsleistungen
IMD	Index of Multiple Deprivation (UK)
InEK	Institut für das Entgeltsystem im Krankenhaus
INKAR	Indikatoren und Karten zur Raum- und Stadtentwicklung
IQTIG	Institut für Qualitätssicherung und Transparenz im Gesundheitswesen
KHEntgG	Krankenhausentgeltgesetz
KV	Kassenärztliche Vereinigung
LDS	Limited Data Set (US)
MBS	Medicare Benefit Schedule (Australien)
MCHP	Manitoba Centre for Health Policy (Kanada)
MHDS	Mental Health Dataset (UK)
MHRA	Medicines and Healthcare Products Regulatory Agency (UK)
MVZ	Medizinische Versorgungszentren
MyHR	My Health Record (Australien)
NHI	National Health Insurance (Südkorea)
NHIS	National Health Insurance Service (Südkorea)
NHS	National Health Service (UK)
NIHR	National Institute for Health Research (UK)

Nivel	Niederländisches Institut für Gesundheitsforschung (Nederlands Instituut voor Onderzoek van de Gezondheidszorg, Niederlande)
OECD	Organisation for Economic Co-operation and Development
ÖGIS	Österreichische Gesundheitsinformationssystem
PBS	Pharmaceutical Benefit Schedule (Australien)
PC	Primary Care
PCORI	Patient-Centred Outcomes Research Institute (USA)
PCORnet	National Patient-Centered Clinical Research Network (USA)
PDSG	Patientendaten-Schutzgesetz
PHARMO	PHARMO Institute for Drug Outcomes Research (Niederlande)
PHE	Public Health England (UK)
PHIN	Personal Health Information Number (Kanada)
PHRN	Population Health Research Network (Australien)
PIN	Persönliche Identifikationsnummer (Schweden)
PopData BC	Population Data British Columbia
PROMs	Patient Reported Outcomes Measures
PUFs	Public Use Files (US)
QoF	Quality and Outcomes Framework (UK)
QOLC	Quality of Life of Colorectal Cancer Survivor: Patient Reported Outcome Measures Survey (UK)
QOLP	Quality of Life of Cancer Survivors in England: Pilot Survey (UK)
QS	Qualitätssicherung
RatSWD	Rat für Sozial- und Wirtschaftsdaten
REGIS	Regionale Gesundheitsinformationssystem (Österreich)
ResDAC	Research Data Assistance Center (USA)
RIFs	Research Identifiable Files (US)
RKI	Robert-Koch-Institut
RKKP	Regionale Programm für die Entwicklung Klinischer Qualität (Regionernes Kliniske Kvalitetsudviklingsprogram, Dänemark)
RSD	Reha-Statistik-Datenbasis
RTDS	National Radiotherapy Dataset (UK)
RUT	Register Utiliser Tool (Schweden)
SACT	Systemic Anti-Cancer Therapy (UK)
SDLE	Social Data Linkage Environment (Kanada)
SEER	Surveillance, Epidemiology, and End Results (USA)
SGB	Sozialgesetzbuch
SPIRE	Scottish Primary Care Information Resource
THL	Finnisches Institut für Gesundheit und Wohlfahrt (Terveyden ja hyvinvoinnin laitos, Finnland)
VPN	Virtuelles Privates Netzwerk
YLL/YLD	Years of Life Lost / Years Lived with Disability
ZfKD	Zentrum für Krebsregisterdaten
ZI	Zentralinstitut für die Kassenärztliche Versorgung in Deutschland

Tabellenverzeichnis

Tabelle	Seite
Tabelle 1.1 Abgrenzung der Disziplinen innerhalb der (bio)medizinischen Forschung	2
Tabelle 1.2 Wichtigste Sekundärdatensätze für die Versorgungsforschung in Deutschland, Stand November 2019	5
Tabelle 2.1 Länderauswahl und jeweilige Einschlussbegründung	13
Tabelle 2.2 Erhebungsinstrument zur Beschreibung von Fallstudien	14
Tabelle 3.1 Übersicht der Fallstudien und deren Inhalte (Datenkategorien korrespondieren zu den Blöcken in Abbildung 2.1)	34
Tabelle 4.1 Umfang relevanter Datenbestände und Beispiele von darauf basierenden Studien	39
Tabelle 4.2 Zusammenführung von Daten in der Primärversorgung	42
Tabelle 4.3 Modelle mit einer zentralen Stelle, die verteilte Datensätze (unterschiedlicher Datenhalter) verknüpft	44
Tabelle 5.1 Übersicht über genutzte Verlinkungsmethoden nach Initiativen	50
Tabelle 6.1 Prozesse für den Datenzugriff für ausgewählte Fallstudien	52
Tabelle 7.1 Zugriffsmöglichkeiten und Patientenrechten bei elektronischen Patientenakten	57

Abbildungsverzeichnis

Abbildung	Seite
Abbildung 1.1 Forschungsstadien und Evidenzgenerierung	2
Abbildung 2.1 Analysekonzept	10
Abbildung 2.2 Mögliche Inhalte unterschiedlicher Datenkategorien	11
Abbildung 3.1 Sundhed.dk: Datenfluss und Akteure	19
Abbildung 3.2 Kernprozesse der Danish Clinical Registries	20
Abbildung 4.1 Übersicht der CMS-Datenformate nach ResDAC	40
Abbildung 5.1 Verlinkungsansatz in Schottland (eDRIS)	48
Abbildung 5.2 Verlinkungsansatz in Manitoba	49
Abbildung 6.1 Beantragung und Nutzung von Daten über SURE (Sax Institute 2019)	53

Boxen-Verzeichnis

Box	Seite
Box 3.1 Gesundheitsdaten in der IDI	24
Box 3.2 Datenbestände, die mit CPRD-Daten verlinkt werden können	30
Box 7.1 Begriffsbestimmungen nach der US-Amerikanischen Behörde für Gesundheits-IT	55

Zusammenfassung

Die Versorgungsforschung („health services research“) hat sich in Deutschland und international bereits als vielfältige Disziplin mit komplexem, transdisziplinärem Charakter etabliert. Die Fragestellungen der Versorgungsforschung umfassen neben der deskriptiven und Inferenz-orientierten Analyse entlang der drei Versorgungsqualitätsdimensionen (Strukturen, Prozesse und Outcomes) auch die Versorgungskonzeption und das Versorgungsmanagement, und somit auch die Implementierung und die Evaluation von Interventionen. Unabdingbare Voraussetzungen für die aussagekräftige Beantwortung aller dieser Fragen, vor allem vor dem Hintergrund der „doppelten Komplexität“ (d.h. komplexe Interventionen mit komplexen Kontextbedingungen), sind das Vorhandensein und die Zugänglichkeit von umfassenden Daten hoher Qualität.

Vor dem Hintergrund der dynamischen Landschaft der unterschiedlichen Datengrundlagen, die für die Versorgungsforschung von Nutzen sein könnten, und der zunehmenden Wichtigkeit von Fragestellungen, für deren Beantwortung ein breites Spektrum an Informationen erforderlich ist, ist davon auszugehen, dass die Versorgungsforschung in den kommenden Jahren auf der Verlinkung (Datenlinkage, „data linkage“), also der Verknüpfung mehrerer Datenquellen auf Basis von geeigneten gemeinsamen Schlüsselvariablen angewiesen sein wird. Die Datenquellen beinhalten Routinedaten, verbunden mit – oder verknüpfbar mit – klinischen Daten und weiteren, teils unstrukturierten Alltags- und Kontextdaten. Der Anspruch an Daten auf hohe Validität, Vollständigkeit, Aktualität, Repräsentativität und Qualität soll davon unberührt bleiben.

In Deutschland werden umfassende Daten hoher Qualität routinemäßig gesammelt, allerdings sind diese für Forschungszwecke nur bedingt geeignet und eingeschränkt verfügbar. Es bestehen viele Datenbestände nebeneinander, die jedoch stark fragmentiert sind. Mit dem Digitale-Versorgung-Gesetz (DVG) soll die Nutzbarkeit der GKV-Routinedaten für Forschungszwecke auf der Grundlage der §§ 303a-e SGB V verbessert werden, die Umsetzung steht jedoch aus.

Das Gutachten ist parallel zum DVG im Jahr 2019 entstanden und entspricht inhaltlich dem damaligen Stand (Ende 2019). Es war geplant, es in Veranstaltungen einer breiteren (Fach-)Öffentlichkeit zu präsentieren und mit dieser zu diskutieren. Die Corona-Krise hat dies nicht nur organisatorisch vereitelt, sondern auch andere (Forschungs-)Prioritäten geschaffen. Aber die Coronakrise hat auch den Bedarf nach zeitnahen Informationen zu Gesundheit und Gesundheitsversorgung exponentiell zunehmen lassen. Seit März 2020 gibt es Corona-Tagesberichte des RKI – vorher gab es noch nicht einmal Jahresberichte zur gesundheitlichen Situation, sondern Daten zu Todesursachen, Krankenhausfällen etc. wurden Jahre nach dem Zeitraum, auf den sie sich beziehen, veröffentlicht (2020 z.B. die Daten zu Hospitalisierungen im Jahr 2018) und standen zusammenhangslos nebeneinander – und viele Daten, etwa zu freien bzw. verfügbaren Betten existierten gar nicht. Jetzt müssen Krankenhäuser täglich ihre intensivmedizinische Bettenbelegung an das Intensivregister melden und alle paar Monate die Daten zu allen Krankenhausbehandlungen an das InEK, das sie zeitnah zur wissenschaftlichen Analyse zur Verfügung stellt. Dies zeigt, dass dies möglich ist – und das politische und öffentliche Interesse wird dies auch in Zukunft einfordern. Die Inhalte des Gutachtens erscheinen den AutorInnen daher noch wichtiger als vorher; sie haben sich daher entschlossen, das Gutachten nunmehr im Frühjahr 2021 in fast unveränderter Form zu veröffentlichen. Lediglich die Ausführungen zur elektronischen Patientenakte und ihrer Nutzung zu Forschungszwecken wurden entsprechend des Patientendaten-Schutzgesetzes aktualisiert.

Im internationalen Vergleich ist erkennbar, dass im Ausland zum Teil vielfältigere, validere und verallgemeinerbare Erkenntnisse aufgrund besserer Datenzugänglichkeit erzielt werden können. Daher liegt der Ursprung dieses Gutachtens in den Grundfragen, welche Fragestellungen der Versorgungsforschung im Ausland auf Basis besserer Daten besser beantwortet werden können, und auf welchen Wegen eine solche Verbesserung der Datengrundlage in Deutschland erzielt werden könnte.

Fallbeispiele interessanter Verlinkungsansätze wurden identifiziert und anhand eines vordefinierten Konzeptes analysiert. Um die Stichprobe der Vergleichsländer für das Gutachten auszuwählen, wurden drei Hauptkriterien zugrunde gelegt: eine Mischung aus Gesundheitssystemstrukturen; ein Spektrum bezüglich der Datenverfügbarkeit und -nutzung; und bekannte interessante Beispiele relevanter Initiativen. Für die Analyse der identifizierten Fallstudien aus dem internationalen Vergleich wurde ein Konzept entwickelt, welches zwischen Datenbeständen mit versorgungsrelevanten Daten i) auf Individualebene und Daten auf Nicht-Individualebene und ii) aus drei bzw. vier großen inhaltlichen Blöcken (Gesundheitsdaten, Gesundheitsversorgungsdaten, soziodemographischen bzw. ökonomischen Daten, Umweltdaten) unterscheidet. Informationen wurden für alle Vergleichsländer bzw. Fallstudien zunächst auf Basis von öffentlich zugänglichen Quellen identifiziert. Anschließend wurde pro Land ein Profil erstellt und geeigneten Länderexperten zur Validierung und Ergänzung weitergeleitet. Das Gutachten stellt eine Übersicht der eingeschlossenen Fallstudien vor (Kapitel 3), sowie im Anhang wichtige Eckdaten pro Land und Fallstudie. Es folgen eine nähere Betrachtung der Möglichkeiten der Nutzung von Routinedaten (Kapitel 4), der Verlinkungsmöglichkeiten mit und ohne eindeutiges personenbezogenes Merkmal (Kapitel 5), der unterschiedlichen Zugriffswege für Forscher (Kapitel 6), der Ausgestaltung des Einsatzes von Daten aus elektronischen Patienten- bzw. Gesundheitsakten für die Forschung (Kapitel 7) und ein Zwischenfazit (Kapitel 8). Das Gutachten endet mit einer Wunschliste für mögliche regulatorische Änderungen in Anbetracht der aktuellen politischen Diskussion (Kapitel 9).

Die dargestellten Fallstudien lassen sich grob in fünf Kategorien einteilen:

- i) große administrative Datensätze, die es aufgrund der zentralen Zulaufstelle ermöglichen, Fragen über den Versorgungsprozess einzelner Patienten sektorenübergreifend abzubilden und Möglichkeiten zur Verlinkung mit anderen Daten aufweisen (z.B. MBS/PBS in Australien, NHIS/HIRA in Südkorea, ÖGIS in Österreich, CMS in den Vereinigten Staaten);
- ii) zentrale Stellen, die vorhandene Datenbestände unterschiedlicher Datenhalter und mit variablem Inhalt zusammenführen; hier zeichnen sich zwei Möglichkeiten aus: die Daten bleiben bei den Datenhaltern und werden auf Antrag verknüpft (z.B. PHRN in Australien, SDLE in Kanada, eDRIS in Schottland) oder in einer konsolidierten Datenbank aufgehoben (z.B. IDI in Neuseeland);
- iii) Datenbestände, in denen routinemäßig erhobene Daten zusammengeführt und für Kliniker und Forscher zur Verfügung gestellt werden (z.B. CPRD in England, Nivel PC Database in den Niederlanden, Kaiser Health Connect in den Vereinigten Staaten);

- iv) elektronische Plattformen, die es (unter anderem) Patienten erlauben, Zugriff auf ihre eigenen Informationen zu gewinnen und die ggf. als potentielle Grundlage für klinische Daten für die Forschung betrachtet werden könnten (z.B. ELGA in Österreich, myHR in Australien bzw. Sundhed.dk in Dänemark, Journalen in Schweden);
- v) Tools, die die Forschung über mehrere Datensätze hinweg erleichtern (z.B. RKKP in Dänemark, Registerforskning in Schweden).

Drei der Fallstudien heben die Möglichkeit hervor, für Forschungsfragen auf vorhandene, integrierte Routinedaten zurückzugreifen. Damit sind Datensätze gemeint, welche bereits für den primären Nutzungszweck in einer integrierten (verknüpften) Form über mehrere Leistungsbereiche erstellt werden. Ein aus Deutschland vergleichbares Beispiel sind dabei die Datensätze, welche den Krankenkassen auf Grundlage der Abrechnungsdaten vorliegen. Diese sind auf Versichertenebene, liegen über mehrere Jahre vor und decken verschiedene Leistungsbereiche ab. In Australien, Südkorea sowie den USA konnten Möglichkeiten der Datennutzung von integrierten Routinedaten identifiziert werden. Es handelt sich dabei in allen Ländern um Datensätze, welche sich überwiegend aus Abrechnungsdaten (Claimsdata) zusammensetzen.

Es ist ersichtlich, dass eine gute Datengrundlage für Forschungszwecke vor allem dann vorliegt, wenn alle Informationen zur Interaktion von einzelnen Patienten mit dem Gesundheitssystem an einer zentralen Stelle gehalten werden; so können leistungserbringer- und ggf. sektorenübergreifende Versorgungsprozesse abgebildet werden, im optimalen Fall auch im Zeitverlauf. Diese Bedingung ist grundsätzlich in Gesundheitssystemen erfüllt, die ein „Gatekeeping“-System in der Primärversorgung verankert haben; ein typisches Beispiel hierfür sind die Hausärzte (general practitioners, GPs) in Großbritannien. Unter den eingeschlossenen Fallstudien ist das typischste Beispiel für eine Datenbank, die von GPs erfassten Daten zusammenführt, der Clinical Practice Research Datalink (CPRD). CPRD führt Daten von Praxen zusammen, die sich in das Programm eingeschrieben haben („opt-in“). Patienten dieser Praxen können sich wiederum von der Datenübermittlung ausschließen lassen („opt-out“). Die von CPRD gehaltenen Daten können für Beobachtungsstudien beantragt werden, oder zur Durchführung von experimentellen Studien unterstützend herangezogen werden. Sie können darüber hinaus mit anderen Datensätzen von NHS Digital verknüpft werden, um reichere, sektorenübergreifende Einsichten zu ermöglichen. Die Datenbank zur Primärversorgung des niederländischen Forschungsinstituts Nivel setzt ebenso Daten aus der Primärversorgung ein, allerdings kommen hier die Daten aus den elektronischen Praxisverwaltungssystemen auch von Leistungserbringern jenseits der Hausärzte zusammen.

In der Recherche wurden mehrere Fallstudien identifiziert, in denen verschiedene Datenquellen mit unterschiedlichem Inhalt von zentralen Stellen über eine oder mehrere Verlinkungsvariablen verknüpft werden. Das übergreifende Ziel der hierunter fallenden Initiativen ist es, einen besseren Datenzugang für die Forschung zu ermöglichen. Dabei werden Daten aus verschiedenen Quellen, insbesondere von verschiedenen öffentlichen Institutionen verlinkt. So können bspw. Daten zur gesundheitlichen Ver-

sorgung mit Daten aus dem Bildungsbereich und Kriminalitätsstatistiken verknüpft werden. Die Möglichkeit des Datenzugriffs ist bei fast allen Initiativen über Antragsverfahren geregelt. Diese Fallstudien heben die Möglichkeiten klar hervor, die sich auf Basis solcher zentralen Stellen für die Zusammenführung unabhängig entstandener Datensätze für die Versorgungsforschung eröffnen. Unterschiedlich unter den hier aufgeführten Fallstudien – und relevant im Hinblick auf die Überlegungen für eine potenzielle Ideengewinnung für den deutschen Raum – ist das Ausmaß der Konsolidierung der unterschiedlichen Datenbestände. So verknüpfen zum Beispiel das australische PHRN und das kanadische SDLE die Daten projektspezifisch, während die Quelldatensätze grundsätzlich verteilt bleiben. Im Gegensatz hinzu fügt die „integrierte“ neuseeländische IDI eingegangene Daten mithilfe eines „Spine“ zusammen.

Eine erfolgreiche Verlinkung setzt die Möglichkeit einer eindeutigen Zuordnung von Informationen aus separaten Datensätzen voraus. Dazu sind geeignete Schlüsselvariablen notwendig, die in allen zu verlinkenden Datensätzen vorhanden sind. Für eine personenbezogene Verlinkung ist es aus datenschutzrechtlichen Gründen häufig notwendig, statt eindeutiger personenidentifizierender Merkmale aus den Originaldaten, wie eine Versicherungs- oder Ausweisnummer, Pseudonyme einzusetzen; dabei muss für alle zu verlinkenden Datenquellen dasselbe Pseudonymisierungsverfahren angewendet werden. Die Verlinkungsinitiativen in den Ländern zeigen, dass auch ohne das Vorhandensein eines eindeutigen Identifikators eine Verknüpfung auf Basis anderer identifizierender Merkmale möglich ist und seit mehreren Jahren erfolgreich eingesetzt wird. Darüber hinaus sind die Verteilungen der Zuständigkeiten variabel – nichtsdestotrotz kann man beobachten, dass der Einsatz von Vertrauensstellen unterschiedlicher Zugehörigkeiten als üblich gilt. Die detaillierte technische Beschreibung möglicher De-Identifikations- und Verlinkungsoptionen geht über den Rahmen dieses Gutachtens hinaus. Nichtsdestotrotz werden beispielhaft vereinzelte Ansätze beleuchtet, um erste Ideen für zukünftige Diskussionen zu erlauben. Die Verlinkungsinitiativen in den Ländern zeigen, dass auch ohne das Vorhandensein eines eindeutigen Identifikators eine Verknüpfung auf Basis anderer identifizierender Merkmale möglich ist und seit mehreren Jahren erfolgreich eingesetzt wird. Darüber hinaus sind die Verteilungen der Zuständigkeiten variabel – nichtsdestotrotz kann man beobachten, dass der Einsatz von Vertrauensstellen unterschiedlicher Zugehörigkeiten als üblich gilt.

Da Daten im Gesundheitssystem grundsätzlich sensibel sind, kommen in der Regel drei Zugangswege in Frage: (1) eine direkte Datenübermittlung, die ein angemessenes Datenschutz- und Datensicherungskonzept voraussetzt; (2) die Inhouse-Auswertung, die an geeigneten Gastarbeitsplätzen der datenhaltenden Institution stattfindet; und (3) die Fernabfrage, die auf separaten Servern mit „remote access“ über VPN-Verbindung möglich ist. Wie Tabelle 6.1 zu entnehmen ist, kommen alle drei Zugangswege in der Stichprobe der eingeschlossenen Fallstudien vor. Der bisher eingesetzte Ansatz für die Bearbeitung der DaTraV-Daten in Deutschland wurde so in keinem anderen Land identifiziert. Hingegen scheinen sowohl die Nutzung von Gastarbeitsplätzen als auch die Fernabfrage über VPN zunehmend Anwendung zu finden.

In den letzten Jahren haben einzelne Länder in der Stichprobe Initiativen eingeführt, die elektronische Patientenakten anbieten und somit potenziell auch neue Datensätze für die Forschung schaffen könnten. Dabei gibt es in Australien sowie in Österreich die Möglichkeit für Patienten, entweder ganz aus der Nutzung der elektronischen Akte auszusteigen oder aber bestimmte Eintragungen nicht vornehmen zu lassen bzw. diese nicht in der Akte freizugeben.

Die thematisierten Fallstudien aus dem internationalen Vergleich heben wichtige Aspekte hervor, die für die Problematik in Deutschland von Bedeutung sein könnten. Neben dem Mehrwert von vergleichsmäßig einfach zugänglichen, umfangreiche(re)n Abrechnungsdaten, die schneller als in Deutschland vorliegen, sektorenübergreifende Analysen erlauben und personenbezogen mit Gesundheits- sowie auch mit Regionaldaten verknüpfbar sind, wird auch das Potential von Datenbanken beleuchtet, die auf Basis der routinemäßigen Dokumentation (i.d.R. elektronischen Gesundheitsakten) von Leistungserbringern entstehen. Eine Investition in diese Richtung, wie sie zum Beispiel in Großbritannien mit CPRD erfolgte, kann nicht nur für das eigene Gesundheitssystem Einsichten ermöglichen, sondern auch die internationale Kooperation fördern und zur internationalen Sichtbarkeit wissenschaftlicher Exzellenz beitragen.

Eine Nebeneinanderstellung der in Deutschland vorhandenen Datenbestände und der in anderen Ländern eingesetzten Daten deutet darauf hin, dass sich die Diskussion hierzulande an erster Stelle mit der Ausgestaltung der zur Verfügung stehenden Daten befassen sollte und ergänzend die Erfassung zusätzlicher Informationen bedenken kann. Das letzte Kapitel des Gutachtens liefert konkrete Empfehlungen, wie dies vor dem aktuellen regulatorischen Hintergrund erreicht werden könnte.

Abstract

Eine unabdingbare Voraussetzung für die Beantwortung der Forschungsfragen der Versorgungsforschung sind das Vorhandensein und die Zugänglichkeit von umfassenden Daten hoher Qualität. Es ist davon auszugehen, dass die Versorgungsforschung in den kommenden Jahren auf der Verlinkung (Datenlinkage, „data linkage“), also der Verknüpfung mehrerer Datenquellen auf Basis von geeigneten gemeinsamen Schlüsselvariablen beruhen wird. In Deutschland werden umfassende Daten hoher Qualität routinemäßig gesammelt, allerdings sind diese für Forschungszwecke nur bedingt geeignet und eingeschränkt verfügbar. Daher liegt der Ursprung dieses Gutachtens in den Grundfragen, welche Fragestellungen der Versorgungsforschung im Ausland auf Basis besserer Daten besser beantwortet werden können, und auf welchen Wegen eine solche Verbesserung der Datengrundlage in Deutschland erzielt werden könnte.

Fallbeispiele interessanter Verlinkungsansätze aus zwölf Ländern wurden identifiziert und anhand eines vordefinierten Konzeptes analysiert. Informationen wurden für alle Vergleichsländer bzw. Fallstudien zunächst auf Basis von öffentlich zugänglichen Quellen identifiziert. Anschließend wurde pro Land ein Profil erstellt und geeigneten Länderexperten zur Validierung und Ergänzung weitergeleitet. Das Gutachten stellt eine Übersicht der eingeschlossenen Fallstudien vor (Kapitel 3), sowie im Anhang wichtige Eckdaten pro Land und Fallstudie. Es folgen eine nähere Betrachtung der Möglichkeiten der Nutzung von Routinedaten (Kapitel 4), der Verlinkungsmöglichkeiten mit und ohne eindeutiges personenbezogenes Merkmal (Kapitel 5), der unterschiedlichen Zugriffswege für Forscher (Kapitel 6), der Ausgestaltung des Einsatzes von Daten aus elektronischen Patienten- bzw. Gesundheitsakten für die Forschung (Kapitel 7) und ein Zwischenfazit (Kapitel 8). Das Gutachten endet mit einer Wunschliste für mögliche regulatorische Änderungen in Anbetracht der aktuellen politischen Diskussion (Kapitel 9).

Wichtig ist hier zu beleuchten, dass die thematisierten Fallstudien aus dem internationalen Vergleich wichtige Aspekte hervorheben, die für die Problematik in Deutschland von Bedeutung sein könnten. Neben dem Mehrwert von vergleichsmäßig einfach zugänglichen, umfangreiche(re)n Abrechnungsdaten, die schneller als in Deutschland vorliegen, sektorenübergreifende Analysen erlauben und personenbezogen mit Gesundheits- sowie auch mit Regionaldaten verknüpfbar sind, wird auch das Potential von Datenbanken beleuchtet, die auf Basis der routinemäßigen Dokumentation (i.d.R. elektronischen Gesundheitsakten) von Leistungserbringern entstehen. Eine Investition in diese Richtung kann nicht nur für das eigene Gesundheitssystem Einsichten ermöglichen, sondern auch die internationale Kooperation fördern und zur internationalen Sichtbarkeit wissenschaftlicher Exzellenz beitragen.

Eine Nebeneinanderstellung der in Deutschland vorhandenen Datenbestände und der in anderen Ländern eingesetzten Daten deutet darauf hin, dass sich die Diskussion hierzulande an erster Stelle mit der Ausgestaltung der zur Verfügung stehenden Daten befassen sollte und ergänzend die Erfassung zusätzlicher Informationen bedenken kann.

Inhaltsverzeichnis

Abkürzungsverzeichnis	v
Tabellenverzeichnis	vii
Abbildungsverzeichnis	vii
Boxen-Verzeichnis	vii
Zusammenfassung.....	ix
Abstract	xv
Kapitel 1: Hintergrund und Problemstellung	1
Kapitel 2: Methodik des Gutachtens.....	9
Kapitel 3: Übersicht der eingeschlossenen Länder und Fallstudien	17
Australien	17
Dänemark	18
Finnland	20
Kanada	21
Neuseeland.....	23
Niederlande	24
Österreich.....	25
Schweden	26
Südkorea.....	27
Vereinigtes Königreich.....	28
Vereinigte Staaten.....	30
Zusammenfassung und weitere Betrachtung einzelner Aspekte im Gutachten	33
Kapitel 4: Erweiterte Möglichkeiten auf Basis von Routinedaten	37
Für den Primärzweck integriert vorliegende Daten.....	37
Zusammenführung von routinemäßig entstandenen Daten in einem neuen Datenbestand	40
Verlinkung von unabhängig entstandenen Datenbeständen	42
Kapitel 5: Unterschiede im Verlinkungsprozess.....	47
Kapitel 6: Zugriff auf gesundheitsrelevante Daten für Forschende.....	51
Kapitel 7: Überlegungen zu elektronischen Patientenakten	55
Kapitel 8: Zwischenfazit zu den Ergebnissen und Limitationen des Gutachtens.....	59
Kapitel 9: Umsetzungsmöglichkeiten für Deutschland	63
Literaturverzeichnis.....	73
Anhang.....	79
Australia (AUS).....	81

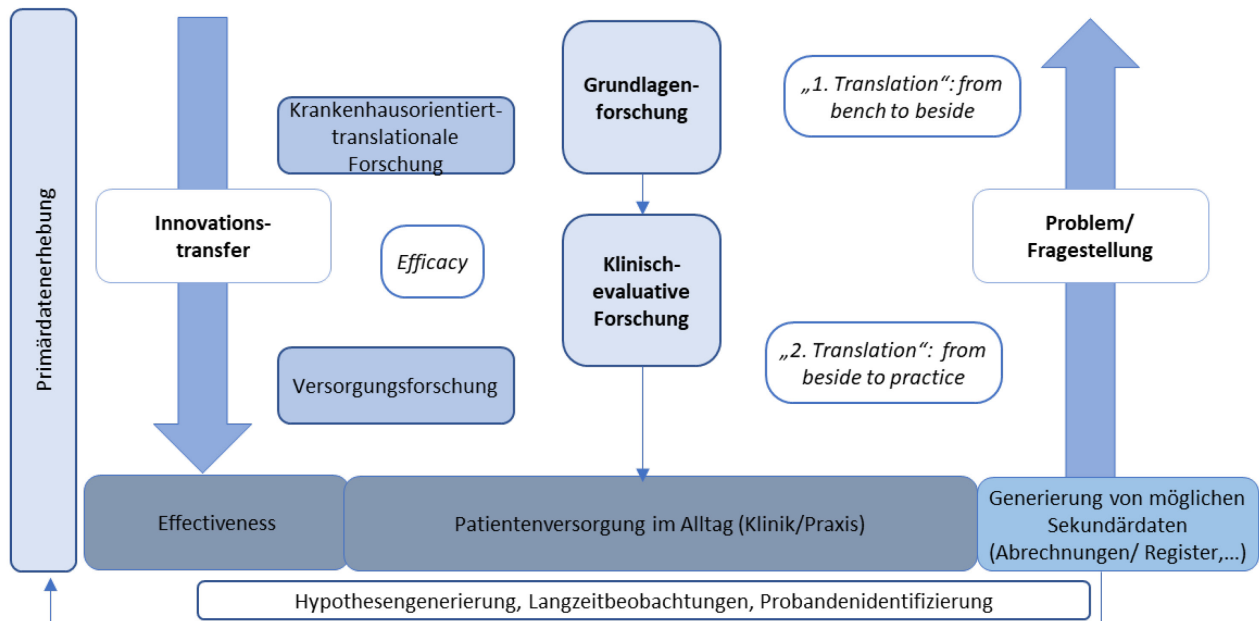
Denmark (DNK).....	87
Finland (FIN)	93
Canada (CAN).....	97
New Zealand (NZL)	105
The Netherlands (NLD)	109
Austria (AUT)	115
Sweden (SWE)	123
Republic of Korea (KOR)	127
United Kingdom (GBR)	131
United States (USA).....	137

Kapitel 1: Hintergrund und Problemstellung

Die Versorgungsforschung („health services research“) hat sich in Deutschland und international bereits als vielfältige Disziplin mit komplexem, transdisziplinärem Charakter etabliert (Pfaff et al. 2017, Raspe et al. 2010). Im internationalen Vergleich war Deutschland nicht unter den Vorreitern, denn erst Ende der neunziger Jahre hat sich hierzulande ein eigenständiges Forschungsgebiet entwickelt, das mit dem inzwischen allgemein anerkannten Begriff „Versorgungsforschung“ einen multidimensionalen, methodisch, theoretisch und inhaltlich differenzierten Aufgabenbereich umfasst (Raspe et al. 2010). Seitdem spiegelt sich die Anerkennung der Wichtigkeit dieses Forschungsgebietes für die Unterstützung und Optimierung des deutschen Gesundheitssystems auf politischer, gesetzlicher und finanzieller Ebene wider (Schrappe & Pfaff 2017). Frühen Förderinitiativen, zum Beispiel durch die Bundesärztekammer (2005-2010), folgte ein Vorhaben zum Ausbau der Disziplin in den Koalitionsverträgen 2009 und 2013, und aktive Investitionen durch die Bundesregierung (z.B. Aktionsplan Versorgungsforschung des BMBF, Ressort Versorgungsforschung des BMG), und die Deutsche Forschungsgemeinschaft (DFG). Nicht zuletzt durch die gesetzliche Verankerung des Innovationsfonds im GKV-Versorgungsstärkungsgesetz im Jahr 2015 wurde die Weiterführung dieser Intention verdeutlicht.

Die DFG stellt drei grundlegende Merkmale der Versorgungsforschung dar, die das Verständnis dieser Disziplin prägen: das besondere Problemfeld der gesundheitlichen Versorgung der Bevölkerung, eine Fokussierung auf die Organisation, Regulierung und Verbesserung dieser Versorgung, und eine konkrete primäre Zielsetzung, die auf eine evidenzbasierte Beratung von Personen bzw. Einrichtungen mit Steuerungsfunktionen im Gesundheitswesen abzielt. Schrappe und Pfaff definieren die Versorgungsforschung als „ein fachübergreifendes Forschungsgebiet, das ausgehend von der Patienten- sowie Populationsperspektive und vor dem Hintergrund komplexer Kontextbedingungen die Versorgungsstrukturen und -prozesse der Gesundheitsversorgung untersucht, den Outcome auf Ebene der Alltagsversorgung beschreibt und komplexe Interventionen zur Verbesserung der Versorgung evaluiert“ (Schrappe & Pfaff 2017). Beide Betrachtungen heben ein grundlegendes Alleinstellungsmerkmal der Disziplin hervor, die erste implizit und die zweite explizit: im Fokus stehen komplexe Interventionen, die unter komplexen Kontextbedingungen eingesetzt werden und entsprechend evaluiert werden müssen (sog. „doppelte Komplexität“; vgl. Abbildung 1.1). Die Abgrenzung der Versorgungsforschung von anderen Stufen der biomedizinischen Forschung wird in Tabelle 1.1 thematisiert.

Die Fragestellungen der Versorgungsforschung umfassen neben der deskriptiven und Inferenzorientierten Analyse entlang der drei Versorgungsqualitätsdimensionen von Avedis Donabedian (Strukturen, Prozesse und Outcomes) (Donabedian, 1980) auch die Versorgungskonzeption und das Versorgungsmanagement, und somit auch die Implementierung und die Evaluation von Interventionen (Schrappe & Pfaff 2017). Unabdingbare Voraussetzungen für die aussagekräftige Beantwortung aller dieser Fragen, vor allem vor dem Hintergrund der doppelten Komplexität, sind das Vorhandensein und die Zugänglichkeit von umfassenden Daten hoher Qualität.



Quelle: in Anlehnung an Schrappe & Pfaff 2017

Abbildung 1.1 Forschungsstadien und Evidenzgenerierung

Tabelle 1.1 Abgrenzung der Disziplinen innerhalb der (bio)medizinischen Forschung

Biomedizinische Grundlagenforschung	Klärung (patho-) physiologischer Mechanismen, Entdeckung/Erfindung einer neuen Gesundheitstechnologie
Translationsforschung	Erste Anwendung an Gesunden bzw. Kranken („ <i>from bench to bedside</i> “, „ <i>proof of principle</i> “; entspricht Phase I- & II-Studien in der Arzneimittelforschung)
Klinische Forschung Transferforschung (<i>clinical evaluative sciences</i>)	Evidenz für Wirksamkeit und Nutzen unter (artifiziellen) Studienbedingungen („ <i>from bedside to market</i> “; entspricht Phase III-Studien in der Arzneimittelforschung)
	Evidenz für Nutzen unter Alltagsbedingungen („ <i>from early market to routine care</i> “; entspricht pragmatischen RCTs, Phase-IV-Studien, Anwendungsbeobachtungen, Outcomes Research)
Versorgungsforschung (<i>health services research</i>)	Beobachtung, Analyse, Prognose, Bewertung, Weiterentwicklung und Evaluation der Routineversorgung

Quelle: Raspe et al. 2010

Je nach Fragestellung unterscheiden sich die angemessenen Datenquellen und die zu untersuchenden Endpunkte. Obwohl gezielt durchgeführte Primärstudien hier eine wichtige Rolle spielen, ist die Versorgungsforschung als Disziplin dadurch charakterisiert, dass häufig auf Sekundärdaten zurückgegriffen wird, also auf Daten, die ursprünglich für einen anderen Zweck generiert wurden (vgl. Abbildung 1.1). Hierzu gehören unter anderem Daten aus klinischen und epidemiologischen Registern und Routinedaten, die während des Versorgungsgeschehens entstehen, wie zum Beispiel die Abrechnungsdaten der gesetzlichen Krankenversicherung. In den letzten Jahren ist das umstritten definierte Konzept von „Big Data“ zu den Evidenzgrundlagen der Versorgungsforschung hinzugekommen. Trotz des Mangels an einer allgemeingültigen Definition gibt es gewisse Elemente, die Big Data eindeutig charakterisie-

ren: es handelt sich um Datenmengen, die so groß und komplex sind, dass sie mit den üblichen Datenverarbeitungsmethoden nicht adäquat gemanagt oder ausgewertet werden können (vgl. McKinsey 2011). Im gängigen Gebrauch des Begriffs sind „Big Data“ zum Teil unstrukturierte Datenmengen, die vorwiegend auf Basis von neuen, digitalen Anwendungen mit großer Geschwindigkeit generiert und ausgewertet werden können. So haben sie theoretisch das Potential, komplexe Versorgungsverläufe abzubilden und zu relevanten Erkenntnissen im Rahmen der Versorgungsforschung beizutragen (vgl. Wegscheider & Koch-Gromus 2017).

Vor dem Hintergrund der dynamischen Landschaft der unterschiedlichen Datengrundlagen, die für die Versorgungsforschung von Nutzen sein könnten, und der zunehmenden Wichtigkeit von Fragestellungen, für deren Beantwortung ein breites Spektrum an Informationen erforderlich ist (z.B. zum Einfluss des Klimawandels auf gesundheitsbezogenen Outcomes), ist davon auszugehen, dass die Versorgungsforschung in den kommenden Jahren auf der Verlinkung (Datenlinkage, „data linkage“), also der Verknüpfung mehrerer Datenquellen auf Basis von geeigneten gemeinsamen Schlüsselvariablen beruhen wird. Die Datenquellen beinhalten Routinedaten, verbunden mit – oder verknüpfbar mit – klinischen Daten und weiteren, teils unstrukturierten Alltags- und Kontextdaten. Der Anspruch an Daten auf hohe Validität, Vollständigkeit, Aktualität, Repräsentativität und Qualität soll davon unberührt bleiben – aber es ist zu erwarten, dass die Forschung zunehmend in größeren Verbänden stattfindet.

In Deutschland werden umfassende Daten hoher Qualität routinemäßig gesammelt (siehe Tabelle 1.2; Nimptsch et al. 2015; DIMDI 2014, Nimptsch & Mansky 2010). Allerdings sind diese für Forschungszwecke nur bedingt geeignet und eingeschränkt verfügbar (Schreyögg 2016, Kreis et al. 2016). Die Datenbestände einzelner gesetzlicher Krankenkassen bieten zwar die Möglichkeit, den Versorgungsprozess sektorenübergreifend abzubilden, können aber keinen Anspruch auf Repräsentativität für die Bevölkerung erheben und beinhalten keine Leistungen, die nicht über die Kasse abgerechnet wurden (Geyer & Jaunzeme 2015); für jedes Forschungsvorhaben ist eine gesonderte Kooperationsbasis mit den entsprechenden Sozialversicherungsträgern erforderlich. Der mit dem GKV-Modernisierungsgesetz (§303a-e SGB V) eingeführte gemeinsame Datenpool der Krankenkassen wurde erst 2012 mit der Datentransparenzverordnung (DaTraV) umgesetzt und ermöglicht seitdem einen kassenübergreifenden Zugang zu versichertenbezogenen Diagnose-, Arzneimittel-, und Ausgabendaten. Jedoch basiert dieser Datensatz auf den für den morbiditätsorientierten Risikostrukturausgleich benutzten Informationen, die die Kassen dem Bundesversicherungsamt (BVA) übermitteln. Das Deutsche Institut für Medizinische Dokumentation und Information (DIMDI) erhält diese Daten mit einem Zeitverzug von ca. vier Jahren und stellt sie dann für Forschungszwecke bestimmten Nutzungsberechtigten zur Verfügung. Bei den Daten fehlen wichtige Variablen, die die Anwendbarkeit für viele Fragestellungen erheblich einschränken, wie zum Beispiel Codes zu Prozeduren oder Kennzeichen für die ambulanten und stationären Leistungserbringer (vgl. Schreyögg 2016, Hoffmann & Glaeske 2017). Der Zugang ist aktuell nur über Datenfernverarbeitung für a priori definierte Fragestellungen und Methoden möglich, welches in der Regel zu längeren Verzögerungen bei der Bearbeitung führt und komplexe Auswertun-

gen erheblich erschwert. Sektorenspezifische Abrechnungsdaten, wie die DRG-Statistik nach § 21 KHEntgG oder die Daten der kassenärztlichen Vereinigungen eignen sich vorwiegend für Fragestellungen, die sich (ausschließlich) auf Leistungen des entsprechenden Sektors beziehen. Aus der (nicht abschließenden) Darstellung in Tabelle 1.2 wird deutlich, dass in Deutschland viele Datenbestände nebeneinander bestehen, die jedoch stark fragmentiert sind.

Mit dem kurz vor Fertigstellung dieses Gutachtens beschlossenen Digitale-Versorgung-Gesetz (DVG) soll die Nutzbarkeit der GKV-Routinedaten für Forschungszwecke auf der Grundlage der §§ 303a-e SGB V verbessert werden. Vorgesehen ist, die Datenaufbereitungsstelle zu einem Forschungsdatenzentrum weiterzuentwickeln und die Zugangsmöglichkeiten zu verbessern. Auch das Datenangebot soll erweitert werden, indem die Krankenkassen definierte Kosten- und Leistungsdaten aus der ambulant-ärztlichen Versorgung, Arzneimittelversorgung, Krankenhausversorgung, der Versorgung durch Hebammen und der Heil- und Hilfsmittelversorgung an den GKV-Spitzenverband, der dann als Datensammelstelle fungieren soll, übermitteln. Das Nähere zu Art und Umfang dieser Daten soll per Rechtsverordnung festgelegt werden, die das Bundesministerium für Gesundheit im Benehmen mit dem Bundesministerium für Bildung und Forschung ohne Zustimmung des Bundesrates bestimmt.

Im internationalen Vergleich ist erkenntlich, dass die Entwicklung der Versorgungsforschung als eigenständiger Disziplin in mehreren anderen Ländern früher stattfand (Raspe et al. 2010, Schrappe & Pfaff 2017, HSR-Europe 2011). Auch vor dem Hintergrund unterschiedlicher Datenschutzbestimmungen hat dies dazu geführt, dass im Ausland zum Teil vielfältigere Erkenntnisse aufgrund besserer Datenzugänglichkeit erzielt werden können (vgl. OECD 2015). Daher liegt der Ursprung dieses Gutachtens in den Grundfragen,

- a) welche Fragestellungen der Versorgungsforschung im Ausland auf Basis besserer Daten beantwortet werden können, und
- b) auf welchen Wegen eine solche Verbesserung der Datengrundlage in Deutschland erzielt werden könnte.

Der Fokus wird dabei auf die Verlinkung von unterschiedlichen Datengrundlagen zur Beantwortung komplexer Fragestellungen gelegt, um den oben beschriebenen aktuellen Herausforderungen nachzugehen. Näheres zur Methodik des Gutachtens wird im nächsten Kapitel beschrieben. Letztlich war es ein Ziel des Gutachtens, vorhandene Arbeit aus dem deutschen und internationalen Raum, die deskriptiv internationale Evidenzgrundlagen für die Versorgungsforschung zusammenfasst, nicht zu replizieren, sondern zu erweitern und zu vertiefen (vgl. OECD 2015, OHE 2015, Gothe 2015).

Tabelle 1.2 Wichtigste Sekundärdatensätze für die Versorgungsforschung in Deutschland, Stand November 2019

Versorgungssektor/ Bereich	Grundlage für die Datenübermittlung	Erfasste Leistungsbereiche/ Sachverhalte	Primäre Datenempfänger	Sekundäre Datenempfänger	Möglichkeit des Datenzugangs	Analyseebene und wesentliche verfügbare Dateninhalte
Abrechnungsdaten, die an die gesetzlichen Krankenversicherungen übermittelt werden						
Stationäre Krankenhausversorgung	§ 301 SGB V	Stationäre Behandlung	Krankenkassen	BVA	DaTraV	Individualebene Entlassungsdatum, Diagnosen
					Nur projektbezogen bei einzelnen Krankenkassen	Aufnahmedatum und -anlass, Prozeduren, Entlassungsgrund
Ambulante Krankenhausversorgung	§ 301 SGB V	Ambulante Behandlung (Institutsleistungen) ¹	Krankenkassen		Nur projektbezogen bei einzelnen Krankenkassen	Individualebene Behandlungsdatum, Diagnosen, Prozeduren
Ambulant-ärztliche Versorgung	§ 295 SGB V	Ambulante Behandlung durch Vertragsärzte, MVZ, ermächtigte Krankenhausärzte	KV -> Krankenkassen bzw. Krankenkassen (Selektivverträge) ²	BVA	DaTraV	Individualebene Quartal, Diagnosen
					Nur projektbezogen bei einzelnen Krankenkassen oder KVen	Prozeduren
Arzneimittel	§ 300 SGB V	In Apotheken ausgegebene, verordnete Arzneimittel	Apothekenrechenzentren -> Krankenkassen	BVA KV	DaTraV	Individualebene Verordnungsdatum, Anzahl, Pharmazentralnummer
Rehabilitation zu Lasten der gesetzlichen Krankenversicherung	§ 301 Abs. 4 SGB V	Stationäre Rehabilitation	Krankenkassen		Nur projektbezogen bei einzelnen Krankenkassen	Individualebene Diagnosen, Aufnahmedatum, Entlassungsdatum und -grund, Angaben zur Leistung
Abrechnungsdaten, die an andere Versicherungsträger übermittelt werden						
Rehabilitation zu Lasten der gesetzlichen Rentenversicherung	§ 301 Abs. 4 SGB V (Datenübermittlungsrahmenvereinbarung)	Stationäre, teilstationäre und ambulante Rehabilitation	DAV der Rentenversicherung -> Rentenversicherungsträger		FDZ der Deutschen Rentenversicherung (RSD)	Individualebene Diagnosen, Aufnahmedatum, Entlassungsdatum und grund, Angaben zur Leistung
Pflege zu Lasten der gesetzlichen Pflegeversicherung	§ 105 SGB XI	Ambulante und stationäre Pflege, Sachleistungen, Pflegegeld	Pflegekassen (bei den gesetzlichen Krankenkassen)		Nur projektbezogen bei einzelnen Pflegekassen	Pflegegrad, Höhe des Pflegegeldes, Abrechnungen der Leistungserbringer

Versorgungs- sektor/ Bereich	Grundlage für die Da- tenüber- mittlung	Erfasste Leis- tungsbereiche/ Sachverhalte	Primäre Daten- empfänger	Sekundäre Daten- empfänger	Möglichkeit des Daten- zugangs	Analyseebene und wesentliche ver- fügbare Dateninhalte
Krankenkassenübergreifende Abrechnungsdaten der Leistungserbringer						
Stationäre Krankenhaus- versorgung (Somatik)	§ 21 KHEntgG	Stationäre Krankenhaus- behandlung	InEK	Statistisches Bundesamt	FDZ-Bund („DRG- Statistik“)	Behandlungsfälle- ebene Diagnosen, Prozeduren, Auf- nahmedatum und - anlass, Entlas- sungsdatum und - grund
Stationäre Krankenhaus- versorgung (Psychiatrie und Psycho- somatik)	§ 21 KHEntgG	Stationäre und stationsäqui- valente Kran- kenhaus- behandlung, psychiatrische Institutsambu- lanzen	InEK	Statistisches Bundesamt	Derzeit noch nicht verfüg- bar	Behandlungsfälle- ebene Diagnosen, Prozeduren, Auf- nahmedatum bzw. Behandlungsdatum
Medizinische Erhebungsdaten						
Bundeskrebs- register	§ 3 BKRg	Krebsneu- erkrankungen	ZfKD beim Robert- Koch- Institut		ZfKD-FDZ	Individualebene Krebsart, Diagno- sezeitpunkt, Be- handlung und Nachsorge, Sterbe- zeitpunkt
Bevölkerungs- repräsentative Gesundheits- surveys		Selbst-berich- tete Angaben der Befragten	Robert- Koch- Institut		FDZ RKI	Individualebene Vorhandensein chronischer Er- krankungen, Inan- spruchnahme
Todes- ursachen	§ 2 BevStatG	Sterbefälle	Statistische Landesämter -> Statisti- sches Bun- desamt		FDZ-Bund	Individualebene Sterbedatum, Fami- lienstand, Todesur- sache
Externe ge- setzliche Qua- litätssicherung	§ 137a SGB V	Krankenhaus- versorgung, ambulante Versorgung (teilweise) für bestimmte QS- Verfahren	IQTIG		IQTIG	Individualebene Informationen zu Indikation, Diagno- sen, Prozeduren

Versorgungssektor/ Bereich	Grundlage für die Datenübermittlung	Erfasste Leistungsbereiche/ Sachverhalte	Primäre Datenempfänger	Sekundäre Datenempfänger	Möglichkeit des Datenzugangs	Analyseebene und wesentliche verfügbare Dateninhalte
Kontextdaten auf aggregierter Ebene						
Bevölkerung	§ 2 BevStatG	Bevölkerungsbewegung	Statistische Landesämter -> Statistisches Bundesamt		FDZ-Bund	Gemeindeebene Lebendgeborene, Verstorbene, Zuzüge, Fortzüge, Alter, Geschlecht
Regionalinformationen		Stadt- und Raumentwicklung	BBR		Online-Atlas INKAR	Gemeindeebene Wirtschaftskraft, öffentliche Finanzen, Sozialleistungen, Arbeit, Bildung, Wohnen, Verkehr, Erreichbarkeit, Flächennutzung

¹ u. a. Ambulantes Operieren, Hochschulambulanzen, psychiatrische Institutsambulanzen und sozialpädiatrische Zentren, integrierte Versorgung, ambulante spezialfachärztliche Versorgung

² Hausarztzentrierte Versorgung, besondere ambulante ärztliche Versorgung, integrierte Versorgung.

SGB: Sozialgesetzbuch, BVA: Bundesversicherungsamt, KV: Kassenärztliche Vereinigung, DAV: Datennahme- und Verteilstelle, DaTraV: Datentransparenzverordnung, MVZ: Medizinische Versorgungszentren, FDZ: Forschungsdatenzentrum, KHEntgG: Krankenhausentgeltgesetz, InEK: Institut für das Entgeltsystem im Krankenhaus BKRg: Bundeskrebsregistergesetz, ZfKD: Zentrum für Krebsregisterdaten, RKI: Robert-Koch-Institut, QS: Qualitätssicherung, IQTIG: Institut für Qualitätssicherung und Transparenz im Gesundheitswesen, BevStatG: Bevölkerungsstatistikgesetz, BBR: Bundesamt für Bauwesen und Raumordnung, INKAR: Indikatoren und Karten zur Raum- und Stadtentwicklung

Kapitel 2: Methodik des Gutachtens

Zur Beantwortung der Grundfragestellungen sollte im Detail untersucht werden, mit welchen Methoden und unter welchen Voraussetzungen in ausgewählten Ländern erfolgreich (gute) Datengrundlagen geschaffen worden sind. Der Aufbau – und damit verbunden die Verlinkungsmöglichkeit – von relevanten Datengrundlagen stehen in engem Zusammenhang mit der organisatorischen Gestaltung des Gesundheitssystems, vor allem der Struktur der Leistungserbringung und der entsprechenden Abrechnungs- und Vergütungsmechanismen. Ihre Zugänglichkeit wird darüber hinaus von Datenschutzbestimmungen und -normen, zum Teil auch kultureller Art, geprägt. Daher wurde die Analyse in diesem Gutachten so konzipiert, dass Erkenntnisse zu Datenstruktur und -Fluss so weit wie möglich systemunabhängig bzw. –übertragbar sind, während möglichst auch länderspezifische regulatorische Rahmenbedingungen erfasst und analysiert werden können.

Analytisches Konzept: Für die Analyse der identifizierten Fallstudien aus dem internationalen Vergleich wurde ein Konzept entwickelt, welches zwischen Datenbeständen mit versorgungsrelevanten Daten i) auf Individualebene („individual level data“, also Mikrodaten) und Daten auf Nicht-Individualebene („non-individual level data“) und ii) aus drei bzw. vier großen inhaltlichen Blöcken (Gesundheitsdaten, Gesundheitsversorgungsdaten, soziodemographischen- bzw. ökonomischen Daten, Umweltdaten) unterscheidet (Abbildung 2.1).

In Abbildung 2.1 sollen die Pfeile die (möglichen) Verknüpfungsmöglichkeiten symbolisieren, die für die Individualdaten entweder über die Speicherung in gemeinsamen Datenbanken (z.B. beruhend auf einer elektronischen Gesundheitsakte) oder über eine Verlinkungsmöglichkeit via eines eindeutigen personenbezogenen Merkmals („[unique] patient identifier“) stattfinden. Genauso relevant sind aber auch die Verknüpfungsmöglichkeiten der Individualdaten mit den nicht-individuellen Daten und von diesen untereinander. Es ist nicht zu erwarten, dass Daten(banken) notwendigerweise wie in Abbildung 1 strukturiert sind; die Kategorien sollen vielmehr dazu dienen, die jeweils enthaltenen Daten inhaltlich zu klassifizieren.

Beispiele für Variablen aus den unterschiedlichen Datenkategorien in Abbildung 2.1 werden in Abbildung 2.2 aufgeführt (vgl. auch Datenschutz-Grundverordnung (DSGVO), Erwägungsgründe 35 und 53).

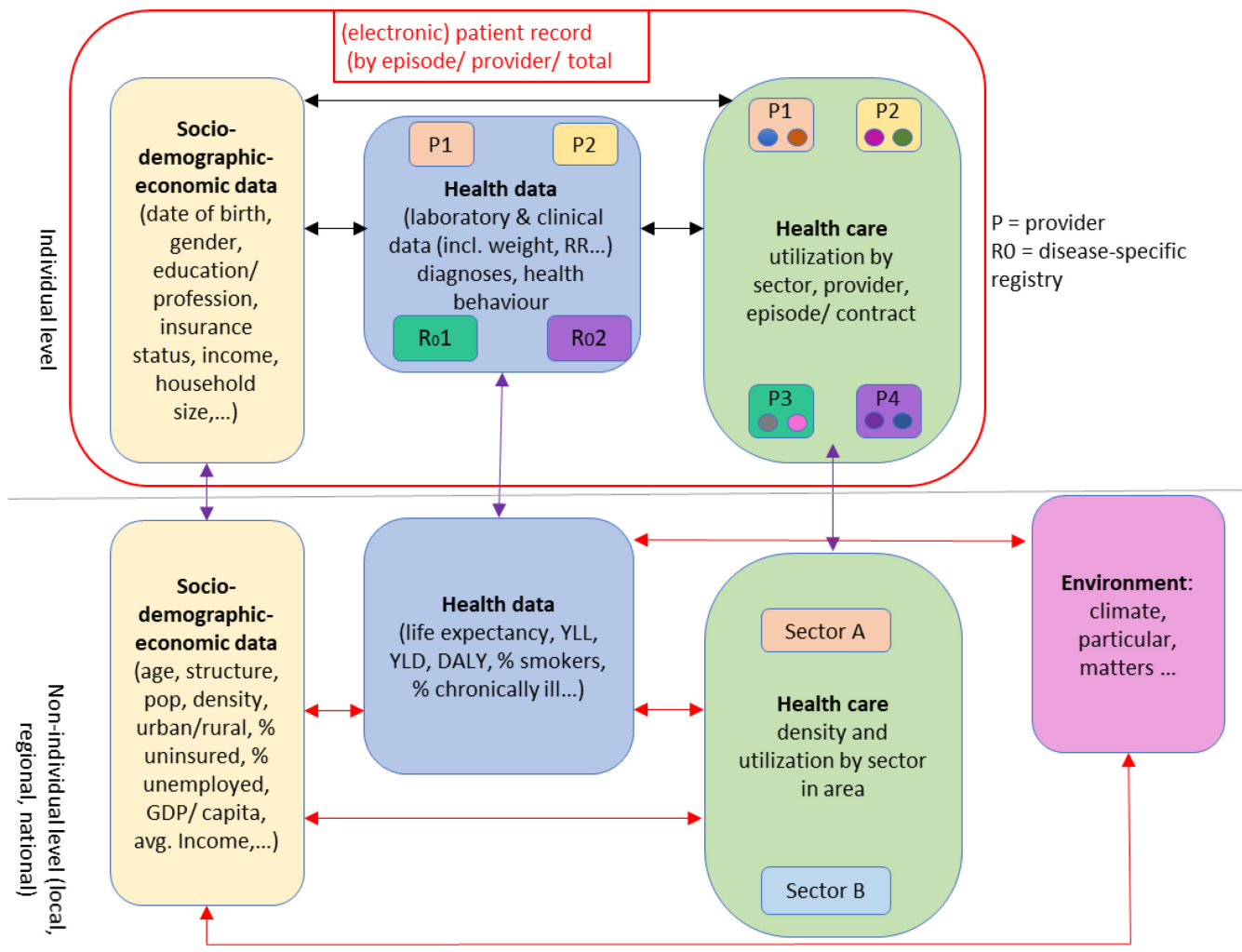


Abbildung 2.1 Analysekonzept

	Sozioökonomische Determinanten	Gesundheit	Gesundheitswesen	Umwelt
Individualebene	<ul style="list-style-type: none"> ✓ Alter ✓ Geschlecht ✓ Andere demographische Faktoren (Bsp. Migrationshintergrund, Familienstand) ✓ Arbeitsverhältnisse ✓ Einkommen ✓ Ausbildung ✓ Versicherungsstatus ✓ ... 	<ul style="list-style-type: none"> ✓ Untersuchungsergebnisse ✓ Laborergebnisse ✓ Diagnose/ Symptome ✓ Risikofaktoren ✓ Durchgeführte Prozeduren ✓ Medikationspläne ✓ Medizingeräte ✓ Pflegerisches Outcome Parameter (Bsp. Stürze, Schmerz Management, Fehlmedikation) 	<ul style="list-style-type: none"> ✓ Untersuchungsergebnisse ✓ Laborergebnisse ✓ Diagnose/ Symptome ✓ Arzttermine ✓ Überweisungen ✓ Einweisungen ✓ Prozeduren ✓ Medikationspläne ✓ Medizingeräte ✓ ... 	X
Aggregierte Ebene	<ul style="list-style-type: none"> ✓ Mittleres Haushaltseinkommen ✓ BIP/Kopf ✓ Bevölkerungsanteil Stadt / Land ✓ % unterhalb der Armutsgrenze ✓ Bildungserfolg ✓ Familienstand ✓ %arbeitslos ✓ Versicherungsstatus ✓ % alleinerziehend ✓ % mit Wohneigentum ✓ % mit Unterstützungsbedarf ✓ Altersstruktur ✓ Div. Aggregate ✓ ... 	<ul style="list-style-type: none"> ✓ Lebenserwartung ✓ Disability-adjusted life years (DALYs) und Bestandteile (YLL/YLD) ✓ Personen unter Risikofaktor ✓ % chronisch kranke Personen mit Einschränkungen ✓ ... 	<ul style="list-style-type: none"> ✓ Versorgungsdichte/ Nutzungsgrad ✓ ... 	<ul style="list-style-type: none"> ✓ Luftreinheitsgrad ✓ Klimatische Indikatoren ✓ Giftstoffe/Abfall ✓ Entfernungskennzahlen ✓ Siedlungsgrad/ Flächennutzung im ländlichen Raum ✓ ...

Abbildung 2.2 Mögliche Inhalte unterschiedlicher Datenkategorien

Auswahl der Vergleichsländer und der Fallstudien: Um die Stichprobe der Vergleichsländer für das Gutachten auszuwählen, wurden drei Hauptkriterien zugrunde gelegt:

- a) eine Mischung aus Gesundheitssystemstrukturen, d. h. sowohl solche, die aus Sozialversicherungssystemen mit mehreren Versicherungsträgern stammen und somit im Aufbau der Datenstruktur in Deutschland ähneln müssten, als auch welche, die ggf. Impulse für eine völlig neue Denkweise anbieten; des Weiteren sollten Systeme mit unterschiedlichen Ausprägungen der Dezentralisierung von Leistungserbringung und Gesundheitssystemsteuerung vertreten sein;
- b) ein Spektrum bezüglich der Datenverfügbarkeit und -nutzung – hierzu wurde die periodische Befragung zur Health Data Governance der OECD herangezogen (OECD 2015), die die Aktivität mehrerer OECD-Länder hinsichtlich der Regulierung, des Aufbaus und der Nutzung von gesundheitsbezogenen Datengrundlagen beschreibt; und
- c) bekannte interessante Beispiele relevanter Initiativen waren ebenfalls abzudecken. Diese wurden auf Basis von Diskussionen der Arbeitsgruppe sowie mit dem Wissenschaftlichen Beirat des Zentralinstituts für die Kassenärztliche Versorgung ausgewählt.

Somit erhebt das Gutachten keinen Anspruch auf eine Vollständigkeit oder eine strenge Systematik hinsichtlich der eingeschlossenen Länder. Ebenso wenig wurde eine erschöpfende Darstellung der Datenbestände der eingeschlossenen Länder angestrebt. Vielmehr ging es darum, wichtige und potentiell nützliche Fallstudien zu identifizieren, die Erkenntnisse zu mehr oder weniger „disruptiven“ Lösungsansätzen für die deutsche Realität liefern.

Nichtsdestotrotz wurden vor dem Hintergrund der in Kapitel 1 geschilderten Limitationen der deutschen Datenbestände, der gesetzgeberischen Dynamik hinsichtlich der Implementierung der elektronischen Patientenakte während der Bearbeitungsperiode (2019) und der Tatsache, dass klinische Daten in Deutschland zumindest aktuell fast ausschließlich aus Registern bzw. Befragungen kommen, Fallstudien priorisiert, die eine Zusammenführung von intersektoralen Routinedaten, Daten aus elektronischen Patienten- bzw. Gesundheitsakten sowie Registern ermöglichen. Zusätzliche Fallstudien wurden im Laufe der Recherche aus der Literatur und nach Empfehlungen der Länderexperten (siehe unten) eingeschlossen.

Tabelle 2.1 zeigt die endgültige Auswahl an Ländern, die im Rahmen dieses Gutachtens herangezogen wurden, den primären Interessensgrund pro Land sowie das Scoring hinsichtlich Datenverfügbarkeit, -reife und -nutzung gemäß OECD (OECD 2015).

Tabelle 2.1 Länderauswahl und jeweilige Einschlussbegründung

Land	Interessensgrund	Ranking OECD (2015)*		
		Daten- verfügbarkeit	Daten- nutzung	Rang**
Australien	Dezentral verteilte Datenquellen werden unter Beteiligung der Datenhalter über eine zentrale Vertrauensstelle studienspezifisch zu „Meta-data“ zusammengeführt.	nicht erfasst		
Dänemark	Alle Routinedaten und zunehmend auch klinische Daten werden zusammengeführt; akademische Einrichtungen erhalten vergleichsweise einfach Zugang.	Hoch	Mittel	5
Finnland	Neue digitale Vernetzungsmöglichkeiten erleichtern die Zusammenführung und Bereitstellung von Gesundheitsdaten.	Hoch	Hoch	8
Kanada	Regionale Verlinkungsinitiativen bringen auf Basisunterschiedlicher strukturellen Modelle mehrere Datenbestände zusammen.	Hoch	Hoch	12
Neuseeland	Integrated Data Infrastructure, die alle öffentlichen Daten (bis hin zum Zensus) zusammenfasst und auf Antrag für Projekte, die in öffentlichem Interesse sind, nutzen lässt.	Mittel	Hoch	10
Niederlande	Daten mehrerer Leistungsanbieter im ambulanten Sektor werden zusammengeführt, um integrierte Versorgung zu ermöglichen und können für Forschungszwecke benutzt werden.	Mittel	Mittel	17
Österreich	Neu eingeführte elektronische Patientenakte mit vorgesehener Forschungszugänglichkeit; mehrere Krankenversicherer, regionale und zentrale Gesundheitsberichtserstattungssysteme.	nicht erfasst		
Schweden	Gut etablierte, verlinkbare Patientenregister; Nutzung verlinkter Daten zur Unterstützung klinischer Leitlinien.	Hoch	Hoch	4
Südkorea	Eine zentrale Krankenkasse macht die Routinedaten aller >50 Millionen Südkoreaner auf einem zentralen Server antragsweise für Forschung zugänglich.	Hoch	Hoch	2
UK England	Clinical Practice Research Datalink sammelt deidentifizierte Daten von mehr als 10 Millionen Patienten. (Quelle: Hausärzte).	Mittel	Hoch	14
UK Schottland	Daten aller Einwohner aus verschiedensten Quellen (Apotheken, Hausärzte, Krankenhäuser, öffentliche Behörden wie Polizei) werden auf regionaler und nationaler Ebene aufbewahrt und können über den Community Health Index verlinkt werden.	Mittel	Hoch	7
USA	Forschungseinrichtungen treffen auf privatrechtlicher Basis Vereinbarungen mit Organisationen und Einzelpersonen, bieten Erkenntnis gegen Datenspende.	Mittel	Mittel	18

*niedrig / mittel / hoch = 0–33 % / 34–66 % / 67–100 % aller Datenbestände, die zur Verfügung stehen bzw. regelmäßig zu Forschungszwecken verknüpft werden; ** unter den 22 erfassten Ländern in Anbetracht einer Reihe von Variablen bezüglich Datenverfügbarkeit, -reife, und -nutzung (Platz 1 entspricht dem höchsten und somit besten Score; dieser wurde 2015 von Island belegt)

Entwicklung des Erhebungsinstruments. Zur Beschreibung der eingeschlossenen Fallstudien und der relevanten Rahmenbedingungen in den Vergleichsländern wurde ein Erhebungsinstrument in mehreren Schritten entworfen, abgestimmt, pilotiert, angepasst und schlussendlich angewendet. Im November 2018 wurde eine erste Version entwickelt und mit dem Wissenschaftlichen Beirat des Zentralinstituts für die Kassenärztliche Versorgung abgestimmt. Anschließend wurden sechs Länder aus der Stichprobe ausgewählt (Australien, England, Neuseeland, Schottland, Schweden, Südkorea), um den Einsatz des Erhebungsinstruments grundsätzlich zu pilotieren, dessen Sinnhaftigkeit und Angemessenheit für die Beantwortung der Fragestellungen zu überprüfen und ggf. Inhalte und/oder Form zu optimieren. Die Auswahl der Pilotländer reflektierte sowohl das Ergebnis von Machbarkeitsanalysen zum Umfang öffentlich zugänglicher Informationen sowie eine Mischung der ursprünglichen Interessensgründe, die die Länderstichprobe für das Projekt bedingt haben (Tabelle 2.1). Tatsächlich erwies es sich als notwendig, den Aufbau des Erhebungsinstruments anzupassen, da es ursprünglich von der Datenstruktur im deutschen Gesundheitssystem geprägt und somit für andere Länder nicht eindeutig zu übertragen war. Nach einer weiteren Beratung mit dem Wissenschaftlichen Beirat entstand die endgültige Version des Erhebungsinstrumentes; diese findet sich (aus dem Englischen übersetzt) in Tabelle 2.2 wieder.

Tabelle 2.2 Erhebungsinstrument zur Beschreibung von Fallstudien

Allgemeine Fragen pro Land	Spezifische Fragen pro Fallstudie
<ul style="list-style-type: none"> • Aufbau des Gesundheitssystems, Organisation der Leistungserbringung • Gesetzlicher Rahmen für Datennutzung und -sicherheit (allgemein und spezifisch zu Gesundheit) • Strategien für Gesundheitsdaten: Implementierung und zuständige Akteure • Auswahl von interessanten Initiativen für Deutschland • Ggf. Zusätzliche Dateninnovationen der letzten Jahre • Publikationen, die als Fallbeispiele für die Datennutzung dienen können 	<ul style="list-style-type: none"> • Primäres Ziel: Forschung, Governance/Qualitätssicherung, Versorgungskoordination, Vergütung usw. • Welche Daten werden erhoben und wer erhebt sie? Sind erhobene Daten strukturiert/ unstrukturiert? • Auf welcher Ebene werden Daten aggregiert (regional/national/?) und wer ist dafür zuständig? • Werden Plausibilitätsprüfungen durchgeführt? Durch wen? Wie ist die Datenqualität? • Wo sind Daten lokalisiert/abgespeichert und für wie lang? Wem gehören die Daten? („ownership“) • Können die Daten über ein UPI verlinkt werden oder werden andere Methoden eingesetzt? Wer ist für die Verlinkung zuständig? • Datenzugang: Wer ist berechtigt? Wie sieht der Prozess aus? Wie lange dauert es, bis Daten zur Verfügung gestellt werden können? Welches ist aktuell das letzte verfügbare Datenjahr? • Bisherige Nutzung der Daten für Forschungszwecke? • Wie repräsentativ ist die durch die Daten abgedeckte Bevölkerung für das Land/die Region? • Für nicht-individuelle Daten: welche ist die kleinste Einheit, für die eine Verlinkung möglich ist? Wie viele dieser Einheiten gibt es?

Informationssammlung und -validierung: Das weitere Vorgehen wurde in Anlehnung an die etablierte Praxis von internationalen Vergleichen am Fachgebiet Management im Gesundheitswesen der Technischen Universität Berlin geplant. Informationen zu den in Tabelle 2.2 aufgeführten Fragen wurden für alle Vergleichsländer bzw. Fallstudien zunächst auf Basis von öffentlich zugänglichen Quellen identifiziert, vorwiegend basierend auf Online-Recherchen nach publizierter und grauer Literatur, sowie nach offiziellen Informationen aus der Internetpräsenz der jeweils beteiligten Akteure des Gesundheitssystems. Für die allgemeinen Informationen wurden zunächst die Gesundheitssystemprofile des Commonwealth Fund sowie die Reihe „Health Systems in Transition“ des European Observatory on Health Systems and Policies herangezogen, ergänzt zumeist durch offizielle Dokumente über Datenregulierungen und -strategien der einzelnen Länder. Informationen zu den spezifischen Fragen pro Fallstudie wurden als erstes über die jeweilige Internetpräsenz aufgesucht, ergänzt von aktuellen publizierten Beschreibungen der entsprechenden Datengrundlagen sowie von Informationen aus publizierten Versorgungsforschungsstudien, die auf Basis der zu untersuchenden Datenbestände durchgeführt wurden. Anschließend wurde pro Land ein Profil erstellt und geeigneten Länderexperten zur Validierung und Ergänzung weitergeleitet. Passende Sachverständige wurden aus den Netzwerken des Fachgebiets Management im Gesundheitswesen und des Wissenschaftlichen Beirates des Zentralinstituts für die Kassenärztliche Versorgung identifiziert.

Die folgenden Kapitel stellen die Erkenntnisse in deskriptiver Synthese dar. Nach einer Übersicht der eingeschlossenen Fallstudien und deren Inhalt (Kapitel 3) erfolgt eine nähere Betrachtung der Möglichkeiten der Nutzung von Routinedaten (Kapitel 4), der Verlinkungsmöglichkeiten mit und ohne eindeutiges personenbezogenes Merkmal (Kapitel 5), der unterschiedlichen Zugriffswege für Forscher (Kapitel 6), der Ausgestaltung des Einsatzes von Daten aus elektronischen Patienten- bzw. Gesundheitsakten für die Forschung (Kapitel 7), und ein Zwischenfazit (Kapitel 8). Das Gutachten endet mit einer Wunschliste für mögliche regulatorische Änderungen in Anbetracht der aktuellen politischen Diskussion (Kapitel 9).

Der Anhang enthält pro Land wichtige Eckdaten (u.a. Bevölkerung und administrativer Aufbau des Landes, Ausgaben für Gesundheit, Aufbau der Leistungserbringung, Datenschutzbestimmungen, Grad der Technologischen Entwicklung) sowie die wichtigsten Informationen zu den einzelnen Fallstudien in Steckbrief-Format. Detailliertere Informationen sind auf Anfrage vom Fachgebiet Management im Gesundheitswesen erhältlich.

Kapitel 3: Übersicht der eingeschlossenen Länder und Fallstudien

Dieser Kapitel liefert eine Übersicht der eingeschlossenen Fallstudien über die Vergleichsländer hinweg; der Fokus liegt hierbei auf der Zielsetzung der jeweiligen Initiativen, auf ihren allgemeinen Aufbau sowie auf den Umfang und die Art der Daten(bestände), die sie zusammenbringen. Tabelle 3.1 am Ende des Kapitels fasst diese Informationen zusammen und weist auf weitere Stellen im Gutachten, die die jeweilige Fallstudie näher betrachten. Eine Darstellung jeder Fallstudie in Steckbrief-Format ist dem Anhang zu entnehmen.

Australien

Australien hat eine föderale Verwaltungsstruktur mit sechs Bundesstaaten, drei Territorien und sieben Außenterritorien. Das Gesundheitssystem ist steuerfinanziert mit universellem Zugang zu Gesundheitsleistungen, die über das **Medicare Benefit Schedule** (MBS) und das **Pharmaceutical Benefit Schedule** (PBS) abgerechnet werden. Medicare versteht sich als ein „öffentliches Gesundheitsversicherungssystem“ (Glover 2019). Private Zusatzversicherung ermöglicht unter anderem eine größere Auswahl an Leistungserbringern, schnelleren Zugang oder niedrigere Zuzahlungen; fast die Hälfte aller Australier macht davon Gebrauch. Die Finanzierung von Medicare wird von der Bundesregierung gesichert, die Organisation und Erbringung der Gesundheitsversorgung ist jedoch dezentralisiert. Hausärzte haben Steuerungsfunktion („gatekeeping“), aber Patienten müssen sich nicht mit spezifischen Hausärzten registrieren (vgl. Anhang, S. 81). Leistungserbringer in der Primärversorgung werden hauptsächlich auf Einzelleistungsbasis vergütet („Fee-for-service“); öffentlich finanzierte Krankenhäuser über globale Budgets und Fallpauschalen, private auf Einzelleistungsbasis.

- Der MBS pflegt ein aktuelles und historisches Datenrepositorium aller MBS-subventionierten, abgerechneten Leistungen (ambulant und stationär); ähnlich hält das PBS die Daten über alle durch das PBS finanzierten, abgerechneten Verordnungen sowie über Verordnungen, die günstiger als die Zuzahlungsgrenze sind und somit der Eigenverantwortung der Bürger obliegen. Die MBS- und PBS-Daten sind versichertenbezogen verfügbar, über die „Medicare Number“ verknüpft, und werden regelmäßig für die Steuerung des Versicherungsprogramms analysiert, zum Beispiel hinsichtlich Inanspruchnahme und Kosten, zur Budgetplanung sowie für die Entwicklung von „risk-sharing agreements“ mit Arzneimittelherstellern.
- Die MBS- und PBS-Daten sind auch für die Versorgungsforschung zugänglich und mit anderen Datenbeständen verknüpfbar, über das Angebot der regionalen Verlinkungsstellen und des **Population Health Research Network** (PHRN), welches das nationale Netzwerk dieser Stellen bildet. Darüber hinaus ist das PHRN ein gesichertes Datenlaboratorium („secure data laboratory“) und ein Koordinationszentrum für elektronische Leistungen zur Datenverarbeitung. Es wurde im Jahr 2009 im Rahmen einer Initiative der Bundesregierung zur kollaborativen Forschungsinfrastruktur gegründet. Während die regionalen Datenbestände teilweise variieren, ist in den verlinkten Datensammlungen des PHRN ein breites Spektrum von Daten abgebildet:

gesundheitsbezogene Daten, einschließlich der MBS/PBS-Abrechnungsdaten, Geburten, Todesfälle, Neugeborenen- und Krebsregister sowie begrenzte klinische Daten und weitere Registerdaten; dazu kommen Daten zu Wohnverhältnissen, Justiz und Bildung. Diese Investition der australischen Regierung in der Verlinkungsinfrastruktur habe zu einem Zuwachs der Nutzung verlinkter Daten für Forschungszwecke geführt (Young & Flack 2018).

- Eine neuere Initiative in Australien stellt die Umsetzung der elektronischen Patientenakte „**my Health Record**“ (MyHR) dar, die erst 2016 pilotiert wurde. Australische Bürger hatten bis Ende Januar 2019 die Möglichkeit auszuschneiden, wenn sie nicht teilnehmen wollten („opt out“), und 10 % der Bevölkerung haben von der „opt-out“-Option Gebrauch gemacht. MyHR soll neben den Routinedaten, die im Rahmen der Versorgung generiert werden, auch Laborbefunde und die Ergebnisse bildgebender Verfahren einschließen, sowie Dokumentation zu Versorgungsplanung und Entlassmanagement umfassen. Die primäre Motivation für die Umsetzung von myHR lag in der Verbesserung der Versorgungscoordination („care coordination“) und der Förderung des Patienten-Empowerment. Nichtsdestotrotz ist die Ermöglichung von Forschung im öffentlichen Interesse und mit dem Ziel der Verbesserung der Versorgung auch unter den Zielen der Initiative genannt. Bis dato sind noch nicht alle technischen Details geklärt; geplant ist ein Komitee für Data Management, welches die Beschaffung, Verarbeitung und Vorbereitung der myHR-Daten regeln soll. myHR-Nutzer sollen darüber hinaus die Möglichkeit haben, von der Nutzung ihrer Daten für die Forschung auszuschneiden, auch wenn sie die Patientenakte in Anspruch nehmen wollen.

Dänemark

In Dänemark wird die Gesundheitsversorgung ebenfalls durch die fünf Regionen, die die mittlere Verwaltungsebene im Land darstellen (Staat, Regionen, Kommunen), geplant, koordiniert und erbracht. Die Sicherung der Gesundheitsversorgung der Bevölkerung über öffentliche Krankenhäuser und die Finanzierung privater Leistungserbringer im ambulanten Bereich entspricht auch der Hauptzuständigkeit der Regionen insgesamt. Hingegen fallen Maßnahmen der Gesundheitsförderung und Prävention sowie der öffentlichen Primärversorgung unter die Verantwortung der Kommunen. Das Gesundheitssystem ist steuerfinanziert mit universellem Zugang; fast 40 % der Bevölkerung entscheidet sich für eine private Zusatzversicherung zur Deckung von Zuzahlungen oder nicht erstatteten Leistungen, während ungefähr 25 % Pakete für eine größere Auswahl an privaten Leistungserbringern abschließt. Hausärzte haben für die große Mehrheit der Bevölkerung eine „gatekeeping“-Funktion für andere ambulante Leistungen und für die stationäre Versorgung (vgl. Anhang, S. 87). Primärversorgung wird auf Basis von Einzelleistungsvergütung bzw. Kopfpauschalen vergütet, Krankenhäuser über globale Budgets und Fallpauschalen. Das dänische Gesundheitssystem weist einen hohen Digitalisierungsgrad auf – schon 1994 wurde die Entwicklung einer eHealth-Strategie in Auftrag gegeben, die den elektronischen Austausch von relevanten Daten zwischen Leistungserbringern ermöglichen sollte. Alle Krankenhäuser und Hausärzte arbeiten inzwischen mit elektronischen Gesundheitsakten: Befunde werden

elektronisch übermittelt, ebenso fast alle Überweisungen und Arzneimittelverordnungen. Daten in der jeweiligen Gesundheitsakte werden lokal gespeichert auf Basis regionaler eHealth-Anwendungen; Daten aus den regionalen Plattformen werden im Dänischen Netzwerk für Gesundheitsdaten zusammengeführt. Alle Einwohner in Dänemark haben eine Personenkennzahl, die Zahl für das Zentrale Personenregister (*Centrale Personregister, CPR*), die auch in allen Registern herangezogen wird. Statistics Denmark ist der Hauptanbieter von Registerdaten für Forschungszwecke und pflegt regelmäßig aktualisierte Informationen auf aggregierter Ebene, die kostenfrei zugänglich sind. Der Zugang zu Mikrodaten setzt eine entsprechende Qualifizierung der anfragenden Forschergruppen voraus.

- Im Jahr 2001 haben die fünf dänischen Regionen, zusammen mit der nationalen Regierung, das nationale eHealth-Portal **Sundhed.dk** gegründet, welches das Zugangsportale zu den Daten des zugrunde liegenden Netzwerkes für Patienten und Leistungserbringer (ambulante Leistungserbringer, Krankenhäuser und Apotheker) darstellt (siehe Abbildung 3.1). Unter anderem können Patienten sich über ihre Diagnosen, Behandlungen und Befunde informieren, Termine buchen, Verordnungen erneuern und die eigene Adhärenz beobachten; darüber hinaus können sie bestimmen, welche Leistungserbringer welche ihrer persönlichen Daten einsehen können. Leistungserbringer können wiederum zusätzlich über die Plattform miteinander kommunizieren, Leistungen abrechnen, und relevante Evidenz (Sachbücher, wissenschaftliche Artikel, klinische Leitlinien) aufrufen.

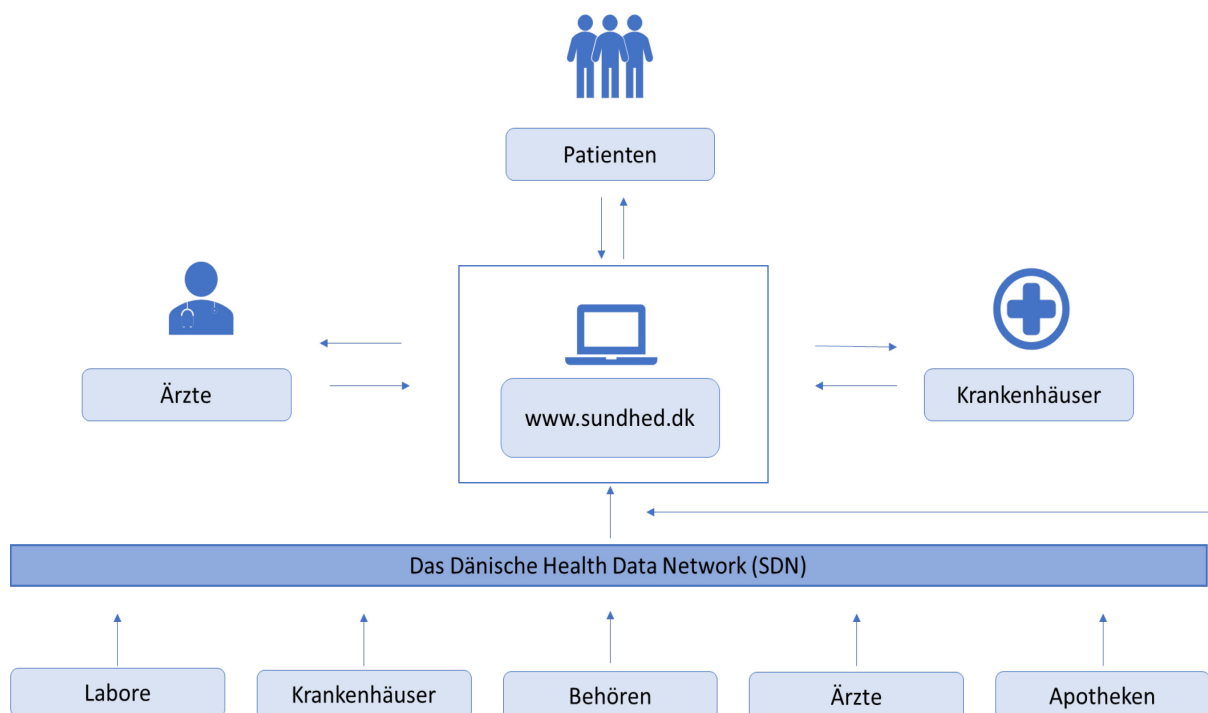


Abbildung 3.1 Sundhed.dk: Datenfluss und Akteure (Quelle: www.sundhed.dk)

- Dänemark hat, wie auch die anderen skandinavischen Länder, eine lange Tradition in der Führung und Nutzung von klinischen und epidemiologischen Registern für die Steuerung der Gesundheitsversorgung und die Verbesserung der Versorgungsqualität. Das Regionale Programm

für die Entwicklung Klinischer Qualität (**Regionernes Kliniske Kvalitetsudviklingsprogram, RKKP**) verwaltet 80–85 klinische Register mit unterschiedlichen Zielen und Populationen (z.B. Schlaganfall, Lungenkrebs, Demenz, Intensivversorgung), die unter anderem klinische Daten, Diagnosen, Behandlungen, Verordnungen, Laborbefunde, patientenberichtete Endpunkte und soziodemographische Informationen beinhalten. Die Register berichten regelmäßig (täglich oder monatlich) für Kliniker und Manager. Aggregierte Daten werden auf jährlicher Basis veröffentlicht (siehe Abbildung 3.2). Kliniker haben Zugang auf Mikrodaten für Patienten, die in ihrer Praxis bzw. Klinik behandelt werden, zur Validierung und Qualitätsverbesserung; andere Kliniker, Verwaltungspersonal und die Öffentlichkeit bekommen nur aggregierte Auswertungen der in RKKP enthaltenen Daten. Die Teilnahme ist vorgeschrieben für alle Leistungserbringer, die Patienten mit den passenden klinischen Indikationen behandeln; für das Programm ist keine Einwilligung der Patienten zur Datenverarbeitung erforderlich.

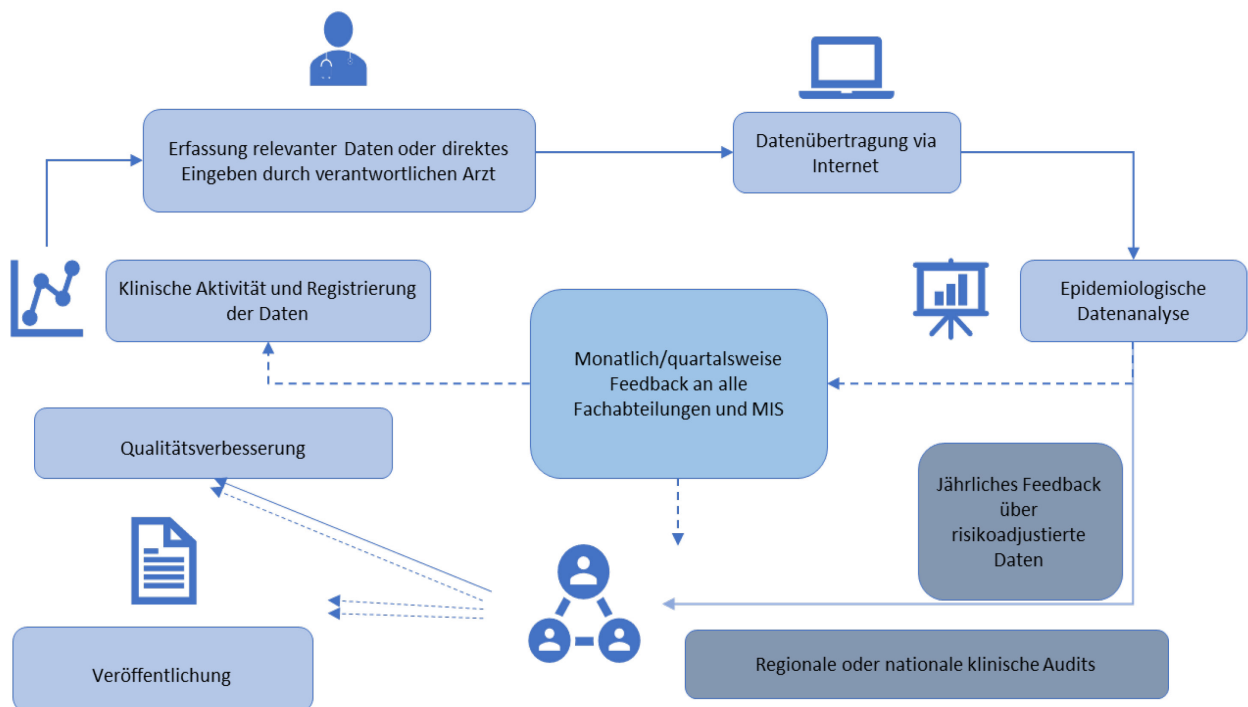


Abbildung 3.2 Kernprozesse der Danish Clinical Registries (Quelle: <https://www.rkkp.dk>)

Finnland

Das finnische Gesundheitssystem ist in seiner Verwaltungsstruktur hoch dezentralisiert, durch mehrere Mechanismen finanziert (Steuer, Versicherungsbeiträge, private Ausgaben) und kombiniert drei parallele Systeme für die Primärversorgung: das kommunale System, das System der nationalen Gesundheitsversicherung und die arbeitsmedizinische Versorgung. Das Hauptsystem wird durch die Kommunen organisiert, die für die Primärversorgung zuständig sind und die Sekundär- und Tertiärversorgung durch die Mitgliedschaft in 20 Krankenhausregionen mitverantworten. Hausärzte haben auch hier eine Steuerungsfunktion („gatekeeping“) für spezialisierte ambulante und stationäre Versorgung, aber Patienten können auswählen, wo sie die entsprechenden Leistungen bekommen (vgl. Anhang, S. 93).

Durch den hohen Dezentralisierungsgrad war die Interoperabilität zwischen den breit eingesetzten elektronischen Gesundheitsakten bisher niedrig (Keskimäki et al. 2019); das nationale elektronische Gesundheitsdatenrepositorium (Kanta) wurde zwischen 2010 und 2016 stufenweise umgesetzt und umfasst elektronische Daten zu Verordnungen, Behandlungen, Befunden und Inanspruchnahme von Gesundheitsleistungen. Leistungserbringer haben darüber Zugang auf landesweiten Daten, Patienten über ihre eigenen Daten aus der Gesundheits- und Sozialversorgung. Mehr als 2 Millionen Finnen machen inzwischen davon Gebrauch. Kanta wird von der Sozialversicherungsanstalt Kela gepflegt; die Finnische Gesundheitsbehörde THL ist verantwortlich für die Strukturierung der Daten nachdem diese von den einzelnen Kontrolleuren eingetroffen sind. Das neue Gesetz für die Sekundäre Nutzung von Gesundheits- und Sozialdaten (2019) hat die Behörde **Findata** etabliert, die auf vorhandenen Datenbeständen, inklusive Kanta, aufbaut und mehrere Ziele für die Verbesserung der Patientenversorgung verfolgt („*Steuerung, Überwachung, Forschung, Statistik und die Weiterentwicklung der Sektoren Gesundheit und Soziales*“). Findata wurde im Anschluss an die finnische Präsidentschaft des Europäischen Rates 2020 gelauncht und. Die Behörde hat die Verantwortung, auf Basis von erfolgreichen Forschungsanträgen Stichproben aus existierenden Datenbeständen zu sammeln und zu verlinken und den beantragenden Forschern Zugriff auf dem verlinkten Datensatz zu gewährleisten.

Kanada

Die zehn kanadischen Provinzen und drei Territorien haben im Land die Hauptzuständigkeit für die Organisation und die Erbringung von Versorgungsleistungen sowie für die Supervision der Leistungserbringer. Sie dürfen die eigenen Versicherungssysteme mit universellem Zugang betreiben, die zentralstaatlichen Vorgaben (dem Canada Health Act von 1986) folgen müssen. Finanziert wird die öffentliche Gesundheitsversorgung über Steuergelder. Ungefähr zwei Drittel aller Kanadier haben eine private Zusatzversicherung, die nicht erstattete Leistungen abdeckt, wie Arzneimittel oder zahnmedizinische Behandlungen. Hausärzte haben Steuerungsfunktion und können von Patienten frei gewählt werden. Das „Gatekeeping“ basiert hier vorwiegend auf finanziellen Anreizen: Fachärzte bekommen eine niedrigere Vergütung für Patienten ohne Überweisung. Die Primärversorgung wird vorwiegend über ein Einzelleistungsmodell („fee-for-service“) vergütet, Krankenhäuser (die Sekundär- und Tertiärversorgung anbieten) über globale Budgets (vgl. Anhang, S. 97).

- Um die Nutzung der schon vorliegenden Mikrodaten aus unterschiedlichen Sektoren in Zusammenführung zu ermöglichen, wurde unter anderem beim Statistischen Bundesamt, Statistics Canada, das **Social Data Linkage Environment (SDLE)** gegründet. Über ein zentral gepflegtes „Schlüsselregister“, welches auf Basis einer Reihe von identifizierenden Merkmalen erstellte neue Identifikatoren enthält (vgl. Kapitel 5, ist ein breites Spektrum von Datensätzen zu verknüpfen, inklusive der Arbeitkräfteerhebung (mit soziodemographischen bzw. -ökonomischen Mikrodaten), nationalen Gesundheitsbefragungen (mit gesundheitlichen, versorgungsrelevanten, soziodemographischen bzw. -ökonomischen Mikrodaten), der Krankenhausentlassung-Datenbank (mit Mikrodaten zu erbrachten Leistungen und Diagnosen), Zen-

sus-Daten, umweltbezogenen Makrodaten und einer Reihe von Provinzdaten über Kriminalität, Bildung, psychische Gesundheit usw. Demnach ist SDLE keine integrierte Datenbank an sich, sondern eine sichere Umgebung, welche die Verlinkung anderer Datenbestände ermöglicht; diese wurden bisher auch häufig für wissenschaftliche Forschungsarbeiten genutzt.

Auch auf der Ebene der Provinzen gibt es wichtige Verlinkungsinitiativen, die in diesem Gutachten herangezogen werden. Als Vorteile dieser Initiativen gelten die große Abdeckung der jeweiligen Bevölkerung (Repräsentativität) und die lange Zeitspanne bestimmter Datenbestände:

- Das ehemalige Institute for Clinical Evaluative Sciences in Ontario wird inzwischen nur mit seinem Akronym **ICES** benannt und hält ein Repositorium von Datensätzen, die für Forschungs- und Analysezwecke zur Verbesserung der Gesundheit der Einwohner von Ontario verlinkbar sind; 94 davon sind öffentlich sichtbar im Datenrepositorium. Dazu zählen Daten zu den Interaktionen der Einwohner von Ontario mit dem Gesundheitssystem (Mikrodaten zur Gesundheitsversorgung), Mikrodaten aus Befragungen zu Gesundheit, Inanspruchnahme von Leistungen, Soziodemographie und Ökonomie, Migrationsdaten (soziodemographisch und -ökonomische Mikrodaten), und Registerdaten (u.a. Krebs-, Schlaganfall-, und Mukoviszidoseregister). Makrodaten aus dem Zensus liefern soziodemographische Informationen auf aggregierter Ebene, die über geographische Merkmale mit Mikrodaten verknüpft werden können. Darüber hinaus ermöglicht ICES die Verknüpfung von Daten aus neuen Forschungsprojekten, die auch außerhalb des Institutes entstanden sind, zu den vorhandenen Datenbeständen.
- Im Gegensatz zu ICES hat **Population Data British Columbia (PopData BC)** keine eigene Forschungsagenda, ermöglicht es aber auch externen Forscher die eigenen Primärdaten mit den in seiner gesicherten Forschungsumgebung gehaltenen Datenbeständen zu verknüpfen. Diese Datensätze umfassen Mikrodaten zur Inanspruchnahme von Leistungen (inklusive Arzneimittelverordnungen und Krankenhausentlassungen), Vitalstatistiken, Migrationsdaten, arbeitsmedizinische Abrechnungsdaten und Bildungsdaten. Einkommensdaten werden per Postleitzahl aggregiert und in Strata aufgeteilt. Auch dieses Angebot wurde bisher in mehreren Forschungsvorhaben eingesetzt.
- Letztlich wird in diesem Gutachten auch das **Manitoba Population Research Data Repository** an der Universität Manitoba herangezogen; die Universität forscht selbst auf Basis der Daten und kann auch Externen den Zugang ermöglichen. Die Daten stammen von mehreren lokalen, provinziellen und nationalen Quellen und enthalten Gesundheitsdaten auf individueller Ebene (aus mehreren Registern), Gesundheitsversorgungsdaten, Daten zu Bildung und Justiz, Zensusdaten und eine Reihe von spezifischeren Sätzen (vgl. Anhang, S. 102). Die Daten werden de-identifiziert, bevor sie das Repositorium erreichen (für einen Vergleich unterschiedlicher Ansätze zur De-Identifikation von Mikrodaten siehe Kapitel 5). Nur aggregierte Daten dürfen in Publikationen und Präsentationen benutzt werden.

Neuseeland

Neuseeland hat ein nationales, steuerfinanziertes Gesundheitssystem, welches durch die zwanzig Bezirksgesundheitsbehörden („District Health Boards“) verwaltet wird. Es ist ein universeller Zugang zu Gesundheitsleistungen für Einwohner erreicht; ungefähr ein Drittel der Bevölkerung hat eine private Zusatzversicherung, vorwiegend um Zuzahlungen bzw. nicht erstattete Leistungen abzudecken. Eine Registrierung bei einem Hausarzt ist nicht zwingend erforderlich und Patienten haben zumindest theoretisch freie Wahl. Hausärzte haben auch hier Gatekeeping-Funktion und werden zu 50 % durch Kopfpauschalen aus dem öffentlichen System und zu 50 % auf Basis von Einzelleistungen über Patientenzahlungen vergütet. Krankenhäuser werden über globale Budgets finanziert (vgl. Gauld 2019). Daten aus dem Gesundheitssystem fließen zusammen mit Daten aus anderen Bereichen (Sozialdienst, Bildung, Einkommen und Beschäftigung, Wohnverhältnisse, Justiz, Soziale Verhältnisse, Bevölkerungsstatistik) in eine integrierte Dateninfrastruktur (Integrated Data Infrastructure, IDI), die vom neuseeländischen statistischen Amt gepflegt wird. Die IDI ist grundsätzlich eine große Forschungsdatenbank, die mit dem Ziel gegründet wurde, die bestmögliche Nutzung vorhandener Daten für die Beantwortung komplexer Fragestellungen zu unterstützen. Die IDI-Datenbestände mit Gesundheitsbezug sind Box 3.1 zu entnehmen.

Innerhalb von Datenbereichen können identifizierende Merkmale genutzt werden, um Daten auf der Mikroebene zu verlinken (zum Beispiel der „National Health Index Number“ für den Gesundheitssektor); um Daten über unterschiedliche Bereiche hinweg zu verlinken wird ein probabilistisches Verfahren eingesetzt (vgl. Kapitel 5). Es gibt die Möglichkeit, über geographischen Bezug auch andere Daten, die nicht in der IDI aufbewahrt sind, mit IDI-Daten zu verknüpfen, wie zum Beispiel Umweltdaten. Die IDI benutzt das sogenannte „five safes“-Framework und Forscher, die Zugang auf IDI-Daten haben möchten, müssen sich daranhalten: „safe people, safe projects, safe settings, safe data, safe output“. Die Daten der IDI werden bereits von mehreren Forschergruppen in Anspruch genommen (Milne 2019). Unter den im Gutachten betrachteten Fallstudien vertritt die IDI den höchsten Grad der Konsolidierung von Daten.

Box 3.1 Gesundheitsbezogene Daten in der IDI

- Ergebnisse aus dem Gesundheitscheck im Vorschulalter (4 Jahre, „B4 School Checks“) – ab 2011
- Krebsregistrierungen – ab 1995
- Chronische Erkrankungen – ab 2007
- Abrechnungsdaten der Hausärzte – ab 2002
- „Health tracker“ (eine Art Gesundheitszensus des Gesundheitsministeriums) – 2006-14
- Impfungen – ab 2006
- Laborabrechnungen – ab 2003
- Leistungen für Wochenbett/Neugeborene – ab 2003
- Todesursachenstatistik – ab 1988
- National Booking Reporting System – ab 2003
- Pflegebedürftigkeit (SOCRATES)
- Abrechnungsdaten für ambulante Patienten – ab 2007
- Verordnungen – ab 2005
- Registrierungen mit Hausärzten – ab 2003
- Demographische Informationen aus dem National Health Index – ab 2004
- Psychische Gesundheit (nur Überweisungen, abgerechnete Leistungen und demographische Informationen; keine Diagnosen oder Befunde) – ab 2008
- Private Krankenhausentlassungen – ab 2001
- Öffentlich finanzierte Krankenhausentlassungen – ab 1988

Niederlande

Das niederländische Gesundheitssystem wird durch Versicherungsbeiträge und Steuergelder finanziert. Alle Einwohner müssen eine gesetzliche Gesundheitsversicherung mit einer privatrechtlichen Krankenkasse abschließen. Die Beiträge werden zentral gesammelt und auf Basis einer Risikoadjustierungsformel an die Kassen verteilt. Ungefähr 85 % der Bevölkerung hat darüber hinaus eine private Zusatzversicherung, die Zuzahlungen bzw. nicht erstattete Leistungen (wie Zahnmedizin) abdeckt. Eine Überweisung vom Hausarzt ist für ambulante fachärztliche und stationäre Versorgung notwendig („gatekeeping“). Obwohl die Versicherten keine Registrierungspflicht haben, schreiben sich die meisten bei ihren Hausärzten freiwillig ein. Diese werden über eine Kombination von Kopfpauschalen und leistungsorientierten Zahlungen vergütet. Fachärztliche Versorgung ist Krankenhausbasiert und wird je nach konkreter vertraglicher Situation entweder auf Einzelleistungsbasis oder über Fallpauschalen finanziert. Alle Patienten verfügen über eine eindeutige Kennziffer („Burgerservicenummer“). An einem zentralen Informationsverwaltungssystem wird gearbeitet (vgl. Wammes et al. 2019); fast alle Hausärzte und Krankenhäuser benutzen elektronische Gesundheitsakten, aber die Interoperabilität ist verbesserungsbedürftig.

- Die Forschungsinstitution Nivel betreibt eine Datenbank für die Primärversorgung (**Nivel Primary Care Database** (Nivel PC Database; *Nivel Zorgregistraties eerste lijn*), welche routinemäßig gesammelte Daten zu gesundheitlichen Problemen und zur Inanspruchnahme (Diagnosen, Ergebnisse klinischer Untersuchungen) von Gesundheitsleistungen aus der Primärversorgung zusammenbringt, um die Gesundheit und die Gesundheitsversorgung der Bevölkerung zu monitoren. Das Patientenkollektiv des Panels der teilnehmenden Praxen wird als repräsentativ gesehen; neben den Hausärzten nehmen auch andere Leistungserbringer der Pri-

märversorgung teil, inklusive Apotheken, Physiotherapeuten, Psychotherapeuten und Ernährungsberater. Teilnehmende Leistungserbringer bekommen regelmäßige Berichte mit aggregierten Daten und Nivel nutzt die Datenbank für eigene Forschungszwecke; auch externe Forscher können die Nutzung der Daten beantragen, die in anonymisiertem Format bekommen können. Es ist möglich, die Daten über Pseudonyme auf Basis der Burgerservicenummer mit anderen Datenbeständen und soziodemographische Informationen zu verknüpfen.

- Eine weitere interessante Möglichkeit für die Nutzung von verknüpften Daten stellt in den Niederlanden **PHARMO** dar, eine unabhängige Forschungsorganisation, die seit ihrer Gründung im Jahr 1999 ein breites Netzwerk an Datenbanken aufgebaut hat. Hier können je nach Forschungsfrage bzw. Antrag angepasste, indikationsspezifische „Kohorten“ erstellt werden und Mikrodaten aus mehreren Quellen verlinkt werden, inklusive Daten aus Hausarztpraxen, ambulanten und stationären Apotheken, Laboren, Krebs-, Mortalitäts- und Neonatalregister, Pathologiebefunden, Krankenhausentlassungen und patientenberichtete Endpunkte. PHARMO forscht selbst zu bestimmten Fragestellungen, hauptsächlich hinsichtlich Pharmakoepidemiologie und der Wirkung von Arzneimitteln nach Marktzugang; externe Forscher können die Daten kostenpflichtig beantragen und erhalten einen maßgeschneiderten Datensatz zur Bearbeitung.

Österreich

Das Gesundheitssystem in Österreich wird aus mehreren Quellen finanziert: der Bund, die neun Länder sowie die Sozialversicherungsträger tragen zum Gesamtbudget bei. Einkünfte für das System stammen aus den allgemeinen Steuereinnahmen (40 %) und den Beiträgen der gesetzlich vorgeschriebenen sozialen Krankenversicherung (60 %), die mehr als 99 % der Bevölkerung umfassen. Die Verfügbarkeit der Krankenhausversorgung (ambulant und stationär) wird von den Ländern gesichert und größtenteils finanziert. Die ambulante Primär- und Sekundärversorgung wird durch Gesamtverträge zwischen dem Hauptverband der Versicherungsträger und den Krankenkassen geregelt. Hausärzte haben keine formale „Gatekeeping“-Funktion und ein Großteil der Versorgung findet im Krankenhaus statt. Patienten haben freie Wahl an Hausärzten, Fachärzten und Krankenhäusern. Die krankenkassenfinanzierte ambulante Versorgung wird über eine Mischung aus Kopfpauschalen und Einzelleistungsvergütung finanziert; privat praktizierende Ärzte müssen zunächst direkt von den Patienten bezahlt werden, die Kosten können aber nachträglich und anteilig von den Krankenkassen zurückerstattet werden. Akutkrankenhäuser werden auf Basis von Fallpauschalen vergütet. Die österreichische e-Health-Strategie von 2007 hatte das Ziel, die Integration und Kompatibilität von bestehenden Informationssystemen unterschiedlicher zu verbessern und einen stabilen Informationsaustausch über die elektronische Versichertenkarte zu ermöglichen.

- Als Grundbestandteil der e-Health Strategie 2007 wurde auch die **Elektronische Gesundheitsakte (ELGA)** eingeführt, die es ermöglichen sollte, Gesundheitsinformationen von Patienten (z. B. medizinische Untersuchungen, Arzneimittelverordnungen, Blutgruppe, Labor-

und Radiologiebefunde, Entlassungsbriefe aus dem Krankenhaus) den Patienten selbst und berechtigten Leistungserbringern strukturiert zur Verfügung zu stellen und historisch darzustellen. Die gesetzliche Grundlage für die Einführung von ELGA wurde im Januar 2013 geschaffen und 2014 wurde das ELGA-Portal aktiviert. Über das Portal können Patienten und Leistungserbringer auf die beschriebenen Daten zugreifen. Patienten können den Zugang zu ausgewählten persönlichen Gesundheitsdaten einschränken und einsehen, wer ihre individuellen Eintragungen aufgerufen hat. Patienten werden für ELGA automatisch angemeldet (kein „opt in“ erforderlich, vgl. Kapitel 7), können sich aber wieder abmelden. Das vollständige, landesweite Rollout von ELGA wird 2021 erwartet (Bachinger et al. 2018).

- Das **Österreichische Gesundheitssystem (ÖGIS)** wird für die Gesundheitsberichterstattung und Planung eingesetzt und als geografisches Informationssystem (GIS) konzipiert. Die im System verfügbaren Daten können aggregiert als Karten und Zeitreihen dargestellt werden und umfassen Bereiche wie Sterblichkeit, Krebsinzidenz, Prozeduren, Krankenhausdichte, subjektiver Gesundheitszustand). Konkret sind Mikrodaten aus dem DRG-System (inkl. Soziodemographischen Informationen, Diagnosen, abgerechnete Leistungen), aus der ambulanten Versorgung (Arztbesuche, abgerechnete Leistungen) und Pflegeleistungen eingeschlossen. Makrodaten aus der Todesursachenstatistik, Krebsstatistik, Krankenhausstatistik, Statistik der ambulanten Leistungserbringer, Befragungen zu Gesundheitsverhalten, subjektiver Gesundheitseinschätzung und Nutzung des Gesundheitssystems (aggregiert nach Alter, Geschlecht und Region der Befragten), demographische und sozioökonomische Daten sowie Daten zur Zugänglichkeit sind ebenso eingeschlossen. Das **Regionale Gesundheitssystem (REGIS)** ist Teil von ÖGIS und ermöglicht die regionale Analyse der ÖGIS-Daten in der Form von Karten.
- Vor dem Hintergrund der aktuellen Diskussion in Deutschland ist eine weitere österreichische Initiative hier erwähnenswert: **das e-Rezept** wird im Jahr 2020 pilotiert, um Papierrezepte abzulösen. Die Daten aller Arzneimittelverordnungen sollen ins System der elektronischen Gesundheitskarte einfließen, in zwei separaten Datenzentren gespeichert werden und bei Bedarf zusammengeführt werden. Die Zugänglichkeit der Daten für die Forschung ist wegen der frühen Implementierungsphase noch nicht ausreichend geklärt.

Schweden

In Schweden sind es die 21 Provinzen (12 Provinziallandtage und 9 Regionalbehörden), die für die Bereitstellung der Gesundheitsversorgung zuständig sind. Kommunen hingegen tragen die Verantwortung für Pflege und Sozialleistungen. Einkünfte für das System kommen von lokalen und regionalen Steuern; ein System für die regionale Neuverteilung der Einkünfte erlaubt es den Regionen, die Bedürfnisse ihrer Einwohner zu erfüllen. Nur ca. 10 % der Bevölkerung haben private Zusatzversicherungspakete für schnelleren Zugang zu fachärztlicher Versorgung sowie nicht erstattete Leistungen (Wahlleistungen). Die Primärversorgung von Patienten wird von privaten (40 %) oder öffentlichen

(60 %) Leistungserbringern (Hausärzten, Pflegekräften, Hebammen, Psychologen, Physiotherapeuten und Gynäkologen, die häufig in Gemeinschaftspraxen organisiert sind) getragen. Eine Registrierung ist in den meisten Regionen erforderlich aber Hausärzte haben keine strenge Steuerungsfunktion („Gatekeeping“); meistens registrieren Patienten sich bei einer (Gemeinschafts-) Praxis eher als bei einem Arzt oder einer Ärztin. Leistungsanbieter in der Primärversorgung werden über eine Mischung aus hauptsächlich Kopfpauschalen, Einzelleistungsvergütung, und leistungsorientierten Zahlungen vergütet. Ambulante und stationäre fachärztliche Versorgung findet in öffentlichen Krankenhäusern oder privaten Kliniken statt; sie wird über Fallpauschalen (DRGs) vergütet, die um leistungsorientierte Komponenten sowie Kosten- bzw. Volumendeckelungen ergänzt werden. Ähnlich wie in Dänemark und in anderen nordischen Ländern hat Schweden eine gut etablierte Datenbasis in Form von epidemiologischen und klinischen Registern.

- Um die Forschung in klinischen Registern, auch mit Daten aus mehreren Datenbeständen zu ermöglichen, hat der Schwedische Forschungsrat die Stelle **Registerforskning** etabliert, die Forscher dabei unterstützen soll, die richtigen Daten zu finden und zu verarbeiten. Ein elektronisches Tool (**Register Utiliser Tool, RUT**) wird aktuell entwickelt, um diesen Prozess noch weiter zu vereinfachen; das Tool wird nicht direkt auf Mikrodaten zugreifen können. Die Register, die über Registerforskning erreichbar sind, umfassen Daten zu Gesundheit und Gesundheitsversorgung, Steuer, Bildung, Justiz, Soziodemographie usw. Die persönliche Identifikationsnummer (PIN) kann als identifizierendes Merkmal für die Verlinkung von Daten über mehrere Register hinweg fungieren.
- Über die Registerforschung hinaus bietet auch die elektronische Patientenakte Schwedens (**Journalen, ehem. SUSTAINS**) interessante Impulse. Leistungserbringer in Schweden haben freie Wahl, welche elektronische Gesundheitsakte sie nutzen. Um die Interoperabilität zwischen unterschiedlichen Systemen zu ermöglichen, hat Schweden eine nationale Plattform, die *Health Information Exchange* (HIE), implementiert, die es Patienten erlaubt, alle ihre Informationen als eine virtuelle Patientenakte zu sehen, obwohl sie aus unterschiedlichen Systemen stammen. Alle Daten, die im Rahmen gesetzlich finanzierter medizinischer und zahnmedizinischer Leistungen erhoben wurden, fließen in Journalen ein, einschließlich verordneter Medikamente, Laborbefunde, Diagnosen, Überweisungen, Impfungen, und Wochenbett. Leistungserbringer können sich weigern, ihre Daten für Journalen bereitzustellen, auch wenn die Provinz sich für die Umsetzung der Akte entschieden hat.

Südkorea

Das **südkoreanische** Gesundheitssystem ist grundsätzlich ein nationales Krankenversicherungssystem; lediglich die Ausgaben für 3 % der Bevölkerung werden durch Steuermittel finanziert. Die Nationale Krankenversicherung (National Health Insurance, NHI) erhält einkommensbasierte Beiträge von Arbeitgeber und Arbeitnehmer bzw. Selbstständigen und finanziert entsprechend die Gesundheitsausgaben Ihrer Versicherten. Obwohl der Leistungskatalog relativ breit ist, können die Zuzahlungen

für bestimmte Leistungsarten hoch sein. Mehr als 60 % der Bevölkerung haben eine private Zusatzversicherung für Zuzahlungen und nicht erstattete Leistungen. Die Primärversorgung findet in Versorgungszentren statt, die vorwiegend privat sind und über Einzelleistungsvergütung finanziert werden. Auch in Krankenhäusern tätige Ärzte können für die Primärversorgung der Patienten Leistungen erbringen. Akutkrankenhäuser werden neben Einzelleistungen über Fallpauschalen für sieben Diagnosegruppen vergütet; Pflegeleistungen werden über Tagespauschalen finanziert. Es gibt keine Steuerungsfunktion („Gatekeeping“) in der Primärversorgung. Der Dienst der nationalen Krankenversicherung (**National Health Insurance Service, NHIS**) zieht die Beiträge ein und vergütet Leistungserbringer, während der Prüfungs- und Bewertungsdienst der Krankenversicherung (**Health Insurance Review and Assessment Service, HIRA**) die beanspruchten abgerechneten Leistungen prüft, die Angemessenheit der geleisteten Versorgung bewertet und für die Gestaltung des Leistungskatalogs und der Zuzahlungen zuständig ist. Als zentrale Stellen für die Bearbeitung der abgerechneten Leistungen für die versicherte Bevölkerung Südkoreas haben NHIS/HIRA umfangreiche Daten zu Untersuchungen, Behandlungen, verordneten Arzneimitteln, Prozeduren und Diagnosen von ca. 50 Millionen Menschen, zusätzlich zu deren soziodemographischen Charakteristika und Informationen zu den beteiligten Leistungserbringern. Die HIRA-Daten sind mit anderen Datenbeständen verlinkbar, so dass fehlende Informationen (Befunde, Todesursachen, Krankheitsschwere usw.) ergänzt werden können. Die HIRA-Daten wurden erst 2009 für Forschungszwecke freigeschaltet; seitdem hat ihre Nutzung zugenommen, so dass Südkorea unter den Ländern mit dem höchsten OECD-Index ist (vgl. Tabelle 1.2). Mögliche Themenbereiche für die Forschung auf Basis von HIRA-Daten umfassen Adhärenz, Verschreibungsmuster, Inanspruchnahme von Leistungen, Krankheitslast, unerwünschte Ereignisse, und die Bewertung von Politikmaßnahmen. Obwohl HIRA-Daten eine „Anamnese“ der ambulanten und stationären Versorgung der Versicherten von Geburt an beinhalten (vgl. Kapitel 4), sind nur die Daten der jeweils letzten 5 Jahre für Forschungszwecke freigeschaltet.

Vereinigtes Königreich

Der National Health Service (NHS) im vereinigten Königreich wurde mit dem Anspruch gegründet, dass Gesundheitsversorgung den Bedarf aller Einwohner entspricht, bedarfsgerecht und unabhängig der Zahlungsfähigkeit von Patienten erbracht wird, und zum Zeitpunkt der Inanspruchnahme von Leistungen kostenfrei ist. Seit seiner Gründung 1948 ist das System entlang dieser Grundprinzipien weiterentwickelt worden. Es besteht aus vier eigenständigen Organisationen für die jeweiligen Landteile des Vereinigten Königreichs (*NHS England, NHS Wales, NHS Scotland, Health and Social Care Northern Ireland*), die unabhängig funktionieren, aber den gemeinsamen Grundprinzipien folgen. Das System wird über Steuergelder finanziert und ist bekannt für die zentrale Rolle der Hausärzte in der Primärversorgung, die als Eingangsstelle für Patienten und „Gatekeeper“ für die weitere Versorgung im System agieren. In England sind es NHS England und die mehr als 200 Clinical Commissioning Groups (CCGs) und in Schottland die insgesamt 21 NHS Boards, die für die Planung und Bereitstellung von Versorgungsleistungen zuständig sind. In England werden Hausärzte über eine Mischung

von Kopfpauschalen, Einzelleistungsvergütung und leistungsorientierten Zahlungen finanziert, während Krankenhäuser über Fallpauschalen vergütet werden, mit gesonderten Budgets für psychische Gesundheit, Ausbildung und Forschung. In Schottland wurde das sogenannte „Purchaser-Provider-Split“ nicht umgesetzt, so dass die Vergütung von Leistungen integrierter verläuft.

- Die wichtige Rolle der Hausärzte spiegelt sich auch in der international bekannten Datenbank **Clinical Research Practice Datalink (CPRD)** wider, welche Daten aus Hausarztpraxen im Vereinigten Königreich auf Basis der jeweiligen elektronischen Praxisinformationssysteme zusammenführt. CPRD wird von der Nationalen Arzneimittel- und Medizinprodukteagentur (Medicines and Healthcare products Regulatory Agency, MHRA) und dem Nationalen Zentrum für Gesundheitsforschung (National Institute for Health Research, NIHR) getragen. Der CPRD sammelt de-identifizierte Mikrodaten von einem Netzwerk aus Hausarztpraxen, die sich für eine Teilnahme entschieden haben und die jederzeit auch wieder ausscheiden können (vgl. Kapitel 4). Die gesammelten Daten beinhalten demographische Informationen, Angaben zu Leistungserbringern, klinischem Erscheinungsbild (Diagnose, Symptome, klinische Zeichen), Überweisungen, Impfungen, Laborbefunde und Arzneimittelverordnungen. Daten aus englischen Praxen können mit weiteren Datensätzen verknüpft werden, so dass der Versorgungsvorgang jenseits der Primärversorgung erforscht werden kann; Praxen müssen vorab einer Teilnahme im Verlinkungsprogramm zugestimmt haben. Die Datenbestände, die aktuell zur Verlinkung herangezogen werden können, sind in Box 3.2 aufgeführt. Die Verlinkung übernimmt NHS Digital; die personenbezogene NHS-Nummer kann als eindeutiges identifizierendes Merkmal eingesetzt werden, aber entsprechend der zu verlinkenden Daten werden auch andere Merkmale herangezogen (vgl. Kapitel 5). Im Jahr 2019 umfassten die CPRD Daten 42 Millionen Patienten und wurden für mehr als 2000 extern begutachtete wissenschaftliche Publikationen genutzt, die unter anderem Fragen der Arzneimittelsicherheit, des Arzneimittelkonsums, der Effektivität von gesundheitspolitischen Maßnahmen, der Organisation der Leistungserbringung und der Entwicklung von Risikofaktoren erforscht haben.
- Schottland hat in den letzten Jahren mit einer Reihe von strategischen Maßnahmen großen Wert auf die Weiterentwicklung und Nutzung von routinemäßig erhobenen Daten gelegt. Der Dienst für Forschung mit elektronischen Daten und Innovation (**electronic Data Research and Innovation Service, eDRIS**) stellt ein „One-Stop-Shop“ für die Nutzung vorhandener administrativer und gesundheitlicher Daten dar. Es soll Forscher und Politikanalytiker dabei helfen, die richtigen Datensätze für ihr Vorhaben zu identifizieren, die Forschungsfragen zu operationalisieren und Zugang zu verlinkten Daten zu bekommen. Alle administrativen Daten im öffentlichen Besitz können für die Forschung beantragt werden, oder auch für Planungs- bzw. Evaluationsprozesse im Gesundheitswesen. Die Verlinkung erfolgt über eine Vertrauensstelle auf Basis eines sogenannten „Population Spine“ (vgl. Kapitel 5), welches die Stammdaten aller Personen enthält, die mit NHS Scotland in Kontakt gekommen sind.

Box 3.2 Datenbestände, die mit CPRD-Daten verlinkt werden können

- Hospital Episode Statistics (HES) Daten
 - für stationäre Patienten
 - HES-Daten für ambulante Leistungen
 - HES- Daten aus den Notfallambulanzen
 - HES Diagnostic Imaging Dataset
- HES Patient Reported Outcomes Measures (PROMs)
- Todesregistrierungen des statistischen Amtes
- Krebsdaten von Public Health England (PHE):
 - Krebsregistrierungen
 - Systemic Anti-Cancer Therapy (SACT) Dataset
 - National Radiotherapy Dataset (RTDS)
 - Cancer Patient Experience Survey (CPES)
 - Quality of Life of Cancer Survivors in England: Pilot Survey (QOLP)
 - Quality of Life of Colorectal Cancer Survivors in England: Patient Reported Outcome Measures Survey (QOLC)
- Psychische Gesundheit (Mental Health Dataset, MHDS)
- Sozioökonomische Daten mit kleinräumigem Bezug: Index of Multiple Deprivation (IMD), Townsend Deprivation Index, Carstairs Index, Patient postcode linked deprivation measures, Practice postcode linked deprivation measures, Rural-Urban Classification

- Analog zum CPRD (siehe oben) wurde in Schottland das **Scottish Primary Care Information Resource (SPIRE)** gegründet, um Daten aus schottischen Hausarztpraxen für Forschungs- und Analysefragen zusammenzuführen. Diese können über eDRIS mit weiteren Datenbeständen des schottischen NHS verknüpft werden. Unter anderem sollen SPIRE-Daten zur Bestimmung der Versorgungsqualität herangezogen werden und den Vergütungsrahmenvertrag (GMS Contract) informieren (vgl. das beendete Quality and Outcomes Framework, ISD Scotland 2019a). SPIRE-Daten werden auf Anfrage aus dem Praxisinformationssystem teilnehmender Praxen extrahiert (die Software wird von der Firma MSDi erstellt und betrieben) und umfassen soziodemographische Daten der Patienten, Befunde, Prozeduren, Arzneimittelverordnungen, Diagnosen, Inanspruchnahme von Leistungen. Nur strukturierte Informationen können für Forschungszwecke extrahiert werden, nie die gesamte elektronische Gesundheitsakte; somit werden zum Beispiel Notizen aus Arzt-Patienten-Diskussionen nicht weitergegeben. Potentiell identifizierende Daten müssen pseudonymisiert werden, bevor sie die Praxis verlassen (vgl. Kapitel 5). Alle Praxen in Schottland nehmen grundsätzlich in SPIRE teil, es sei denn sie entscheiden sich auszuschneiden („opt-out“).

Vereinigte Staaten

Das US-Amerikanische Gesundheitssystem ist für seine Fragmentierung hinsichtlich der Finanzierung und der Organisation der Leistungserbringung und des fehlenden universellen Zugangs zu Gesundheitsleistungen bekannt. Zwei Drittel der Bevölkerung haben eine private Krankheitsversicherung mit variierenden Leistungskatalogen und Zuzahlungsanforderungen. Die Mehrheit der privat versicherten Personen sind über arbeitgeberfinanzierte Pläne abgedeckt (56 % der Bevölkerung im Jahr 2017, vgl.

Commonwealth Fund 2019). Ungefähr ein Drittel der Bevölkerung ist über die öffentlichen Sozialversicherungssysteme abgedeckt; die wichtigsten sind Medicare für Personen über 65 Jahre (17 % der Bevölkerung) und Medicaid für Personen mit niedrigem Einkommen (19 % der Bevölkerung). Durch den 2010 verabschiedeten Affordable Care Act wurde der individuelle Erwerb von Gesundheitsversicherungspaketen für alle sonst nicht versicherte Personen eingeführt und dies hat zu einer Abnahme der Nichtversichertenquote beigetragen (8,8 % im Jahr 2017). Das Mandat wurde inzwischen wieder abgeschafft. In der Primärversorgung sind überwiegend in kleinen Einzel- oder Gruppenpraxen tätige Ärzte aktiv; sie haben meistens keine „Gatekeeping“-Funktion und Patienten haben freie Wahl, solange die Leistungserbringer dem Netzwerk ihrer Versicherung zugehören. Die Vergütung variiert je nach Versicherungssystem (vgl. Rice et al. 2013). In der Primärversorgung umfasst sie vorwiegend Einzelleistungszahlungen (mit regulierten oder verhandelten Gebühren, für die öffentlich bzw. privat finanzierten Leistungen entsprechend), Kopfpauschalen und leistungsorientierte Anreize (in manchen Versicherungssystemen). Fachärztliche Versorgung findet in Arztpraxen und Krankenhäusern statt. Große Unternehmen bieten integrierte Versorgungssysteme an, die Gesundheitsversicherung, ambulante und stationäre Gesundheitsversorgung im „Managed Care“-Modell zusammenbringen. Das größte davon ist aktuell Kaiser Permanente, mit Niederlassungen in acht Staaten und mehr als 12 Millionen Versicherten.

Ähnlich wie der Aufbau des Systems ist auch die Verfügbarkeit von Daten für die Forschung fragmentiert; nichtsdestotrotz mangelt es nicht an Initiativen, die der Transparenz und der Ermöglichung von Forschungs- und Analyseaktivitäten dienen sollen. Selektiv werden in diesem Gutachten folgende weiter beleuchtet:

- Die **Centers for Medicare and Medicaid Services** sind Teil des Department of Health and Human Services (entspricht dem Ministerium für Gesundheit und Soziales auf Bundesebene) und pflegen mehr als 400 Datensätze (**CMS Data**) zum öffentlich geregelten Versorgungsgeschehen im amerikanischen System. Diese umfassen sozio-ökonomische/demographische, Gesundheits- und Gesundheitsversorgungsdaten auf Individual- und Makroebene. Viele dieser Datensätze sind auf aggregierter Ebene frei zugänglich als „public use files (PUFs)“ über data.cms.gov. Deidentifizierte Mikrodaten sind entweder als „limited data set (LDS)“ oder als „research identifiable files (RIFs)“ auf Antrag zugänglich; letztere können je nach Forschungsfrage maßgeschneidert werden, bevor sie die Institution verlassen und ihre Nutzung ist viel kostenintensiver im Vergleich zu den LDS (vgl. Abbildung 4.1). Die CMS-Datenbanken beinhalten mehrere identifizierende Merkmale, die für eine Verlinkung herangezogen werden können; CMS veröffentlicht aber kein Data Dictionary dazu (vgl. Kapitel 4).
- Auf Basis der elektronischen Gesundheitsakten (EPIC) der bei Kaiser Permanente erfassten Patienten bringt das **Kaiser Health Connect System** Informationen aus unterschiedlichen Ebenen zusammen: klinische Befunde, Untersuchungen, Diagnosen und Behandlungen mit Terminbuchungen, Registrierungen und Abrechnungsdaten. Es ist somit das größte private klinische Informati-

onssystem der Welt. Es bietet zwei Oberflächen an: MyChart, auf dem Patienten ihre Informationen einsehen können; und EpicCare, das von Gesundheitspersonal für die koordinierte Versorgung des Patienten eingesetzt wird.

- Der Affordable Care Act von 2010 hat unter anderem das Patient-Centred Outcomes Research Institute (PCORI) gegründet. PCORI leitet die Initiative **PCORnet** (National Patient-Centered Clinical Research Network); diese hat das Ziel, die Durchführung von klinischer Forschung auf Basis großer Datenbestände und starker partizipativer Netzwerke zu vereinfachen und zu optimieren, indem auch die Bedürfnisse der Patienten expliziter berücksichtigt werden. PCORnet stellt ein Dachnetzwerk von sogenannten “Research Networks” dar; aktuell (2019) sind neun von diesen bei klinischen Versorgungssystemen und zwei bei Versicherungsorganisationen angesiedelt; die ursprünglich gleichermaßen eingebundenen patientengesteuerten Netzwerke werden ab April 2019 beratend herangezogen. PCORnet funktioniert über ein gemeinsames Datenmodell (Common Data Model, CDM), welches für die Aufbereitung der Daten teilnehmenden Partner ausschlaggebend ist. Das PCORnet Coordinating Center verwaltet das sogenannte „Front Door“, welches es potentiellen Datennutzern erlaubt, über die Möglichkeit der PCORnet Daten beraten zu werden und Nutzungsanträge zu stellen. Die PCORnet-Daten umfassen demographische Variablen, Beobachtungsdauer im Netzwerk, Kontakte mit dem Gesundheitssystem, Diagnose, Prozeduren und diagnostische Tests, Vitalparameter, verordnete Arzneimittel, Laborbefunde, Gesundheitszustand (diagnostiziert und selbstberichtet), patientenberichtete Endpunkte, Todesfälle und Todesursachen, Informationen über die eingebundenen Leistungserbringer, ggf. Daten aus klinischen Studien zu den eingeschlossenen Patienten.
- Die Verlinkungsinitiative **SEER-Medicare** verknüpft die von den Krebsregistern des SEER-Programms (*Surveillance, Epidemiology, and End Results*) erfassten Daten zu Inzidenz, Staging und Überlebensraten von Krebserkrankungen mit den Abrechnungsdaten des Medicare-Programms (siehe oben). Die SEER-Register werden vom National Cancer Institute finanziert und stellen den umfassendsten bevölkerungsbasierten Datenbestand der Staaten in der Onkologie dar. Daten umfassen demographische Informationen, Tumor-bezogene Daten inklusive Identifizierung und Staging sowie versorgungsrelevante Abrechnungsdaten. Mögliche Forschungsfragen, die anhand der verknüpften Daten beantwortet werden können, umfassen das Versorgungsgeschehen für Patienten vor und nach einer Krebsdiagnose (unmittelbar und mit Langzeitperspektive), die Nutzung von Diagnostika und Prozeduren sowie die Kosten der Krebsversorgung.
- Auch im Bereich der Onkologie versteht sich **CancerLinQ** als eine „big data analytics platform“, die auf Basis von Daten aus der alltäglichen Versorgung von Krebspatienten die Versorgungsqualität verbessern möchte. CancerLinQ ist eine gemeinnützige Tochtergesellschaft der amerikanischen Fachgesellschaft für klinische Onkologie (American Society of Clinical Oncology, ASCO) und wird somit von Onkologen geleitet; Kliniker, die die Datenbank abonnieren, können patientenbezogene Informationen aufrufen. Für Forschung stehen nur de-identifizierte Daten zur Verfü-

gung, die je nach Projektantrag bereitgestellt werden. Von mehr als 100 teilnehmenden Praxen erhebt CancerLinQ folgende Daten: Patienten und Praxeninformationen, Diagnosen, Pathologiebefunde, Staging, Vitalparameter, klinische Bewertungen, Laborbefunde, Radiologiebefunde und Bildgebung, Behandlungspläne, verordnete und vor Ort dargereichte Medikamente, Radiotherapiebehandlungen, Operationen, weitere Outcomes sowie die Teilnahme an klinischen Studien. Es gibt darunter auch unstrukturierte Daten, die eine Kuration erfordern. CancerLinQ-Daten dürfen nicht mit anderen Datenbeständen verlinkt werden.

Zusammenfassung und weitere Betrachtung einzelner Aspekte im Gutachten

Die dargestellten Fallstudien lassen sich grob in fünf Kategorien einteilen:

- i) große administrative Datensätze, die es aufgrund der zentralen Zulaufstelle ermöglichen, Fragen über den Versorgungsprozess einzelner Patienten sektorenübergreifend abzubilden und Möglichkeiten zur Verlinkung mit anderen Daten aufweisen (z.B. MBS/PBS, NHIS/HIRA, ÖGIS, CMS);
- ii) zentrale Stellen, die vorhandene Datenbestände unterschiedlicher Datenhalter und mit variablem Inhalt zusammenführen; hier zeichnen sich zwei Möglichkeiten aus: die Daten bleiben bei den Datenhaltern und werden auf Antrag verknüpft (z.B. PHRN, SDLE, eDRIS) oder in einer konsolidierten Datenbank aufgehoben (z.B. IDI);
- iii) Datenbestände, in denen routinemäßig erhobene Daten zusammengeführt und für Kliniker und Forscher zur Verfügung gestellt werden (z.B. CPRD, Nivel PC Database, SPIRE, Kaiser Health Connect);
- iv) elektronische Plattformen, die es (unter anderem) Patienten erlauben, Zugriff auf ihre eigenen Informationen zu gewinnen und die ggf. als potentielle Grundlage für klinische Daten für die Forschung betrachtet werden könnten (z.B. ELGA, myHR bzw. Sundhed.dk, Journalen);
- v) Tools, die die Forschung über mehrere Datensätze hinweg erleichtern (z.B. RKKP, Registerforskning).

Die folgenden Kapitel des Gutachtens greifen selektiv auf einzelne Fallstudien zu, um wichtige Aspekte für eine Verbesserung der Situation in Deutschland hervorzuheben (Näheres siehe Kapitel 2).

Tabelle 3.1 Übersicht der Fallstudien und deren Inhalte (Datenkategorien korrespondieren zu den Blöcken in Abbildung 2.1)

Land	Fallstudie	Primäres Ziel	Mikrodaten			Makrodaten				Weitere
			SD	G	GV	SD	G	GV	U	
Australien	PHRN	Forschung – Ermöglichung der Zusammenführung von Daten aus ganz Australien	✓	✓	✓	✓	✓	✓	(✓) *	✓
	MyHR	Versorgungskoordination , Patientenempowerment, Unterstützung von Forschung	✓	✓	✓					
	MBS/PBS	Abrechnungsdaten werden für Forschung bereitgestellt um Steuerung im Gesundheitssystem zu ermöglichen	✓		✓					
Dänemark	RKKP	Verbesserung der klinischen Versorgung , sekundär Versorgungsforschung	✓	✓	✓					
	Sundhed.dk	Koordination von Leistungserbringern, Patientenempowerment, Transparenz	✓	✓	✓					
Finnland	Findata	Forschung: One-stop-shop für die Nutzung von sozialen und gesundheitlichen Sekundärdaten	✓	✓	✓	✓	✓	✓		
Kanada	SDLE	Ermöglichung der Forschung für das ganze Land ohne Bedarf an zusätzlicher Datenerhebung	✓	✓	✓	✓	✓	✓	✓	✓
	ICES	Unterstützung der Versorgungskontinuität , und Forschung für Gesundheitsplanung	✓	✓	✓	✓				✓
	PopData BC	Forschung über Gesundheitsdeterminanten, Wohlbefinden und Entwicklung	✓	✓	✓	✓				✓
	Manitoba Population Research Data Repository	Forschung zur Information von Entscheidungen in der Gesundheits- und Sozialpolitik	✓	✓	✓	✓				✓
Neuseeland	IDI	Forschung zu multidimensionalen Fragen, Monitoring der Effektivität von Maßnahmen und Leistungen	✓	✓	✓					✓

Land	Fallstudie	Primäres Ziel	Mikrodaten			Makrodaten				Weitere
			SD	G	GV	SD	G	GV	U	
Niederlande	PHARMO	Forschung – Nutzung von Real World Evidence, Innovationsförderung im Gesundheitswesen		✓	✓					
	Nivel PC Database	Forschung über Entwicklungen in Gesundheit und Inanspruchnahme von Leistungen der Primärversorgung auf Basis von Routinedaten	✓	✓	✓					
Österreich	ELGA	Versorgungskoordination: Unterstützung einer besseren Versorgung, wenn mehrere Leistungserbringer eingebunden sind	✓	✓	✓					
	ÖGIS/REGIS	Gesundheitsmonitoring, -berichterstattung, und -planung (Steuerung im System)	✓		✓	✓	✓	✓		
	eRezept	Vereinfachung von Verwaltungsprozessen, Patientenempowerment			✓					
Schweden	Registerforskning	Unterstützung von Forschung durch die Bereitstellung von Informationen zu Registerdaten (Gesundheit, Steuer, Justiz, Bildung usw.)	✓	✓	✓					
	Journalen	Patienteninformation und -empowerment, Versorgungsqualität	✓	✓	✓					
Südkorea	NHIS/HIRA	Bewertung und Verbesserung der Versorgungsqualität durch die Unterstützung von Gesundheitssystemsteuerung auf Basis von Abrechnungsdaten	✓		✓					
UK	CPRD	Forschung über Primärversorgung und Möglichkeit der Verlinkung mit anderen Daten	✓	✓	✓	(✓)				
UK - Schottland	eDRIS	Forschung und Politikanalyse: Information über und Bereitstellung von (verlinkten) Daten	✓	✓	✓	✓	✓	✓		✓
	SPIRE	Forschung und Versorgungsoptimierung durch die Zusammenführung von Daten aus der Primärversorgung	✓	✓	✓					

Land	Fallstudie	Primäres Ziel	Mikrodaten			Makrodaten				Weitere
			SD	G	GV	SD	G	GV	U	
USA	CancerLinQ	Forschung – Zusammenführung von Beobachtungsdaten ("real world cancer data") aus mehreren Quellen		✓	✓					
	CMS Data	Transparenz für Versorgungssteuerung und -Forschung	✓	✓	✓	✓	✓	✓		
	Kaiser Health Connect	Verbesserung der klinischen Versorgung für individuelle Patienten, Forschung zur Verbesserung der zukünftigen Versorgung insgesamt	✓	✓	✓					
	PCORnet	Forschung – zentrale Anlaufstelle für weiter dezentral gehaltene Daten von Leistungserbringern, Krankenversicherern und sog. „patient-powered research networks“	✓	✓	✓					
	SEER-Medicare	Forschung über die Versorgung von Krebspatienten	✓	✓	✓					

*Manche der regionalen Verlinkungseinheiten ziehen auch Umweltdaten heran

Kapitel 4: Erweiterte Möglichkeiten auf Basis von Routinedaten

Für den Primärzweck integriert vorliegende Daten

Drei der Fallstudien heben die Möglichkeit hervor, für Forschungsfragen auf vorhandene, integrierte Routinedaten zurückzugreifen. Damit sind Datensätze gemeint, welche bereits für den primären Nutzungszweck in einer integrierten (verknüpften) Form über mehrere Leistungsbereiche erstellt werden. Ein aus Deutschland vergleichbares Beispiel sind dabei die Datensätze, welche den Krankenkassen auf Grundlage der Abrechnungsdaten vorliegen. Diese sind auf Versichertenebene, liegen über mehrere Jahre vor und decken verschiedene Leistungsbereiche ab; die Generierung dieser Datensätze erfolgt durch die jeweilige Krankenkasse und wird primär für Abrechnungszwecke bzw. interne Analysen der Krankenkassen genutzt. Abzugrenzen sind davon Datensätze, welche erst für Forschungszwecke (sekundärer Nutzungszweck) erstellt bzw. verknüpft werden (diese werden im nächsten Abschnitt vorgestellt). In Australien, Südkorea sowie den USA konnten Möglichkeiten der Datennutzung von integrierten Routinedaten identifiziert werden (vgl. Tabelle 4.1). Es handelt sich dabei in allen Ländern um Datensätze, welche sich überwiegend aus Abrechnungsdaten (Claimsdata) zusammensetzen; die tragenden Institutionen unterscheiden sich entsprechend der jeweiligen Gesundheitssystemstruktur.

In Australien sind insbesondere die beiden großen Datensätze Medicare Benefit Schedule (MBS) und Pharmaceutical Benefit Schedule (PBS) zu nennen. In beiden Datensätzen sind Informationen zu allen erbrachten Leistungen enthalten, welche unter MBS bzw. PBS fallen. Da fast die gesamte australische Bevölkerung den Zugang zu gesundheitlicher Versorgung über MBS/PBS erhält, sowie ein Großteil der Leistungen vollständig oder anteilig abgedeckt sind, können die Datensätze als sehr umfangreich bzw. „vollständig“ betrachtet werden. Ein Vergleich mit primär erhobenen Daten hat gezeigt, dass die PBS-Abrechnungsdaten eine gute Alternative für die Erforschung des Versorgungsgeschehen darstellen (Harris et al. 2017), obwohl sie tendenziell die Nutzung unterschätzen (Gisev et al. 2018). Zudem ist in beiden Datensätzen die Medicare-ID enthalten, welche eine Verlinkung zwischen beiden Datensätzen sowie auch zu anderen Daten ermöglicht (siehe dazu PHRN, weiter unten). Der Zugang zu den Daten ist für die Forschung klar reguliert und die Daten werden auch dementsprechend für wissenschaftliche Analysen genutzt.

In Südkorea fließen in die Datenbank der nationalen Versicherungsagentur alle Abrechnungsdaten für über 50 Millionen Versicherte (vgl. Kapitel 3). Diese Daten werden im Zuge der Vergütung von Leistungserbringern generiert: Abrechnungen werden elektronisch von Leistungserbringern an HIRA übermittelt, und werden dort geprüft. Die Abrechnungsdaten der erstatteten Claims landen im „Data Warehouse“ und werden dort für fünf Jahre gespeichert. Der Zugang für Forschungszwecke wurde im Jahr 2009 freigegeben. Entsprechend des Speicherungszeitraumes sind aktuell nur Projekte mit Daten der letzten 5 Jahre machbar, obwohl eine Erweiterung auf 10 Jahre angestrebt ist. Identifizierungsziffern von Versicherten und Leistungserbringern werden verschlüsselt abgespeichert, um persönliche Informationen zu schützen. Da sich die HIRA-Daten auf Abrechnungsdaten (also Gesundheitsversorgungsdaten) beschränken, haben bisher einzelne Forschungsprojekte weitere Datenbestände (z.B. die

nationalen Gesundheitssurveys mit Gesundheitsdaten) herangezogen und mit verlinkten Daten gearbeitet. Das Linkage wurde dabei beispielsweise anhand regionaler Variablen oder basierend auf mehreren Variablen, welche in beiden Datensätzen vorkamen (z.B. Geburtsdatum, Geschlecht etc.), vorgenommen (vgl. Beispiele in Tabelle 4.1). Mikrodaten stehen Forschern in Universitäten und staatlichen Behörden zur Verfügung, bedingt und in engerem Umfang auch Interessenten von der Industrie und dem Privatsektor, also z.B. auch von privaten Versicherungsfirmen. Diese Zugänglichkeit für eigenütziges Interesse ist nicht ohne Kontroverse (vgl. z.B. Song 2017). Der Zugang zu den unbeschränkten Mikrodaten ist nach Vereinbarung und auf Basis eines Antrag- und Reviewprozesses über sieben Forschungszentren oder Fernzugang möglich. Darüber hinaus bietet HIRA auch bestimmte Stichproben zur Krankenhausversorgung (13 % der Versicherten), zur älteren Bevölkerung (20 %) und der pädiatrischen Bevölkerung (10 %), die ohne Reviewprozess zugänglich sind (Kim et al. 2014). Aggregierte Daten sind als zusammenfassende Statistiken öffentlich zugänglich.

Eine ähnliche Skalierung der Zugriffsmöglichkeiten ist auch bei den Routinedaten der öffentlichen Versicherungsprogramme in den USA zu finden (siehe Abbildung 4.1). Exemplarisch werden hier die Centers for Medicare and Medicaid Services (CMS) Daten von Medicare beschrieben (vgl. Kapitel 3 zum Kontext im Gesundheitssystem). CMS erlaubt Forscher Zugang zu den Medicare-Abrechnungsdaten; diese umfassen Kontakte mit dem Gesundheitssystem, therapeutische Behandlungen inkl. Arzneimittel, Prozeduren, und weitere Leistungen. Eine Übersicht der mit Medicare-Daten durchgeführte Studien aus dem Zeitraum 1979-2016 zeigt eine wesentliche Zunahme der Nutzung, vor allem in den Jahren ab 2010 (im Jahr 2016 alleine 205 Publikationen; Mues et al. 2017). Die untersuchten Fragestellungen umfassen die Erforschung von Muster von Morbidität, Mortalität und Krankheitslast, die vergleichende Einschätzung des Effektes medikamentöser Therapien, die Abbildung der Kosten unterschiedlicher Versorgungspfade, die Analyse der Effekte des Verhaltens von Leistungserbringern auf die Versorgung sowie des Einflusses von gesundheitspolitischen Maßnahmen auf Leistungserbringer und Patientenendpunkte. Somit ermöglichen diese Daten eine umfassende Einsicht in die Versorgung älterer Patienten. Um den Zugang zu den (richtigen) Daten zu ermöglichen und eine optimale Nutzung der vorhandenen Datenbestände zu unterstützen, kooperiert CMS mit dem Research Data Assistance Center (ResDAC), welches als Eingangstor und Wegweiser für Interessenten dient. ResDAC ist auch dafür zuständig, Anträge auf Wissenschaftlichkeit und Machbarkeit zu prüfen (vgl. dazu Kapitel 6). Trotz der großen Abdeckung der Bevölkerung über 65 Jahre im Medicare-Programm bilden die Medicare Daten nicht das gesamte Kollektiv der dazugehörigen Patienten ab; es gibt spezifische Versicherungsoptionen, die nicht eingeschlossen sind. Darüber hinaus sind diese Daten, ähnlich wie auch andere hier beschriebene Abrechnungsdaten dadurch limitiert, dass sie nur abgerechnete und somit erstattete Leistungen abbilden.

Tabelle 4.1 Umfang relevanter Datenbestände und Beispiele von darauf basierenden Studien

Land	Datensatz	Was decken die Daten ab?	Zugriff für Forschung	Beispiele für veröffentlichte Analysen
Australien	Medicare Benefit Schedules (MBS)	Informationen zu allen erbrachten Leistungen, die vollständig oder anteilig unter den Medicare Benefit Schedule fallen, d.h. ein Großteil der ambulanten und stationären Versorgung. Von Medicare ausgenommen sind z.B. zahnärztliche Leistungen. Die Daten umfassen aktuelle und Archiv-Daten.	Ja Verknüpfung mit anderen Datenbeständen möglich Letztes verfügbares Datenjahr: 2018	i) Beschreibung der Inanspruchnahme und der Kosten pädiatrischer Versorgung und regionaler Vergleich (Freed & Allen 2018) ii) Zeitverlauf der Autismus-Diagnosen nach Geschlecht (May & Williams 2018) iii) Inzidenz, regionale Variation und Versorgungstrends für keratinozytische Neoplasien (Adelson et al. 2018)
	Pharmaceutical Benefit Schedule (PBS)	Informationen zu allen erbrachten Leistungen, die vollständig oder anteilig unter den Pharmaceutical Benefit Schedule fallen. Die Daten können über die individuelle Medicare-ID auch mit den Daten des MBS verknüpft werden.	Ja Verknüpfung mit anderen Datenbeständen möglich Letztes verfügbares Datenjahr: 2018	i) Zunahme des Einsatzes von Antikoagulantien und der Kosten nach Marktzugang neuer Medikamente (Morgan et al. 2018) ii) MBS und PBS Daten: Vorhersage der Initiierung einer Zweitlinientherapie für Diabetes in der australischen Bevölkerung (Fiorini et al. 2019)
Südkorea	National Health Insurance (NHIS) & NHIDB	Behandlungen, verordnete Arzneimittel, Prozeduren, Diagnosen, Untersuchungen, soziodemographische Charakteristika der Patienten	Ja Verknüpfung mit anderen Datenbeständen möglich (auf Basis einzelner Projekte) Letztes verfügbares Datenjahr: 2018	i) Überwachung des Einsatzes von Antibiotika im Langschnitt (Park et al. 2017) ii) Einfluss von Versicherungsschutz auf Nutzung von Antikoagulantia (Ko et al. 2018) iii) Umweltverschmutzung und Schwangerschaftshypertension (verlinkte Daten, Choe et al. 2018; für den Aufbau der benutzten NHIS-Kohorte siehe Lee et al. 2017) iv) Verlinkung klinischer Registerdaten (Schlaganfall) und HIRA-Daten (Kim et al. 2018)
USA	CMS Medicare Data	Kontakte mit dem Gesundheitssystem, therapeutische Behandlungen inkl. Arzneimittel, Prozeduren, und weitere Leistungen	Ja Verknüpfung mit anderen Datenbeständen möglich (z.B. SEER-Medicare, individuelle Projekte) Letztes verfügbares Datenjahr: 2017	i) Unterschiede in Kosten zwischen Universitäts- und anderen Krankenhäusern (Burke et al. 2019) ii) Effekt einer jährlichen Vorsorgeuntersuchung auf Screeningraten, Überweisungen, Inanspruchnahme und Kosten (Ganguli et al. 2019)









	Public Use File (PUF)	Limited Data Sets (LDS)	Research Identifiable File (RIF)
Detailgrad	Aggregierte Daten	De-identifizierte Mikrodaten	Mikrodaten
Kosten	Keine		
Individuelle Gestaltung	Keine	Keine	Gestaltbar
Dauer bis zum Erhalt der Daten	 0-2 Wochen	 2-3 Wochen	 3-5 Monate
Benötigter Speicherplatz			

Abbildung 4.1 Übersicht der CMS-Datenformate nach ResDAC (in Anlehnung an die ResDAC-Webseite)

Zusammenführung von routinemäßig entstandenen Daten in einem neuen Datenbestand

Es ist ersichtlich, dass eine gute Datengrundlage für Forschungszwecke vor allem dann vorliegt, wenn alle Informationen zur Interaktion von einzelnen Patienten mit dem Gesundheitssystem an einer zentralen Stelle gehalten werden; so können leistungserbringer- und ggf. sektorenübergreifende Versorgungsprozesse abgebildet werden, im optimalen Fall auch im Zeitverlauf. Diese Bedingung ist grundsätzlich in Gesundheitssystemen erfüllt, die ein „Gatekeeping“-System in der Primärversorgung verankert haben; ein typisches Beispiel hierfür sind die Hausärzte (general practitioners, GPs) in Großbritannien. Kapitel 3 und die Länderprofile im Anhang bieten länderspezifisch Informationen zum Thema Gatekeeping an.

Unter den eingeschlossenen Fallstudien ist das typischste Beispiel für eine Datenbank, die von GPs erfassten Daten zusammenführt, der Clinical Practice Research Datalink (CPRD; zur Entstehungsgeschichte und Einbettung im britischen System vgl. auch Gothe 2015). CPRD führt Daten von Praxen zusammen, die sich in das Programm eingeschrieben haben („opt-in“). Patienten dieser Praxen können sich wiederum von der Datenübermittlung ausschließen lassen („opt-out“), im Einklang mit der nationalen opt-out Regelung des NHS (Vgl. Kapitel 8). Die von CPRD gehaltenen Daten können für Beobachtungsstudien beantragt werden, oder zur Durchführung von experimentellen Studien unterstützend herangezogen werden. Sie können darüber hinaus mit anderen Datensätzen von NHS Digital verknüpft werden, um reichere, sektorenübergreifende Einsichten zu ermöglichen (siehe auch nächsten Abschnitt, Kapitel 3 und Kapitel 5). Bereits mehr als 2000 wissenschaftliche Studien haben CPRD-Daten genutzt.

In Schottland wurde 2016 der Scottish Primary Care Information Resource (SPIRE) gegründet, um Daten aus schottischen Praxen zusammenzubringen und schottischen GPs die Möglichkeit zu geben, sich über ihr Patientenkollektiv im Vergleich zu informieren und Versorgung besser zu planen, sowie den schottischen NHS dabei zu unterstützen, die Qualität der Versorgung zu monitoren und zu verbessern. Auch hier können Patienten die Nutzung ihrer Daten von SPIRE ablehnen. Die Initiative betont, dass es bei SPIRE nicht um eine umfassende zentrale Datenbank aller Informationen zu schottischen Patienten geht, sondern um die zielspezifische Zusammenführung von Parametern aus GP-Daten für konkrete, bewilligte Zwecke. Daten werden nicht länger gespeichert, als es für die Bearbeitung der Anträge notwendig ist und Informationen werden danach vernichtet. Die vorliegenden Informationen zu SPIRE sind noch neu und nicht vollständig (vgl. SPIRE Rollout, SPIRE 2019). Es bietet sich aber an, den Aufbau der Initiative zu vergegenwärtigen (vgl. Anhang 136) und in nächster Zeit erneut zu bewerten.

Die Datenbank zur Primärversorgung des niederländischen Forschungsinstituts Nivel setzt ebenso Daten aus der Primärversorgung ein, allerdings kommen hier die Daten aus den elektronischen Praxisverwaltungssystemen auch von Leistungserbringern jenseits der Hausärzte zusammen (vgl. Kapitel 3). In den Niederlanden wird dieses Modell auch von anderen Institutionen bzw. Initiativen genutzt (vgl. zum Beispiel Smeets et al. 2018). Allen Initiativen ist gemein, dass neben den Forschungsmöglichkeiten die Daten auch explizit für Feedback an die teilnehmenden Praxen genutzt werden, inklusive einer Möglichkeit eigene Ergebnisse mit den von anderen Praxen zu vergleichen. Ein systematisches Review zu den Einflussfaktoren für die Entwicklung von Datenbanken auf Basis von elektronischen Praxisverwaltungssystemen in der Primärversorgung englischsprachiger Länder identifizierte weitere Initiativen in England, Kanada und den USA und fand, dass die wichtigsten Voraussetzungen für eine erfolgreiche Umsetzung eine robuste technische Infrastruktur sowie starke wissenschaftliche und politische Unterstützung unabdingbar sind (Gentil et al. 2017).

Das Zi-Praxispanel folgt einem ähnlichen Prinzip im deutschen Gesundheitssystem. Ärztliche Praxen sind aufgefordert, dem Panel zuzutreten. Ziele sind die Beobachtung von Qualität und Kosten, aber auch die Nutzung der Daten für die Versorgungsforschung. Der Aufbau der fachärztlichen Versorgung in Deutschland bedeutet, dass im Gegensatz zu den oben beschriebenen Beispielen auch Fachärzte im Praxispanel teilnehmen.

Tabelle 4.2 Zusammenführung von Daten in der Primärversorgung

Land	Initiative	Abdeckung	Opt-in/ Opt-out	Zugriff für For- schung	Beispiele
UK	CPRD	Ca. 650 Pra- xen in 2016	Patienten kön- nen aus Data- sharing ausstei- gen; GPs neh- men freiwillig an der Übermitt- lung an CPRD teil. Ebenso ist die Zustimmung der Praxis für die Verlinkung der Daten not- wendig	Zugriff ist für Wissenschaftler, Gesetzgeber/ NHS-Beschäftigte und andere (inkl. Industrie) möglich	>2300 Artikel, https://www.cprd.com/bibliography i) Änderungen in den Todesur- sachen von COPD-Patienten 2005-2015 (Gayle et al. 2019) ii) Vergleichende Nutzenbe- wertung in der COPD- Arzneimitteltherapie (Suissa et al. 2019) iii) Internationaler Vergleich von „high-cost, high-need“- Patienten (Tanke et al. 2019)
UK Schottland	SPIRE	Rollout in Umsetzung	Opt-in für Pra- xen; Patienten können die Nutzung ihrer Daten aus- schließen	Ja, auf Antragsba- sis (geplant)	Noch keine
Niederlande	Nivel PC Database	Repräsentati- ve Stichprobe von ca. 10 % der Bevölke- rung	Opt-in für Pra- xen; opt-out für Patienten	Ja, auf Antragsba- sis; Nivel eigene Forschung	i) Nutzung von außerplanmäßi- ger Primärversorgung während der Pflegereform nach Ein- kommen (Jansen et al. 2018) ii) Nutzung von E-Mail- Beratung in der Primärversor- gung (Huygens et al. 2018)

Verlinkung von unabhängig entstandenen Datenbeständen

In der Recherche wurden mehrere Fallstudien identifiziert, in denen verschiedene Datenquellen mit unterschiedlichem Inhalt über eine oder mehrere Verlinkungsvariablen verknüpft werden. Dies bedeutet im Gegenzug zu den in den vorigen Abschnitten des Kapitels beschriebenen Datenbeständen, dass die Daten nicht bereits für den primären Nutzungszweck integriert erhoben werden oder zu einem neuen Datenbestand zusammengeführt werden, sondern erst für den sekundären Nutzungszweck (in diesem Fall für die Forschung) verknüpft werden müssen. Dies kann entweder notwendig sein, da die Daten für den primären Zweck nur einen sehr abgegrenzten Bereich abdecken (für Deutschland wären hier u.a. die §21er KHEntgG Daten ein Beispiel) oder da noch eine zusätzliche Erweiterung eines bereits umfangreichen Datensatzes notwendig ist, um Forschungsfragestellungen entsprechend zu beantworten (bspw. eine Verknüpfung von Daten der gesundheitlichen Versorgung mit Daten zur Berufstätigkeit/Arbeitslosigkeit oder von klinischen Daten mit Abrechnungsdaten).

Das übergreifende Ziel der hierunter fallenden Initiativen ist es, einen besseren Datenzugang für die Forschung zu ermöglichen. Dabei werden Daten aus verschiedenen Quellen, insbesondere von verschiedenen öffentlichen Institutionen verlinkt. So können bspw. Daten zur gesundheitlichen Versor-

gung mit Daten aus dem Bildungsbereich und Kriminalitätsstatistiken verknüpft werden. Die Möglichkeit des Datenzugriffs ist bei fast allen Initiativen über Antragsverfahren geregelt (siehe Tabelle 4.3). Einen Überblick über die durchgeführte Forschung auf Basis der Daten bzw. über alle positiv evaluierten Anträge wird ebenso von fast allen Initiativen auf der entsprechenden Homepage gegeben, so dass eine hohe Transparenz erzielt wird und auch für die breite Öffentlichkeit ersichtlich ist, wie die Daten genutzt werden.

Sowohl die entsprechenden Informationen in Kapitel 3 (Tabelle 3.1) sowie die in Tabelle 4.3 aufgeführten Beispiele heben die Möglichkeiten klar hervor, die sich auf Basis solcher zentralen Stellen für die Zusammenführung unabhängig entstandener Datensätze für die Versorgungsforschung eröffnen. Hier nicht näher beleuchtet, aber im Prinzip ähnliche Überlegungen gestützt, sind auch Fallstudien wie eDRIS in Schottland oder PHARMO in den Niederlanden, sowie die neue Behörde Findata Authority in Finnland. Unterschiedlich unter den hier aufgeführten Fallstudien – und relevant im Hinblick auf die Überlegungen für eine potenzielle Ideengewinnung für den deutschen Raum – ist das Ausmaß der Konsolidierung der unterschiedlichen Datenbestände. So knüpfen zum Beispiel das PHRN und das SDLE die Daten projektspezifisch zusammen, während die Quelldatensätze grundsätzlich verteilt bleiben. Im Gegensatz hinzu fügt die „integrierte“ IDI eingegangene Daten mithilfe eines „Spine“ zusammen (Vgl. Kapitel 5). Interessant im Ansatz des PHRN ist auch, dass die regionalen Behörden auf Bundesstaatsebene eigene Verlinkungsstellen haben, die im Netzwerk des PHRN kooperieren, die aber auch eigene Verlinkungen durchführen. Dieser Gedanke könnte auch für das deutsche Gesundheitssystem in Erwägung gezogen werden, vielleicht in Kombination mit einer Ausgestaltung des German Record Linkage Center.

Tabelle 4.3 Modelle mit einer zentralen Stelle, die verteilte Datensätze (unterschiedlicher Datenhalter) verknüpft

Land	Fallstudie	Was wird verlinkt und wo liegen die Daten?	Wer führt das Linkage durch?	Zugriff für Forschung	Beispiel für veröffentlichte Analysen
Australien	Population Health Research Network (PHRN)	Routinemäßig verknüpft werden Daten aus: Geburtenregister, Perinatale Daten (Schwangerschaft, Geburt etc.), Sterberegister, Krebsregister, Mental Health, öffentliche Notaufnahmen (Public Emergency Departments), private und öffentliche Krankenhausfälle. Es können auch Daten zusätzlich verknüpft werden, dabei gibt es Unterschiede nach verantwortlicher Region (Jurisdiction). Die Daten sind bei den entsprechenden Institutionen angesiedelt.	In Abhängigkeit von den angefragten Datensätzen und der zuständigen Region (Jurisdiction) gibt es eine verantwortliche Data-Linkage Unit (DLU). Die DLUs sind alle Teil des PHRN.	Zugriff für die Forschung ist klar reguliert und läuft über ein Antragsverfahren. Zudem fallen Gebühren für die Verknüpfung bzw. Datennutzung an.	Überblick über einzelne Studien: https://www.phrn.org.au/for-the-community/case-studies/ Beispiele: i) Suizidrisiko nach Inhaftierung in Queensland (Spittal et al. 2016) ii) Verbesserung der Schätzung von Mortalitätsratios in Krankenhäusern (Spilsbury et al. 2017)
Kanada	Social data linkage environment (SDLE)	Zur Verfügung stehen Daten aus der Arbeitserhebung, nationalen Gesundheitsbefragungen, Krankenhausentlassungen, Zensus-Daten, umweltbezogene Makrodaten und eine Reihe von Provinzdaten über Kriminalität, Bildung, psychische Gesundheit usw. Die Daten sind in einer gesicherten Umgebung von Statistics Canada gespeichert; es gibt eine Trennung des DRD, der Source Data files und des Key Registry (vgl. Kap 3).	Statistics Canada	Ja, Zugang zu manchen verknüpften Daten bzw. zur Datenverknüpfung kann beantragt werden.	Übersicht positiv evaluierter Anträge: https://www.statcan.gc.ca/eng/record/2018

Land	Fallstudie	Was wird verlinkt und wo liegen die Daten?	Wer führt das Linkage durch?	Zugriff für Forschung	Beispiel für veröffentlichte Analysen
Kanada – Ontario	ICES linked data	Daten aus verschiedenen administrativen Quellen, Surveydaten sowie Registerdaten werden verknüpft. Es ist ausschließlich die Bevölkerung aus Ontario in den Daten abgebildet. Die Daten sind auf den Servern von IC/ES in einem geschlossenen Computing-System gehalten.	ICES	Der Zugang ist für Forscher außerhalb des IC/ES seit 2009 möglich und setzt einen erfolgreichen Antrag voraus.	Übersicht Publikationen: https://www.ices.on.ca/Publications/Journal-Articles Beispiel: i) Luftverschmutzung und Risiko für Vorhofflimmern und Schlaganfall (Shin et al. 2019)
Kanada – British Columbia	Population Data BC (PopData BC)	Daten auf Individual-ebene, de-identifiziert und im Längsschnitt zur Bevölkerung von British Columbia. Daten kommen von verschiedenen „Partnern“ (u.a. Krebsregistern, Gesundheitsministerium, Statistics Canada). Es können eigene Datensätze der beantragenden Forscher auch verknüpft werden. Daten sind auf dem zentralen Server („Secure Research Environment“) gespeichert.	PopData BC	Ja, Zugang ist für die Forschung geregelt. Es gibt ein Antrags-system.	Liste Forschungsprojekte: https://www.popdata.bc.ca/ria/projects Beispiel: i) Neugeborenen-Outcomes bei Vätern unter Medikation für Multiple Sklerose (Lu et al. 2014)
Kanada – Manitoba	Manitoba Population Research Data Repository	Administrative, survey, register- und weitere Daten, werden für die Einwohner von Manitoba verlinkt. Daten kommen von verschiedenen Organisationen. Daten sind auf den Servern des Manitoba Center for Health Policy gespeichert.	Manitoba Population Research Data Repository	Ja, Zugang ist für die Forschung geregelt. Es gibt ein Antrags-system.	Beispiele: i) Autismus-Risiko bei pränataler Antibiotikaexposition (Hamad et al. 2019) ii) Zusammenhang zwischen den sozialen Determinanten von Gesundheit und Versorgungsqualität (Katz et al. 2018)

Land	Fallstudie	Was wird verlinkt und wo liegen die Daten?	Wer führt das Linkage durch?	Zugriff für Forschung	Beispiel für veröffentlichte Analysen
Neuseeland	Integrated data infrastructure (IDI)	Beinhaltet Mikrodaten der Einwohner und Haushalte in Neuseeland. Die Daten sind im Längsschnitt und de-identifiziert. Die Daten kommen von verschiedenen Ministerien, Befragungen sowie von Nicht-Regierungs-Organisationen. Die Daten decken ein sehr breites Feld ab, u.a. den Gesundheitsbereich, Bildungsbereich und Polizeistatistiken. Daten sind auf den Servern der IDI gespeichert (>166 Milliarden Beobachtungen)	Stats NZ	Ja, Zugang ist für die Forschung geregelt. Es gibt ein Antrags-system.	Übersicht aller Forschungsprojekte: https://cdm20045.contentdm.oclc.org/digital/collection/p20045coll17 Beispiele: i) Zusammenhang zwischen Ruralität, Freilandexposition und Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung (Donovan et al., fortl.) ii) Finanzielle Belastung durch Verletzungen für ältere Arbeitnehmer (Davie & Lilley 2018)

Kapitel 5: Unterschiede im Verlinkungsprozess

Kapitel 4 hat das Potential der Verlinkung von Datenbeständen im Gesundheitswesen grundsätzlich beleuchtet. Dabei kann es sowohl um eine Verlinkung bestehender Datensätze gehen, als auch um das Zuspielen neuer Ergebnisse aus der Primärforschung auf vorhandene administrative Daten (oder umgekehrt). Eine erfolgreiche Verlinkung setzt die Möglichkeit einer eindeutigen Zuordnung von Informationen aus separaten Datensätzen voraus. Dazu sind geeignete Schlüsselvariablen notwendig, die in allen zu verlinkenden Datensätzen vorhanden sind. Für eine personenbezogene Verlinkung ist es aus datenschutzrechtlichen Gründen häufig notwendig, statt eindeutiger personenidentifizierender Merkmale aus den Originaldaten, wie eine Versicherungs- oder Ausweisnummer, Pseudonyme einzusetzen; dabei muss für alle zu verlinkenden Datenquellen dasselbe Pseudonymisierungsverfahren angewendet werden. Die Zuständigkeiten für die Pseudonymisierung der Daten sowie der Prozess der De-Identifikation der Datenbestände können auch erheblich zur Bewahrung der Anonymität und zum Schutz der Privatsphäre beitragen. In Deutschland war es bisher datenschutztechnisch herausfordernd, einen breiten Einsatz der Verlinkung von gesundheitsbezogenen Daten zu etablieren. Nichtsdestotrotz zeigen erfolgreiche Fallbeispiele, dass die entsprechenden Möglichkeiten besser ausgenutzt werden können (vgl. March, Stallmann & Swart 2015). Mit der „Guten Praxis Datenlinkage“ wurden methodische Standards zur Verlinkung personenbezogener Daten im Rahmen von Forschungsvorhaben formuliert (March et al. 2019).

Die detaillierte technische Beschreibung möglicher De-Identifikations- und Verlinkungsoptionen geht über den Rahmen dieses Gutachtens hinaus. Nichtsdestotrotz erscheint es sinnvoll, beispielhaft vereinzelte Ansätze zu beleuchten, um erste Ideen für zukünftige Diskussionen zu erlauben. So gibt es in der Stichprobe der hier erhobenen Fallstudien ein Cluster, welches ein Verfahren basierend auf einer zugrundeliegenden, „verankernden“ Liste von Personen nutzt, um mittels identifizierender Merkmale diese Personen in den zu verlinkenden Datensätzen zu identifizieren und Daten somit zu verknüpfen. In der neuseeländischen IDI ist das der „IDI Spine“. Dieser stellt eine Liste von Individuen dar, die Informationen der Finanzämter mit Geburten- und Visum-Angaben kombiniert. Ziel des Spines ist es, alle möglichen Individuen zu identifizieren, die jemals in Neuseeland gewohnt haben (Zielpopulation), und ein einziges Mal zu erfassen (Black 2016; Statistics NZ 2014). Die Mikrodaten von den in der IDI abgebildeten Sektoren sind dann mit dem Spine verknüpft und können somit auch miteinander verlinkt werden, auch ohne ein eindeutiges identifizierendes Merkmal über die unterschiedlichen Sektoren hinweg. Ähnlich funktioniert das „Population Spine“ in Schottland, welches für die Verlinkungen über eDRIS herangezogen wird. Es bildet alle Personen in Schottland ab, die in Kontakt mit NHS Scotland waren, samt ihrer Identifikatoren (ISD Scotland 2019b). Abbildung 5.1 zeigt den Prozess in Schottland, der damit beginnt, nur die Identifikatoren in den Datensätzen eingeschlossener Personen mit dem Spine zu matchen.

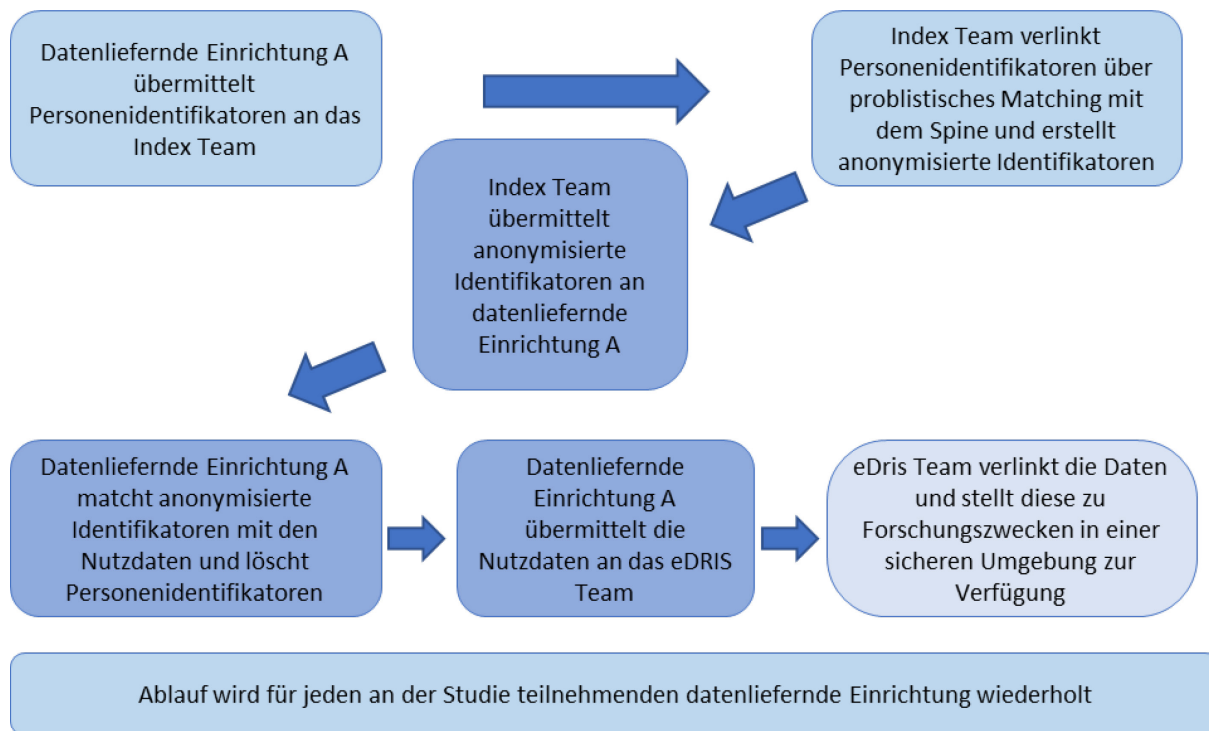


Abbildung 5.1 Verlinkungsansatz in Schottland (eDRIS)

Das gleiche Konzept einer Liste der Zielpopulation mit Identifikationen liegt auch dem Derived Record Depository (DRD) in Kanada zugrunde, welches für die Verlinkungen über das SDLE eingesetzt wird. Das DRD kann über seine Quelldaten aktualisiert werden. Neben dem DRD gibt es beim SDLE auch das „key registry“, wo dann die verknüpften IDs der verschiedenen Datensätze abgekoppelt gespeichert werden, so dass dann über das key registry später die entsprechenden Fälle in den ursprünglichen Datensätzen identifiziert werden können (Trudeau 2017). Datensätze, die denselben eindeutigen Identifikator („unique person identifier“) enthalten, können darüber verknüpft werden; dies kann zum Beispiel für Datenbestände aus dem Gesundheitsbereich erfolgen, die die „Health Insurance Number“ enthalten. Sonst werden deterministische und probabilistische Ansätze eingesetzt, um die Datensätze anhand von nicht-eindeutigen Identifikatoren zu matchen. Das Prinzip der Trennung zwischen den Daten mit den Informationen zur Verknüpfung und den ursprünglichen Daten, sowie auch ein „population repository“, werden auch bei PopData BC angewendet. Es entsteht ein Datensatz mit den Verknüpfungsinformationen, über diesen können dann Fälle aus anderen Datensätzen identifiziert werden.

Der Ansatz zur Verlinkung von Daten mit sensiblen Variablen für das Manitoba Population Research Data Repository wird hier hervorgehoben, da die Grundidee zur Diskussion über die Rollenverteilung bei der De-Identifizierung von Datenbeständen beitragen könnte. Die Datensätze kommen beim Repository ohne identifizierende Merkmale wie Namen, Adressen und Telefonnummern an. Diese schicken Datenhalter getrennt an Manitoba Health (das Landesgesundheitsministerium von Manitoba), während sie an Manitoba Centre for Health Policy (MCHP, dem Host des Repository) nur Programm- und Referenznummern schicken. Somit haben weder Manitoba Health noch MCHP Zugang zum vollständigen

digen Datensatz. Manitoba Health verlinkt unter Benutzung deterministischer und probabilistischer Methoden die identifizierenden Merkmale mit persönlichen Gesundheitsinformations-ziffern (personal health information number, PHIN); letztere werden dann verschlüsselt zusammen mit den Referenznummern als „crosswalk files“ an MCHP geschickt. Jede Person, die in Manitoba gesetzlich versichert ist (vgl. Kapitel 3), sowie jede Person, für die eine Leistung in Manitoba abgerechnet wurde, bekommt eine eigene PHIN. Sobald MCHP diese Informationen hat, kann es anhand der Referenznummern die Programmdateien mit den verschlüsselten PHINs verknüpfen. Das zugrundeliegende Prinzip ist, dass keiner der drei betroffenen Akteure alle Informationen hat: die Datenhalter haben keinen Zugriff auf die verschlüsselten PHINs, das Repository hat keinen Zugriff auf die identifizierenden Merkmale und Manitoba Health hat die Programmdateien nicht (vgl. Abbildung 5.2).

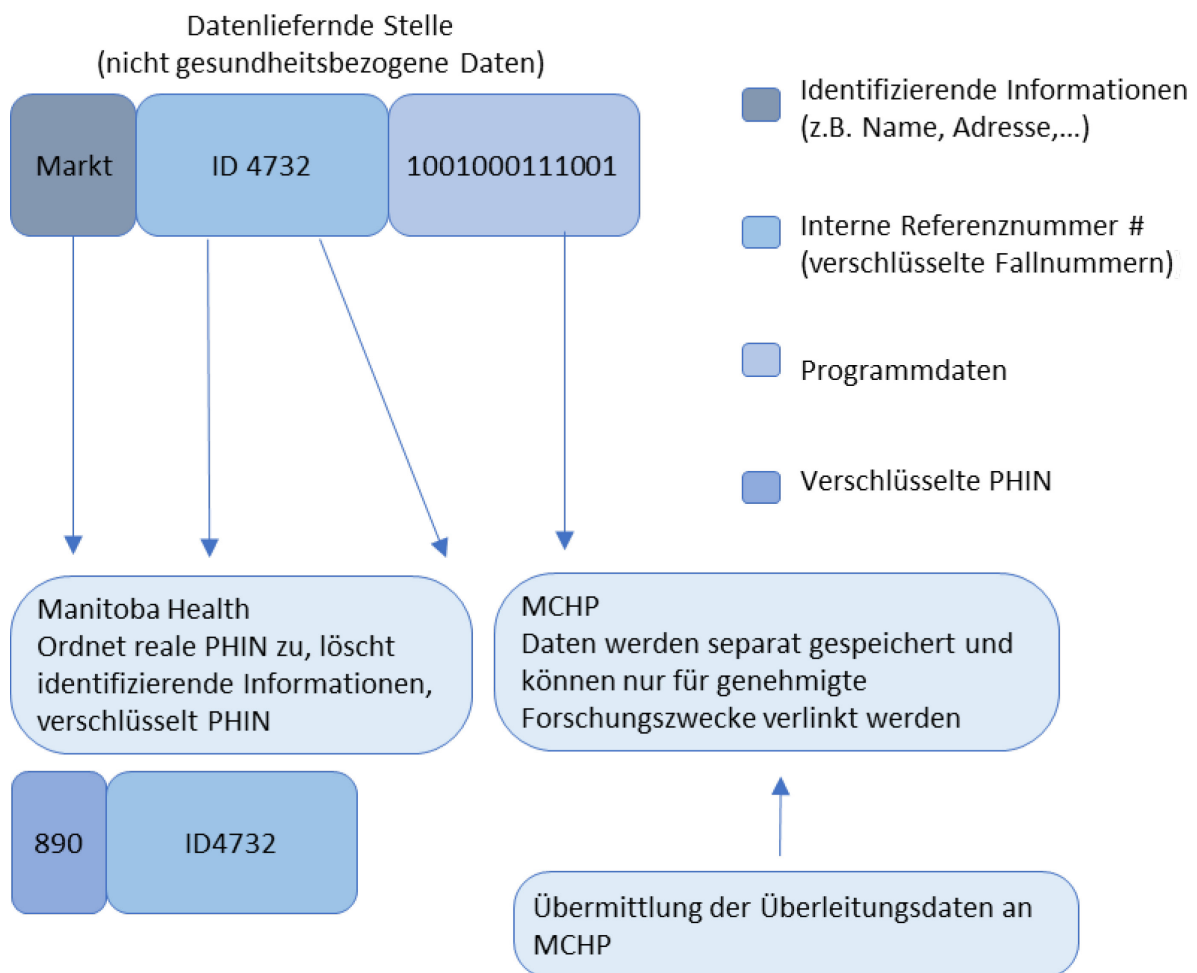


Abbildung 5.2 Verlinkungsansatz in Manitoba

Ein interessantes Beispiel liefert darüber hinaus die Verlinkung der Daten für den CPRD und der CPRD-Daten mit anderen Datenbeständen (vgl. Kapitel 3 und 4). Bevor Daten aus den Praxen in den CPRD fließen, werden identifizierende Merkmale (Namen, Geburtsdaten, Postleitzahlen und National Health Service Numbers) über die Praxissoftwareanbieter entfernt und mit pseudonymisierten Patienten- und Leistungserbringeridentifikatoren ersetzt. Für Praxen, die der Datenverlinkung zugestimmt

haben, schicken die Softwareanbieter die Identifikatoren an NHS Digital, welches die Validität der Daten überprüft und diese mit vorhandenen Daten aus der CPRD zusammenführt, um eine „Kohorte“ zu bilden. Datenhalter der zu verlinkenden Datensätze (vgl. Box 3.2) schicken identifizierende Merkmale ebenso an NHS Digital zusammen mit einem pseudonymisierten Patientenidentifikator, den „Link ID“. NHS Digital matcht die Identifikatoren der externen Datenhalter zu den CPRD-Identifikatoren und generiert eine „Linker“-Datei mit Paaren von pseudonymisierten Identifikatoren (Praxis und Patienten-ID, Link ID) für jeden Patienten (Padmanabhan 2019).

Die Verlinkungsinitiativen in den Ländern zeigen, dass auch ohne das Vorhandensein eines eindeutigen Identifikators eine Verknüpfung auf Basis anderer identifizierender Merkmale möglich ist und seit mehreren Jahren erfolgreich eingesetzt wird. Tabelle 5.1 hebt dies beispielhaft hervor. Darüber hinaus sind die Verteilungen der Zuständigkeiten variabel – nichtsdestotrotz kann man beobachten, dass der Einsatz von Vertrauensstellen unterschiedlicher Zugehörigkeiten als üblich gilt.

Tabelle 5.1 Übersicht über genutzte Verlinkungsmethoden nach Initiativen

Land	Initiativen	Typ und Vorgang der Verlinkung
Australien	PHRN	Probabilistische Verlinkung ohne eindeutigen Identifikator auf Basis anderer identifizierender Merkmale (Vorname, Nachname, Geburtsdatum, Adresse). PHRN unterstützt die entsprechenden lokalen (<i>state, territory</i>) Institutionen bei Datenverknüpfungen.
Kanada	SDLE	StatCan verlinkt auf sicherem Server; deterministisch oder probabilistisch basierend auf Vorhandensein von identifizierenden Informationen der zu verknüpfenden Datensätze.
	Manitoba Population Research Data Repository	Personal Health Number; De-Identifizierung und Verlinkung der Daten wird von derselben Einheit des Ministeriums durchgeführt.
Neuseeland	Integrated Data Infrastructure (IDI)	Innerhalb der Sektoren ggf. eindeutige Identifikatoren. Alternativ probabilistische Verknüpfung, basierend auf vorhandene Personendaten im Spine. Die Verlinkung wird von Stats NZ durchgeführt.
UK	Clinical Practice Research Data-link (CPRD)	NHS Number und andere Personen- und Praxisbezogene Daten; deterministisches Vorgehen
UK Schottland	electronic Data Research and Innovation Service (eDRIS)	Community Health Index und andere identifizierende Merkmale, probabilistisches matchen über Spine.

Kapitel 6: Zugriff auf gesundheitsrelevante Daten für Forschende

Neben den eingeschränkten Auswertungsmöglichkeiten der zur Verfügung stehenden Daten im deutschen Gesundheitswesen (vgl. Kapitel 1) wird auch der Datenzugang als herausfordernd bzw. nicht optimal kritisiert. Insbesondere für Projekte, die eine Verlinkung unterschiedlicher Datenbestände benötigen, sind die Berechtigungen intransparent und die Antragsverfahren häufig kompliziert und langwierig. Somit können die Projektdurchführung und somit ggf. auch die Relevanz von gewonnenen Einsichten erheblich erschwert werden. Aber auch die konkreten Möglichkeiten für einen Zugriff auf erfolgreich beantragten Daten schöpfen das Spektrum der zur Verfügung stehenden Mechanismen nicht aus.

Für alle relevanten in der Stichprobe enthaltenen Fallstudien ist eine Beantragung der Datennutzung bei der zuständigen Institution notwendig. Dies verläuft in vielen Fällen digital und wird für viele der geschilderten Fallstudien von einer zentralen Stelle unterstützt, auch wenn mehrere Datenhalter eingebunden werden müssen (z.B. PHRN, eDRIS). Zum Beispiel zählt der Vorgang der Antragstellung bei PopData BC zu den zugänglichsten und transparentesten in Kanada und verläuft in drei Schritten: der Antrag muss inklusive Ethik- und Reviewprozesse sowie Einwilligungsdokumente eingereicht werden; die zuständigen PopDataBC-Mitarbeiter helfen die Antragsteller, die Genehmigung der notwendigen Datenhalter zu bekommen; nach erfolgreichen Genehmigungen können Nutzungsverträge (inklusive Kosten) und Geheimhaltungsvereinbarungen unterschrieben werden. Schlussendlich müssen erfolgreiche Antragsteller ein Datenschutztraining absolvieren und ihre Arbeitsumgebung innerhalb der gesicherten Forschungsumgebung von PopData BC einrichten. Für alle Fallstudien, die einen Zugang für Forscher vorsehen, waren die Beantragungsregelungen entweder explizit auf der Webseite beschrieben oder es wurde auf die entsprechende Stelle verwiesen.

Da Daten im Gesundheitssystem grundsätzlich sensibel sind, kommen in der Regel drei Zugangswege in Frage: (1) eine direkte Datenübermittlung, die ein angemessenes Datenschutz- und Datensicherungskonzept voraussetzt; (2) die Inhouse-Auswertung, die an geeigneten Gastarbeitsplätzen der datenhaltenden Institution stattfindet; und (3) die Fernabfrage, die auf separaten Servern mit „remote access“ über VPN-Verbindung möglich ist (vgl. Ihle et al. 2015). Wie Tabelle 6.1 zu entnehmen ist, kommen alle drei Zugangswege in der Stichprobe der eingeschlossenen Fallstudien vor. Der bisher eingesetzte Ansatz für die Bearbeitung der DaTraV-Daten in Deutschland (vgl. Kapitel 1) wurde so in keinem anderen Land identifiziert. Hingegen scheinen sowohl die Nutzung von Gastarbeitsplätzen als auch die Fernabfrage über VPN zunehmend Anwendung zu finden.

Ein interessantes Beispiel zur Vergegenwärtigung der Beantragungs- und Nutzungsprozesse bietet das Secure Unified Research Environment (SURE), welches für die Nutzung der PHRN-verlinkten Daten in Australien zum Einsatz kommt (Abbildung 6.1). Es wird vom Sax Institute in Sydney gemanagt und stellt eine Datenverarbeitungsumgebung dar. Hier können Datenverwalter und Forscher Daten hochladen und Forscher Daten analysieren. Es wird in der Regel zur Bearbeitung von de-identifizierten

Mikrodaten und verlinkten Daten benutzt. Während der Bearbeitung bleiben die Daten auf den Sax-Servern in Sydney, nach erfolgreichem Abschluss der Analyse können aber aggregierte Daten auf lokale Computer heruntergeladen werden und gehören den Forschern. Autorisierte Nutzer von SURE können überall in der Welt arbeiten und die Umgebung mittels sicheren Einloggens betreten.

Tabelle 6.1 Prozesse für den Datenzugriff für ausgewählte Fallstudien

Fallstudie	Wer kann die Daten beantragen?	Wo/von wem müssen die Daten beantragt werden?	Wer entscheidet über Anträge?	Wie lang dauert der Bewerbungsprozess?	Wie bzw. Wo können Forscher mit den Daten arbeiten?	Wem gehören die Daten während der Analyse?
PHRN (Australien)	Forscher, deren Projekte inhaltliche und ethische Vorschriften erfüllen	PHRN Online Application System	Data linkage unit	Variiert entspr. der zu verlinkenden Datenbestände	SURE, remote-access data research laboratory	Data bleibt auf den Servern in Sydney
PopData BC (Kanada)	Kanadische Forscher die an einer Forschungsinstitution arbeiten	PopData's Data Access Unit (DAU)	Datenhalter	Variiert entsprechend der Rückmeldungen von Datenhaltern	Secure Research Environment (SRE), kann über einen Firewall mittels verschlüsseltem VPN benutzt werden	Datenhalter
Nivel PC Database (Niederlande)	Forscher, die bestimmte Vorschriften erfüllen (inkl. Verpflichtung zur Veröffentlichung der Ergebnisse)	Nivel über ein Formular (auf Webseite verfügbar)	Nivel auf Basis der Verfahrensordnung (Spitzenverbände der Leistungserbringer und Datenschutzkomitee)	Durchschn. 3 Wochen, aber abhängig von Komplexität und Datenschutzbesonderheiten	Je nach Antrag, entweder Fernzugang-Einrichtung („remote access facility“) oder Datensatz direkt erhältlich	Nivel
IDI (Neuseeland)	Analysten und Forscher im öffentlichen Sektor, Universitäten und NGOs	Statistics NZ (online)	Statistics NZ, + ggf. andere Behörden	Reviewprozess: 6 Wochen	Statistics NZ secure data lab environment (bei den Niederlassungen von StatsNZ; eigene data labs können beantragt werden)	Statistics NZ
HIRA (Südkorea)	Wissenschaftler, öffentlicher Sektor, Privatsektor (Umfang der Daten variiert)	HIRA	HIRA	2-4 Wochen nach Antrag bis Daten vorhanden sind	Sieben Forschungszentren oder Fernabfrage (unter Auflagen)	HIRA

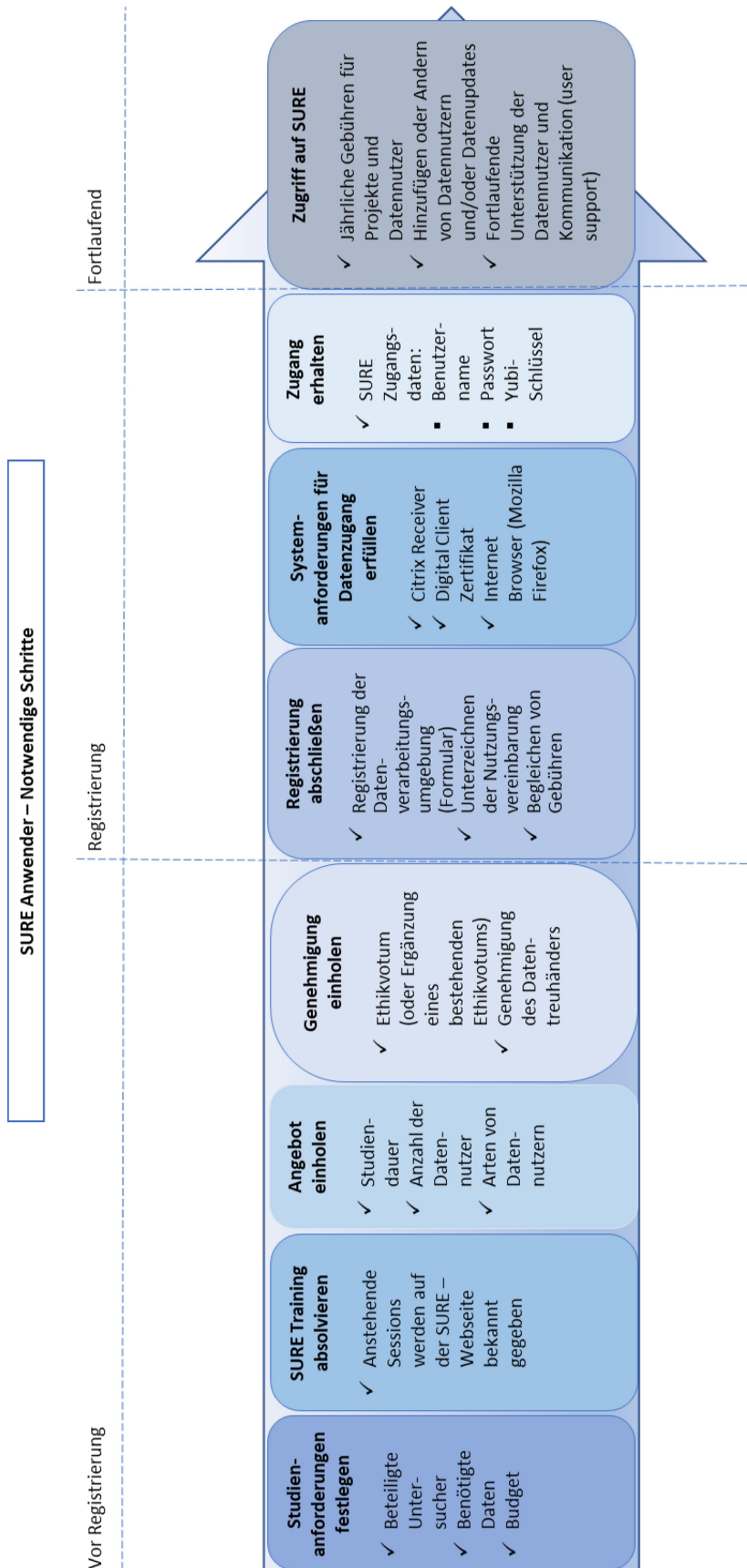


Abbildung 6.1 Beantragung und Nutzung von Daten über SURE (Sax Institute 2019)

Kapitel 7: Überlegungen zu elektronischen Patientenakten

Daten aus elektronischen Patienten- bzw.- Gesundheitsakten bieten eine weitere Quelle für Analysen für die Versorgungsforschung. Einleitend müssen die Begrifflichkeiten jedoch auseinandergezogen werden. Die US-Amerikanische Behörde für Gesundheits-IT differenziert zwischen elektronischen Krankenakten („electronic medical record“), elektronischen Gesundheitsakten („electronic health record“) und elektronischen Patientenakten („personal health record“; vgl. Box 7.1 und zum Inhalt von elektronischen Gesundheitsakten auch Häyrinen et al. 2008). Während nach dieser Aufteilung die elektronische Krankenakte der Patientenakte im elektronischen Praxisverwaltungssystem einzelnen Leistungserbringer entsprechen würde, unterscheidet man hierzulande zwischen den als Satzungsleistung nach § f 68 SGB V von einzelnen Krankenkassen angebotenen elektronischen Gesundheitsakten, welche unterschiedliche Funktionalitäten und Möglichkeiten der ärztlichen Einbindung umfassen und der in § 291a SGB V geregelte elektronische Patientenakte, die einheitliche Möglichkeiten für alle Versicherte anbieten soll und das zentrale Element der vernetzten Gesundheitsversorgung und der Telematikinfrastruktur darstellt (KBV 2019).

Box 7.1 Begriffsbestimmungen nach der US-Amerikanischen Behörde für Gesundheits-IT

Elektronische Krankenakte („electronic medical record“): vereinfacht eine digitale Version der traditionell papierbasierten Krankenakte; primärer Zweck: Diagnose und Behandlung bei Klinikern in derselben Organisation (Praxis, Krankenhaus). Nicht für den Austausch mit weiteren Leistungserbringern gedacht.

Elektronische Gesundheitsakte („electronic health record“): breiterer Einsatz, enthält Informationen von allen Leistungserbringern, die in der Versorgung der Patienten beteiligt sind und alle diese Leistungserbringer dürfen auf die darin enthaltenen Informationen zugreifen. Elektronische Gesundheitsakten „folgen den Patienten“.

Elektronische Patientenakte („personal health record“): sie enthalten die gleichen Informationen wie elektronische Gesundheitsakten, sind aber so aufgebaut, dass sie von Patienten abgerufen und gemanagt werden können. Ziel ist, dass Patienten ihre Informationen in einer privaten, gesicherten und vertraulichen Umgebung aufbewahren und verwalten können.

Quelle: Office of the National Coordinator for Health Information Technology 2019

Tatsächlich sind elektronische Patientenakten, bzw. Anwendungen, die es Patienten und Leistungserbringern erlauben, alle gesundheitsbezogenen Informationen zusammenhängend aufzurufen, in vielen Ländern ein zentrales Element der Umsetzung von e-Health. Dänemark ist in der Länderstichprobe hierfür das beste Beispiel. Über das zugrunde liegende Netzwerk für Gesundheitsdaten, welches Gesundheits- und Gesundheitsversorgungsinformationen von allen Leistungserbringern enthält (vgl. Abbildung 3.1), können Patienten ihre Daten auf Sundhed.dk aufrufen. Schweden hat ein im Prinzip ähnliches, aber technisch abweichendes System mit einer im Vergleich eingeschränkteren Umsetzung (vgl. Kapitel 3, Davoody et al. 2019, Hägglund et al. 2018). Die primären Zwecke solcher Systeme liegen in der Verbesserung der Koordination der Versorgung und in einer erhöhten Transparenz für die Patienten, nicht zuletzt, um deren Empowerment zu fördern. Regulären Zugriff auf die Daten haben dabei die Patienten selbst sowie die Leistungserbringer. Ein Zugriff für Forscher würde eine zusätzli-

che bzw. sekundäre Funktionalität darstellen. Zum Beispiel bietet das schwedische Health Information Exchange theoretisch auch Anwendungen für Forscher, ihre Umsetzung ist aber unklar (Sellberg & Eltes 2017).

In den letzten Jahren haben einzelne Länder in der Stichprobe Initiativen eingeführt, die elektronische Patientenakten anbieten und somit potenziell auch neue Datensätze für die Forschung schaffen könnten. Dabei gibt es in Australien sowie in Österreich die Möglichkeit für Patienten, entweder ganz aus der Nutzung der elektronischen Akte auszusteigen oder aber bestimmte Eintragungen nicht vornehmen zu lassen bzw. diese nicht in der Akte freizugeben (Vgl. Tabelle 7.1). In Australien werden darüber hinaus die Patienten informiert, wenn ein Leistungserbringer auf ihre Akte zugreift. In Australien können die Daten der „MyHR“ für Forschungszwecke genutzt werden, vorausgesetzt die Patienten haben einer Nutzung ihrer Daten für die Forschung nicht widersprochen. Das österreichische Modell sah zwar eine ähnliche Nutzung für die Forschung vor, aber der politische Wille für ihre Umsetzung soll ausgeblieben sein. Die aktuellen Planungen für die deutsche elektronische Patientenakte (ePA; §291a SGBV) weichen von den identifizierten Modellen in den anderen Ländern ab. Aktuell ist ein doppeltes „opting-in“-Verfahren geplant, d.h. die Versicherten erhalten über ihre Krankenkasse die Möglichkeit an der ePA teilzunehmen und melden sich für diese an. Darüber hinaus können die Daten für die Forschung bereitgestellt werden, auch hierfür melden sich die Versicherten an. International wurden keine Erfahrungen mit einer vergleichbaren Nutzung der Daten aus elektronischen Patientenakten für die Versorgungsforschung identifiziert; entsprechende Überlegungen und Vorschläge sind daher eher tentativ. Es leuchtet aber ein, dass je nachdem wie viele und welche Patienten sich für oder gegen die Nutzung von solchen Anwendungen entscheiden und gemäß den genauen Gestaltungsmöglichkeiten der Zugriffsrechte, die Vollständigkeit und Repräsentativität der Daten maßgeblich eingeschränkt sein könnten (vgl. Kapitel 8).

Tabelle 7.1 Zugriffsmöglichkeiten und Patientenrechten bei elektronischen Patientenakten

Land	Initiative	Zugriff	Opt-out / Bestimmung der Datensichtbarkeit	Zusätzliche Möglichkeiten für die Forschung
Australien	MyHealthRecord (MyHR)	Patienten und Leistungserbringer haben regulär Zugriff auf die Daten	Patienten: opt-out of MyHR, ausgewählte Leistungserbringer einschränken, bestimmte Unterlagen verbergen; Benachrichtigung, wenn Leistungserbringer auf myHR zugreifen. Opt-out aus der Nutzung für Forschung. Leistungserbringer: Opt-out	Zugriff auf die Daten zukünftig möglich gegen Gebühr, ausschließlich der Daten von Bürgern die nicht MyHR vollständig oder für die Forschung widersprochen haben.
Dänemark	Sundhed.dk	Patienten und Leistungserbringer haben regulär Zugriff auf die Daten	Nein	Nein
Österreich	ELGA (Elektronische Gesundheitsakte)	Patienten und Leistungserbringer (Umsetzungsphase), letztere mit Zeiteinschränkung	Patienten: Ganz und partiell (letzteres für einzelne Anwendungen, z.B. eBefunde)	Reguliert, aber nicht umgesetzt
Schweden	Journalen (Anwendung des PAEHR – “patient-accessible electronic health record”)	Patienten und Leistungserbringer	Implementierung bei Region. Für alle Patienten, Leistungserbringer können Beitrag einschränken	Nein

Kapitel 8: Zwischenfazit zu den Ergebnissen und Limitationen des Gutachtens

Die in den bisherigen Kapiteln thematisierten Fallstudien aus dem internationalen Vergleich heben wichtige Aspekte hervor, die für die Problematik in Deutschland von Bedeutung sein könnten. Eine Nebeneinanderstellung der in Deutschland vorhandenen Datenbestände (Tabelle 1.2) und der in anderen Ländern eingesetzten Daten (hauptsächlich in Kapiteln 3 und 4 dargestellt) deutet darauf hin, dass sich die Diskussion hierzulande an erster Stelle mit der Ausgestaltung der zur Verfügung stehenden Daten befassen sollte und ergänzend die Erfassung zusätzlicher Informationen bedenken kann. Konkrete Empfehlungen werden in Kapitel 9 entwickelt.

Neben dem Mehrwert von vergleichsmäßig einfach zugänglichen, umfangreiche(re)n Abrechnungsdaten, die schneller als in Deutschland vorliegen, sektorenübergreifende Analysen erlauben und personenbezogen mit Gesundheits- sowie auch mit Regionaldaten verknüpfbar sind (z.B. MBS/PBS in Australien, NHIS/HIRA in Südkorea, CMS Data in den USA), beleuchtet Kapitel 4 auch das Potential von Datenbanken, die auf Basis der routinemäßigen Dokumentation (i.d.R. elektronischen Gesundheitsakten) von Leistungserbringern entstehen (z.B. CPRD, Nivel PC Database). Eine Investition in diese Richtung, wie sie zum Beispiel in Großbritannien mit CPRD erfolgte (Gothe 2015; Gentil et al. 2017), kann nicht nur für das eigene Gesundheitssystem Einsichten ermöglichen, sondern auch die internationale Kooperation fördern und zur internationalen Sichtbarkeit wissenschaftlicher Exzellenz beitragen (vgl. Vezyridis & Timmons 2016); teilnehmende Praxen können von periodischen Berichten sowie eine Möglichkeit zum Benchmarking profitieren. Allerdings muss unbedingt beachtet werden, dass eine Reihe von Verzerrungsquellen die Forschung auf Basis von Daten aus elektronischen Gesundheitsakten erschweren können. Dies betrifft unter anderem die Dokumentation und Kodierung in der elektronischen Akte, die Extraktion und Zusammenführung der Daten sowie die Aufbereitung für die wissenschaftliche Praxis (Agniel, Kohane & Weber 2018; Verheij et al. 2018). Die Repräsentativität der Stichprobe bei Netzwerken teilnehmender Praxen (wie beim CPRD und der Nivel Database) muss beachtet werden, sowie die Vollständigkeit der Daten; beides wird davon beeinflusst, welche Möglichkeiten Patienten zur Verfügung stehen, ihre Daten aus dem Forschungszugriff auszuschließen (vgl. auch Kapitel 7). Zum Beispiel wurde im NHS zusammen mit der Umsetzung der DSGVO im Jahr 2018 das sogenannte „National Data Opt Out“ eingeführt, ein Dienst, der es Patienten ermöglicht, den Zugriff auf ihre vertraulichen Daten zu regeln; alle Institutionen im Gesundheitssystem müssen bis März 2020 mit der Regelung konform sein.

In der in Deutschland vorgesehenen, aber noch nicht abschließend vorliegenden Regelung zur elektronischen Patientenakte sollen Patienten die Datenhoheit behalten und sich freiwillig sowohl für die Akte an sich als auch für die Nutzung ihrer Daten für die Forschung entscheiden (doppeltes „opt-in“). Somit würden Daten, die über die ePA generiert bzw. spendiert wurden, ähnlichen Limitationen unterliegen.

Die Vielfältigkeit der Datenbestände zu den unterschiedlichen Kategorien im Analysekonzept des Gutachtens (vgl. Abbildung 2.1) ist kein Alleinstellungsmerkmal Deutschlands. In diesem Gutachten wird eine Reihe von Initiativen dargestellt, die genau aufgrund eines Bedarfs an verlinkten Daten den Zugang für die Forschung vereinfachen sollen. Diese umfassen sowohl integrierte Modelle (wie die IDI in Neuseeland) als auch zentrale Zugangsstellen für die Beantragung und Verlinkung verteilter Datenbestände unterschiedlicher Datenhalter (z.B. PHRN in Australien, SDLE in Kanada, eDRIS in Schottland; in allen Ländern sind parallel auch regionale Initiativen aktiv). Unter der Rubrik „Datenangebot“ bieten die Forschungsdatenzentren der statistischen Ämter der Bundesländer in Deutschland auf ihrer Webseite eine Übersicht vorhandener Daten¹; hierauf ließe sich in diese Richtung aufbauen. Zusätzlich zu den schon genannten Beispielen könnten Plattformen wie das Forschungszentrum (ResDAC) und der Data Navigator des CMS in den USA herangezogen werden. Aus den strategischen Bemühungen in den Vergleichsländern (vgl. Profile im Anhang und Kapitel 3) wird deutlich, dass ein starkes politisches Engagement für eine erfolgreiche Umsetzung neuer Wege für die Versorgungsforschung ausschlaggebend ist (vgl. auch Tew et al. 2017).

Das Konzept des „one-stop-shop“ für Daten, also einer Stelle, die alle bürokratischen Schritte koordiniert bzw. durchführt und somit die Prozesse für Beantragende vereinfacht, dient nicht nur der Entlastung der Forscher bzw. Politikanalysten, sondern auch einer besseren und effizienteren Nutzung aller Daten und somit der Formulierung und Beantwortung weitreichender Fragestellungen. In diese Richtung gibt es auch neuere Bemühungen, wie etwa in Finnland mit Findata. Auf europäischer Ebene steht die Möglichkeit eines European Health Data Space vor, welches die Möglichkeiten bestehender und neuer Daten auch grenzüberschreitend einsetzen soll. Erste Initiativen für den grenzüberschreitenden Austausch gesundheitsbezogener Daten sind bereits unterwegs (Paun 2019). Die deutsche Präsidentschaft des Europäischen Rates in der zweiten Hälfte 2020 soll das Konzept aktiv fördern (von der Leyen & Spahn 2019).

Das Gutachten hat sich auf die Vorstellung und Diskussion selektierter Fallstudien beschränkt; somit ist davon auszugehen, dass auch andere Länder interessante Impulse für Deutschland anbieten können, oder sogar weitere Initiativen in den eingeschlossenen Ländern. Darüber hinaus war eine detaillierte Analyse von datenschutzrechtlichen Regelungen nicht Teil der Zielsetzung des Gutachtens; wichtige Regelungen wurden pro Land erfasst, um eine Grundlage für weiterführende Recherchen bereitzustellen. Das Gutachten hat nicht ein Benchmarking der eingeschlossenen Fallstudien beabsichtigt, weder im Sinne der Qualität, Vollständigkeit und Repräsentativität der Daten noch hinsichtlich ihrer a priori Übertragbarkeit für Deutschland. Vielmehr wurde es als zielführend erachtet, hilfreiche Aspekte zu beleuchten und ihre Umsetzung über die Fallstudien hinweg zu diskutieren. Eine Limitation des Gutachtens liegt in dem Umfang der Erfahrungsberichte mit den eingeschlossenen Fallstudien; es ist nicht auszuschließen, dass über die öffentlich zugänglichen Quellen bzw. über die Konsultation der Landesexperten nur selektiv Informationen zu prozeduralen Aspekten oder zu Datensicherheitsabwägungen

¹ <https://www.forschungsdatenzentrum.de/de#datenangebot>

identifiziert wurden. Die Nutzung und Auswertung von unstrukturierten Daten sind unter den eingeschlossenen Initiativen nicht häufig vertreten, spielen aber in den Diskussionen um Big Data eine wichtige Rolle. Auch die explizite Beschäftigung mit Umweltdaten konnte nicht ausführlich beleuchtet werden, obwohl diese in vereinzelt Projekten verlinkt wurden (vgl. Tabellen 4.1 und 4.3). Hierzu sind weiterführende Recherchen erforderlich.

Kapitel 9: Umsetzungsmöglichkeiten für Deutschland

Obwohl in Deutschland große Datenbestände vorliegen, welche relevante Informationen für die Versorgungsforschung enthalten (vgl. Tabelle 1.2), sind die praktischen Möglichkeiten zur Nutzung dieser Daten als „eingeschränkt“ zu bezeichnen. Idealerweise ist für die Versorgungsforschung ein Zugriff auf Daten möglich, welche die nachfolgenden Kriterien erfüllen. Die Daten sollen:

- auf Individualebene vorliegen, d.h. nicht nur auf aggregierter Ebene („Mikrodaten“)
- möglichst alle relevanten Versorgungssektoren umfassen, und innerhalb der Sektoren sowohl die Leistungen als auch die Leistungserbringer; somit wäre auch eine Aggregation auf entsprechender Ebene möglich
- über längere Zeiträume vorliegen, und über die Jahre verknüpfbar sein, so dass die jeweils aktuellsten Daten zugespielt werden können
- die gesamte Bevölkerung abdecken
- auch Informationen zu nicht-medizinischen, d.h. etwa soziodemographischen/ sozioökonomischen Parametern sowie regionale Informationen (z.B. PLZ oder Wohnort) enthalten
- auch auf aggregierter (etwa regionaler) Ebene verknüpfbar sein (etwa mit Informationen zur Lokalisation oder Dichte von Leistungserbringern, regionalen/ KV-Daten zur durchschnittlichen Inanspruchnahme oder Leistungserbringung, aber auch mit anderen Datenbeständen, z.B. hinsichtlich regionaler Umweltfaktoren)

Des Weiteren sollten für die wissenschaftliche Datennutzung bestimmte Kriterien erfüllt werden. Insbesondere sollen:

- transparente Regelungen hinsichtlich der Nutzungsberechtigungen vorliegen
- gleiche Bedingungen für alle wissenschaftlichen Institutionen für einen Datenzugriff gelten
- eine möglichst geringe zeitliche Differenz zwischen Datenerhebung und möglichem Datenzugriff bestehen
- eine zeitnahe Bearbeitung des beantragten Datenzugriffs erfolgen
- Möglichkeiten für einen direkten Zugriff auf die Mikrodaten – unter den entsprechenden datenschutzrechtlichen Bestimmungen – vorhanden sein (in Ergänzung zur Datenfernverarbeitung)

Die in Deutschland vorliegenden und für die Versorgungsforschung nutzbaren Datenbestände sind aktuell weit davon entfernt, diese Kriterien zu erfüllen. Die verfügbaren Daten sind fragmentiert, teilweise unvollständig und erlauben nur selten sektorenübergreifende Analysen oder eine Verknüpfung mit anderen Daten. Darüber hinaus stellen die (teilweise nicht vorhandenen bzw. intransparenten) Zugriffsregelungen die VersorgungsforscherInnen vor weitere Hürden und verursachen einen immensen Arbeitsaufwand. Basierend auf internationalen Erfahrungen, wie sie in Kapiteln 3 bis 9 und im Anhang dieses Gutachtens dargestellt werden, wurden daher vier (sich z.T. ergänzende bzw. aufeinander aufbauende) Szenarien entwickelt, welche einzeln bzw. in Kombination, die Datenverfügbarkeit

für die Versorgungsforschung in Deutschland verbessern könnten. Der grundsätzliche Anspruch an höchste Datenschutzstandards und an die Bewahrung der Anonymität von Personen wird durchgehend als unabdingbare Voraussetzung betrachtet.

a) Ausweitung des bestehenden verknüpften GKV-Datenkranzes (DaTraV Daten) und Verbesserung des Datenzugriffs

Der umfangreichste, sektorenübergreifende und im Längsschnitt verknüpfbare Datensatz in Deutschland liegt den gesetzlichen Krankenkassen (jeweils für ihre eigenen Versicherten) vor. Mit der Neuregelung der §§ 303a-e SGB V im Jahr 2012 wurde ein Datenbestand geschaffen, in dem die Daten aller gesetzlichen Krankenkassen zusammengeführt sind und dessen Nutzung der Aufgabenerfüllung verschiedener Institutionen des Gesundheitswesens sowie der wissenschaftlichen Forschung dienen soll. Mit der Aufbereitung und Bereitstellung der Daten ist das Deutsche Institut für Medizinische Dokumentation und Information (DIMDI) beauftragt. Der unter der Bezeichnung „Informationssystem Versorgungsdaten“ bereitgestellte Datenbestand beruht auf den jährlichen Datenübermittlungen der gesetzlichen Krankenkassen an das Bundesversicherungsamt zum Zweck der Weiterentwicklung des morbiditätsorientierten Risikostrukturausgleichs (Morbi-RSA).

Da bislang jedoch nur die für die Berechnung des Morbi-RSA relevanten Informationen von den Krankenkassen, bzw. vom Bundesversicherungsamt an das DIMDI übermittelt werden, ist dieser Datensatz für die Versorgungsforschung nur eingeschränkt nutzbar. So sind u.a. derzeit die Daten verstorbener Personen im Sterbejahr nicht im Datensatz enthalten, ebenso fehlen Operations- und Prozedurencodes aus der stationären und ambulanten Versorgung. Auf Grund dieser fehlenden Informationen kann beispielsweise nicht analysiert werden, wie sich nach 30, 90 oder 365 Tagen die Mortalität/Morbidität nach einem Schlaganfall zwischen Patienten mit Behandlung in einer Stroke-Unit und Patienten ohne Behandlung in einer Stroke-Unit (abbildbar über den OPS-Code) unterscheidet.

Aktuell kann diese Frage nur mit anderen Datensätzen analysiert werden, die jeweils ebenfalls spezifische Einschränkungen aufweisen: entweder nur für den Zeitraum der Krankenhausbehandlung selbst (Krankenhausfalldaten gemäß §21 KHEntgG), in Kooperation mit einzelnen Krankenkassen nur für die entsprechende Versichertenpopulation oder anhand der Daten der Schlaganfall-Register, die aufgrund der freiwilligen Teilnahme der Kliniken i.d.R. keine vollständige Erhebung darstellen.

Bei einer Ausweitung des DaTraV Datensatzes (auch rückwirkend für die bereits vorliegenden Jahre) um alle den Krankenkassen vorliegenden Informationen könnten zahlreiche wichtige Forschungsfragen vollständig analysiert werden (vgl. HIRA in Südkorea). Zur angemessenen Evaluation von Innovationen in der Gesundheitsversorgung (z.B. gesetzliche Neuregelungen, Einführung von Medikationsplänen) sind longitudinale Studien unabdingbar; darüber hinaus sind die Effekte von vielen Interventionen, wie z.B. Stroke Units, unter Umständen erst über einen längeren Zeitraum erkennbar. Es ist also essenziell, dass die Daten auch als Panel analysiert werden können. Wünschenswert wäre, dass die Krankenkassen einen definierten Auszug ihrer Daten gemäß § 284 SGB V direkt an die Datenaufbereitungsstelle übermitteln. Dies würde auch dazu beitragen, die große Zeitverzögerung, mit der die

Daten derzeit zur Verfügung stehen, zu verringern, wie es in anderen Ländern (wie zum Beispiel in Kanada, Neuseeland und den USA) bereits erfolgt ist. Eine weitere Verringerung der Zeitverzögerung ließe sich mit unterjährigen (beispielsweise quartalsweisen) Datenlieferungen erreichen.

Neben den Einschränkungen des Datensatzes erschweren derzeit auch lange Bearbeitungszeiten für Nutzungsanträge und prozessuale Hürden bei der Durchführung von Analysen die Datennutzung. Aufgrund der Erfahrungen und bisherigen Ergebnisse wäre es wünschenswert, die in den §§ 303b und 303d SGB V genannten Aufgaben des Forschungsdatenzentrums einer Stelle zu übertragen, die mit der Bereitstellung von Forschungsdaten Erfahrung hat und über die erforderlichen technischen und personellen Ausstattungen sowie Prozesse verfügt. Eine Akkreditierung des Forschungsdatenzentrums durch den Rat für Sozial- und Wirtschaftsdaten (RatSWD) wäre vorteilhaft. Als (bereits akkreditiertes) Forschungsdatenzentrum wäre z.B. das Statistische Bundesamt geeignet, dessen Forschungsdatenzentrum zahlreiche Arten von Forschungsdaten für wissenschaftliche Auswertungen nutzerfreundlich auf verschiedenen Wegen zugänglich macht. Daneben könnten die DaTraV Daten dann auch zur Erstellung von Bundesstatistiken verwendet werden. Ähnliche Modelle sind international schon umgesetzt (zum Beispiel in Kanada, Neuseeland und Schottland).

Zur Auswertung der DaTraV Daten ist zudem eine Ausweitung der Zugriffsmöglichkeiten über die aktuell bestehende Datenfernverarbeitung wünschenswert. Dies betrifft u.a. die Schaffung von Gastwissenschaftlerarbeitsplätzen (wie an den „safe havens“ des Britischen NHS, in Neuseeland oder auch in Südkorea) sowie idealerweise einen direkten Datenzugriff mittels sicherer Verbindung (bspw. über einen Terminalserver, wie das SURE-System in Australien oder der VPN-Zugang zum kanadischen *Population Data BC*). Technische und datenschutzrechtliche Möglichkeiten, wie die Daten der gesetzlichen Krankenkassen leichter für die Forschung zugänglich gemacht werden können, haben dabei schon einzelne Krankenkassen ausgearbeitet und umgesetzt (ein Beispiel bietet das *Wissenschafts-Data-Warehouse* der BARMER).

Das Digitale-Versorgung-Gesetz (DVG, BGBl. vom 9.12.2019) soll die Nutzbarkeit der GKV-Routinedaten für Forschungszwecke auf der Grundlage der §§ 303a-e verbessern. Es sieht die Schaffung einer Datenbasis vor, welche auf der Individualebene Angaben zu Alter, Geschlecht und Wohnort, Angaben zum Versicherungsverhältnis, Kosten- und Leistungsdaten nach den §§ 295, 295a, 300, 301, 301a und 302 SGB V, Angaben zum Vitalstatus und zum Sterbedatum sowie Angaben zu den abrechnenden Leistungserbringern enthalten soll. Zur Wahrnehmung der Aufgaben des Forschungsdatenzentrums soll eine öffentliche Stelle des Bundes per Rechtsverordnung bestimmt werden. Auch Art und Umfang der Daten, sowie die konkreten Zugriffsregelungen werden per Rechtsverordnung neu bestimmt. Da diese noch nicht vorliegt, können diese Aspekte noch nicht beurteilt werden. Wünschenswert wäre, dass im Vorfeld der neuen Rechtsverordnung die zukünftigen Datennutzer (insbesondere aus dem Bereich der Versorgungsforschung) mit einbezogen oder zumindest angehört werden. Um die Möglichkeiten einer Verlinkung zwischen Mikrodaten und Regionaldaten entfalten zu können, wäre bezüglich der Angaben zum Wohnort der Einschluss von Postleitzahlen notwendig.

Auch wenn die Neuerungen des Digitale-Versorgung-Gesetzes auf Verbesserungen des Datenangebots und der Zugangsmöglichkeiten hoffen lassen, muss jedoch beachtet werden, dass sich die DaTraV Daten nur auf Versicherte und Leistungsbereiche der gesetzlichen Krankenversicherung beziehen und damit Teile der Bevölkerung (Privatversicherte) sowie bestimmte Leistungsbereiche der Gesundheitsversorgung (z.B. Rehabilitation zu Lasten der gesetzlichen Rentenversicherung) nicht erfasst werden.

b) Koordinierende Instanz, die eine Verlinkung bestehender Daten ermöglicht (Clearingstelle Gesundheitsdaten?)

Eine darüber hinaus gehende Möglichkeit, die Datenbasis für die Versorgungsforschung zu verbessern, ist die Verknüpfung bereits bestehender Datenbestände (innerhalb der GKV und darüber hinaus). Aktuell stellt die Verknüpfung bestehender Datenbestände Forscher vor zwei wesentliche Herausforderungen: i) die variierenden Zugangsregelungen zu den verschiedenen Daten stellen eine Hürde an sich dar und ii) viele der Datensätze liegen pseudonymisiert und ohne eindeutigen Verknüpfungsschlüssel vor, so dass Forscher vor der Herausforderung des Datenlinkage stehen. Der Zugriff auf die Daten sowie die Durchführung eines Linkage gehen bei jedem Forschungsprojekt mit umfangreichen Anträgen zur Datennutzung, Datenschutzkonzepten sowie ggf. notwendigen Ethikanträgen – und somit einem hohen Arbeitsaufwand – einher.

Mittels einer Koordinierungs- und Datenverlinkungsstelle könnten diese Probleme sowie der Aufwand für Zugriff und Verlinkung erheblich minimiert werden. International liegen viele positive Beispiele für die Nutzung von Koordinierungs- und Datenverlinkungsstellen vor (u.a. die „*electronic data research infrastructure*“, das „*social data linkage environment*“ (SDLE) in Kanada, *Population Data BC* in Kanada sowie das *Population Health Research Network* in Australien). Während sich der Umfang der zur Verfügung stehenden Daten sowie das genaue Prozedere sowohl der Datenverknüpfung als auch des Datenzugriffs unterscheiden, sind die Leitideen der verschiedenen Initiativen identisch:

- Ziel: eine verbesserte Datenverfügbarkeit für die Wissenschaft, um so eine evidenzbasierte Gesundheitspolitik zu fördern und die Gesundheit der Bevölkerung zu verbessern
- Verwendung von bereits bestehenden Datenbeständen
- Ein Zugriffspunkt für Wissenschaftler, über den Informationen zu den vorhandenen Daten sowie den Analysemöglichkeiten gegeben werden, sowie die Verknüpfung verschiedener Datenbestände durchgeführt wird
- Über die Koordinierungs- und Datenverlinkungsstelle wird die Einhaltung datenschutzrechtlicher Auflagen sichergestellt
- Zusammenfassend sollen die Hürden sowie die Arbeitslast für die Nutzung der Datenbestände reduziert werden und so eine effizientere Nutzung der Ressourcen ermöglicht werden

Die Verknüpfung bestehender Daten ermöglicht insbesondere sektorenübergreifende Langzeitanalysen. Beispielsweise könnte die Versorgung bei Krebserkrankungen von der Diagnose über die Therapie (stationär, ambulant, Rehabilitation, Medikation etc.) bis zu einem definierten Endpunkt (z.B. 5-Jahres-Überleben) analysiert werden (vgl. auch Beispiele aus den Vergleichsländern in Tabelle 4.3).

Für Deutschland wäre der Aufbau einer solchen Koordinierungs- und Verlinkungsstelle wichtig, da hierbei prinzipiell alle verfügbaren Daten berücksichtigt werden könnten. Zu fordern wäre ein verknüpfbarer Datenbestand von Daten aller Sozialversicherungsträger und möglichst auch der Privaten Krankenversicherungen, sowie von Daten der Todesursachenstatistik und weiteren Daten (etwa von Leistungserbringern oder Registern).

Möglichkeiten zur Datenverknüpfung sind gegeben. Wie bereits internationale Beispiele als auch Beispiele aus Deutschland zeigen, ist ein Datenlinkage auch ohne einheitliche Schlüsselvariable, beispielsweise durch die Kombinationen verschiedener Informationen (bspw. mittels probabilistischem Linkage) möglich. Wünschenswert wäre jedoch die Verwendung eines einheitlichen Pseudonyms, das bei der Datengenerierung an verschiedenen Stellen in einheitlicher Weise vergeben wird und damit eine sichere Verknüpfung ermöglicht.

Die koordinierende Instanz sollte die folgenden Daten verknüpfbar zur Verfügung stellen bzw. den Zugriff und die Verknüpfung koordinieren:

- *auf Individualebene verlinkbar*
 - Daten aller gesetzlichen Krankenkassen
 - Daten der privaten Krankenversicherungen nach gleicher Schnittstellendefinition
 - Daten der gesetzlichen Rentenversicherung
 - Daten der gesetzlichen Unfallversicherung
 - Ggf. zukünftig Daten der elektronischen Patientenakte
 - Ggf. zukünftig Befragungsdaten (z.B. hinsichtlich subjektiver Gesundheit, Erfahrungen mit dem Gesundheitssystem), welche notwendige Informationen für die Verlinkung sowie das Einverständnis der Befragten erfassen
- *auf aggregierter Ebene verlinkbar (bspw. Regionen oder auch per Institutskennezeichen)*
 - Auf Ebene der Leistungserbringer aggregierte Daten von kassenärztlichen Vereinigungen und Krankenhäusern (auch IQTIG)
 - Daten der Todesursachenstatistik
 - Ggf. weitere Daten (z.B. Krebs- und Implantateregister, bestehende Befragungsdaten (bspw. des RKI))

Des Weiteren sollte angestrebt werden, dass auch Daten aus einzelnen Forschungsprojekten mit diesem Datenbestand verknüpft werden können. So könnten z.B. Daten aus Befragungen oder auch aus Randomisierten Kontrollierten Studien (nach Zustimmung der Studienteilnehmer) mit den vorliegenden Routinedaten verknüpft werden, um die Effizienz der Datenerhebung zu erhöhen und ggf. weitere Analysen zu ermöglichen. Ein gutes Beispiel bietet hier die Initiative *Population Data BC* aus Kanada. Ein denkbare Vorgehen wäre, dass Studienteilnehmer um die Bereitstellung ihrer Krankenversicherungsnummer oder anderer Informationen gebeten werden, die eine Verknüpfung der späteren Daten über die Datenverlinkungsstelle möglich machen. Über die Datenverlinkungsstelle bzw. die Vertrauensstelle könnten dann die Daten mittels der bereitgestellten Informationen verlinkt werden und an-

schließlich ein verknüpfter und pseudonymisierter Datensatz für die Auswertung bereitgestellt werden. All dies natürlich nur unter der Zustimmung der Studienteilnehmer und Einhaltung der datenschutzrechtlichen und ethischen Bestimmungen. Dieser pseudonymisierte Datensatz sollte auch anderen Wissenschaftlern zugänglich gemacht werden (ggf. nach Einhaltung einer Schutzfrist), um so den größtmöglichen Nutzen aus den generierten Daten zu erhalten.

Für Forschungszwecke wäre es somit ausreichend, wenn die Daten bei der Datenaufbereitungs- und Verlinkungsstelle in pseudonymisierter Form vorlägen. Für Zwecke der ggf. projektspezifischen Verknüpfung mit weiteren Daten könnte dann zusätzlich auch die Möglichkeit einer De-Pseudonymisierung angedacht werden.

Ein Datenbestand in der hier beschriebenen Art würde einen erheblichen Fortschritt für die Versorgungsforschung, für die Weiterentwicklung der medizinischen Qualität und eine evidenzbasierte Gesundheitspolitik in Deutschland beinhalten. Die gesetzlichen Regelungen für eine solche umfassende Datenzusammenführung müssten vermutlich größtenteils neu geschaffen werden.

c) Erweiterung des Datenkranzes durch freiwillige Bereitstellung der Daten der elektronischen Patientenakte

Bis Januar 2021 soll in Deutschland die elektronische Patientenakte (ePA) eingeführt werden. Die bisherigen Überlegungen zur ePA waren Teil des Referentenentwurfs des Digitale Versorgungsgesetzes (DVG) vom 15.5.2019. Nach der Revision des Entwurfs (Kabinettsentwurf vom 10. Juli 2019, siehe oben), wurden die Regelungen zur ePA jedoch entfernt, da diese in einem eigenen Datenschutzgesetz ausgeführt werden sollten. Dies geschah 2020 mit dem Gesetz zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur (Patientendaten-Schutzgesetz, PDSG).

Nach Einführung der ePA können Versicherte auf eigenen Wunsch die Daten ihrer Akte für die Nutzung zu Forschungszwecken zur Verfügung stellen (§ 363 SGB V). Somit steht zukünftig eine weitere potenzielle Datenquelle für die Versorgungsforschung zur Verfügung.

In der ePA „sollen den Versicherten auf Verlangen Informationen, insbesondere zu Befunden, Diagnosen, durchgeführten und geplanten Therapiemaßnahmen sowie zu Behandlungsberichten, für eine einrichtungs-, fach- und sektorenübergreifende Nutzung für Zwecke der Gesundheitsversorgung, insbesondere zur gezielten Unterstützung von Anamnese und Befunderhebung, barrierefrei elektronisch bereitgestellt werden“ (§341 SGB V i.d.F. des PDSG). Daneben können u.a. auch Daten zum Nachweis der regelmäßigen Inanspruchnahme zahnärztlicher Vorsorgeuntersuchungen, Daten zur Früherkennung von Krankheiten bei Kindern (elektronisches Untersuchungsheft für Kinder), Daten über die ärztliche Betreuung während der Schwangerschaft und nach der Entbindung (elektronischer Mutterpass), Daten der Impfdokumentation (elektronische Impfdokumentation), Gesundheitsdaten, die durch den Versicherten zur Verfügung gestellt werden, Daten des Versicherten aus einer von den Krankenkassen finanzierten elektronischen Akte des Versicherten oder bei den Krankenkassen gespeicherte Daten über die in Anspruch genommenen Leistungen des Versicherten, Daten, die der Versicherte seiner Krankenkasse für die Nutzung in zusätzlichen von der Krankenkasse angebotenen Anwendun-

gen nach § 345 Absatz 1 Satz 1 zur Verfügung stellen kann, Daten zur pflegerischen Versorgung des Versicherten, Daten elektronischer Verordnungen, die ausgestellten Bescheinigungen über eine Arbeitsunfähigkeit und sonstige von den Leistungserbringern für den Versicherten bereitgestellte Daten gespeichert werden.

Alle gesetzlichen Krankenkassen werden verpflichtet, ihren Versicherten eine ePA anzubieten. Die Nutzung der ePA durch die Versicherten ist freiwillig. In den Forschungsdaten basierend auf der ePA sind somit nur Personen enthalten, die a) gesetzlich versichert sind, b) sich für die Nutzung der ePA entschieden haben und c) ihre Daten freiwillig der Forschung zur Verfügung stellen. Es gibt somit ein Doppeltes „opting in“ Verfahren, erst für die Nutzung der ePA und anschließend für die Bereitstellung der Daten.

Dieses Vorgehen kommt im internationalen Vergleich nicht häufig vor. In Australien läuft aktuell die Implementierungsphase von „*My Health Record*“ (MyHR). Im Gegensatz zum deutschen Vorgehen sind initial alle Personen, welche die grundlegenden Kriterien erfüllen, in MyHR eingeschlossen. Es ist dann möglich, der Teilnahme in MyHR zu widersprechen (opting out). Aktuell wird berichtet (Expertenaussage), dass ca. 10 % der Bevölkerung der Teilnahme widersprochen haben. Darüber hinaus ist es möglich, ausschließlich der Nutzung der MyHR-Daten für die Forschung zu widersprechen. Hier sind bisher keine Zahlen bekannt. Bei MyHR gehören die Daten dem Patienten. Dadurch besteht auch die Möglichkeit, der Weitergabe einzelner Daten bzw. der Eintragung bestimmter Informationen zu widersprechen (wenn bspw. die psychische Erkrankung nicht Teil der Akte werden soll). Somit kann es ggf. zu einer Verzerrung durch „soziale Erwünschtheit“ in den dokumentierten Daten kommen. In Österreich können Patienten von der elektronischen Patientenakte ELGA ähnlich wie in Australien aussteigen, vollständig oder für bestimmte Anwendungen (z.B. eBefunde). Eine Verfügbarkeit der ELGA-Daten zu Forschungszwecken ist zwar regulatorisch avisiert, könnte aber bei der Umsetzung schwerfallen. Anekdotischer Evidenz zufolge haben sich die Ausscheidungen nach Bekanntgabe der Forschungszugänglichkeit verdreifacht.

Da sowohl die australischen als auch die österreichischen Erfahrungen nur äußerst eingeschränkt auf Deutschland und das doppelte „opting in“ übertragen werden können, können aktuell keine Aussagen zur späteren „Vollständigkeit“ des deutschen ePA-Datensatzes getroffen werden. Es muss jedoch in jedem Fall bedacht werden, dass die Wahrscheinlichkeit einer Verzerrung durch Nicht-Teilnahme hoch ist. So ist es durchaus denkbar, dass gerade Personen, die sich z.B. aufgrund von Krankheit wenig am öffentlichen Leben beteiligen, und Patienten mit Erkrankungen, die als stigmatisierend empfunden werden, nicht teilnehmen. Dies ist selbstverständlich auch bei einem Opting-Out-Verfahren nicht auszuschließen.

Insgesamt können die Daten damit eher als eine Ergänzung zu bestehenden Daten bzw. zu den aufgeführten Szenarien betrachtet werden. Möglichkeiten die ePA-Daten mit anderen, bestehenden Daten zu verknüpfen, sollten daher zeitnah überprüft werden bzw. idealerweise direkt bei den anstehenden Entscheidungen zur Bereitstellung der ePA-Daten mitgedacht werden. Die Nutzung der ePA-Daten als

Ergänzung würde die Analyse weiterer Fragestellungen ermöglichen. So erlauben die Daten einen detaillierteren Einblick in das Versorgungsgeschehen, als dies über reine Abrechnungsdaten möglich ist. Insbesondere sind in der ePA klinische Befunde (z.B. Laborparameter) enthalten. Damit könnten beispielsweise Krankheitszustände wie Niereninsuffizienz, Herzinsuffizienz oder chronische Lebererkrankung unabhängig von einer (insbesondere im ambulanten Sektor wenig verlässlichen) ICD-Kodierung identifiziert und hinsichtlich der Erkrankungsschwere kategorisiert werden. Damit wären unter anderem Analysen der Versorgung bei chronischen Erkrankungen in Abhängigkeit von Erkrankungsschwere möglich. Daneben könnten auch Leistungen, die vom Patienten selbst gezahlt werden (z.B. Over-the-counter Medikation, IGeL-Leistungen), analysiert werden. International gibt es auch Beispiele, wie das SHARE-Register in Schottland (nicht im Gutachten analysiert), die es Patienten erlauben, ihre Daten bereitzustellen um als potenzielle Kandidaten für wissenschaftliche Studien identifiziert werden zu können.

d) Erweiterung des Datenkranzes durch freiwillige Datenbereitstellung von Referenznetzen von selektierten Leistungserbringern

Während die unter c) angedachte Erweiterung des Datenkranzes durch die ePA detaillierte Daten zum Versorgungsgeschehen auf Patientenebene liefert, ist zudem eine Ergänzung um Referenznetze von selektierten Leistungserbringern möglich, um so detailliertere Daten zum Versorgungsgeschehen auf Ebene der Leistungserbringer zu erhalten (dies könnte sowohl detaillierte medizinische Daten als auch betriebswirtschaftliche Informationen umfassen). Beim Aufbau von Referenznetzwerken kann dabei ein Ansatz gewählt werden, der auf in Deutschland bestehende Netzwerke aufbaut (z.B. den Praxen, die an den Sentinel-Erhebungen des RKI teilnehmen, oder Praxen des ZI Praxispanels). Darüber hinaus ist auch eine Erschließung von neuen, z.B. regionalen Referenznetzen denkbar. Bei einem Aufbau neuer Referenznetze kann u.a. tiefergehend die Koordination der Versorgung analysiert werden und um detaillierte Informationen der Leistungserbringer ergänzt werden. Eine ähnliche Zielsetzung hat die Nivel PC Database in den Niederlanden. Auf Basis von Routinedaten der ambulanten Versorgung wird unter anderem untersucht, inwiefern Leistungserbringer für die Versorgung von einzelnen Patienten zusammenarbeiten und wie sich gesundheitspolitische Maßnahmen auf die Versorgung auswirken. Teilnehmende Praxen bekommen Zugang zu periodischen Feedbackberichten, die ein Benchmarking erlauben und zur Qualitätssicherung beitragen können. Zudem könnte anhand selektierter Referenzpraxen auch getestet werden, welche weiteren Informationen sinnvolle Ergänzungen zu bereits möglichen Analysen (bspw. durch Umsetzung der Szenarien a) und/oder b)) darstellen und daher ggf. innerhalb größerer Studien erfasst werden sollten. Ein ähnliches Modell bietet der häufig für die Forschung herangezogene Clinical Practice Research Datalink (CPRD) im Vereinigten Königreich. Auch hier ist wichtig, dass sowohl Leistungserbringer als auch die bei ihnen registrierten Patienten die Möglichkeit haben, sich für oder gegen eine Teilnahme zu entscheiden. Die Datenlieferung durch selektierte Leistungserbringer wird vermutlich immer hohe Einschränkungen bezüglich der Repräsentativität aufweisen. Daher sollte dieser Vorschlag, ebenso wie Vorschlag c), ausschließlich als Ergänzung zu den Vor-

schlägen a) und b) gedacht werden und nicht als alleinstehende Lösung behandelt werden. Für Gesundheitssysteme ohne „Gatekeeping“-System kommt eine zusätzliche Schicht an Komplexität hinzu: hierzulande sind systembedingt insbesondere auch Limitationen bezüglich der Datenvollständigkeit zu erwarten, da in der Regel kein einzelner Leistungserbringer (anders als z.B. der Hausarzt in Großbritannien) im Besitz aller relevanten Daten ist.

Dieses Gutachten hat das Ziel verfolgt, explorativ interessante Ansätze aus einem internationalen Vergleich zu identifizieren, die für eine Verbesserung der Datengrundlage für die Versorgungsforschung in Deutschland informativ sein können. Eine gründlichere Auseinandersetzung mit einzelnen strukturellen, technischen und politischen Aspekten ist erforderlich, um eine potentielle Umsetzung der hier geschilderten Überlegungen optimal zu unterstützen.

Literaturverzeichnis

- Adelson P, Sharplin GR, Roder DM, Eckert M. (2018). Keratinocyte cancers in South Australia: Incidence, geographical variability and service trends. *Australian and New Zealand Journal of Public Health*, 42(4), 329–333. <https://doi.org/10.1111/1753-6405.12806>
- Agniel D, Kohane I, Weber M. (2018). Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. *BMJ* 360:k1479
- Bachner F, Bobek J, Habimana K, Ladurner J, Lepuschütz L, Ostermann H, et al. (2018). *Austria: Health system review 2018. Health Systems in Transition*. 20(3), 1–256.
- Black A, Statistics New Zealand. (2016). *The IDI prototype spine's creation and coverage*. Retrieved from: http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/idi-resources/linking-methodology-statsnz-idi.aspx (last accessed: 18.11.2019)
- Burke LG, Khullar D, Zheng J, Frakt AB, Orav EJ, Jha AK. (2019). Comparison of Costs of Care for Medicare Patients Hospitalized in Teaching and Nonteaching Hospitals. *JAMA Network Open*, 2(6), e195229. <https://doi.org/10.1001/jamanetworkopen.2019.5229>
- Choe SA, Jun YB, Kim SY. (2018). Exposure to air pollution during preconceptional and prenatal periods and risk of hypertensive disorders of pregnancy: A retrospective cohort study in Seoul, Korea. *BMC Pregnancy and Childbirth*, 18(1). <https://doi.org/10.1186/s12884-018-1982-z>
- Cole A, Garrison L, Mestre-Ferrandiz J, Towse A & others. (2015). Data governance arrangements for real-world evidence. *Office of Health Economics, London, UK*. Retrieved from: <https://www.ohe.org/publications/data-governance-arrangements-real-world-evidence#> (last accessed: 18.11.2019)
- Commonwealth Fund. (2019). The U.S. Health Care System. *The Commonwealth Fund*. Retrieved from: https://international.commonwealthfund.org/countries/united_states/ (last accessed: 18.11.2019)
- Davie G, Lilley R. (2018). Financial impact of injury in older workers: Use of a national retrospective e-cohort to compare income patterns over 3 years in a universal injury compensation scheme. *BMJ Open*, 8(4), e018995. <http://dx.doi.org/10.1136/bmjopen-2017-018995>
- Davoody N, Koch S, Krakau I, Hägglund M. (2019). Accessing and sharing health information for post-discharge stroke care through a national health information exchange platform—A case study. *BMC Medical Informatics and Decision Making*, 19(1). <https://doi.org/10.1186/s12911-019-0816-x>
- Donabedian A (1980). *The Definition of Quality and Approaches to Its Assessment*. Vol 1. Explorations in Quality Assessment and Monitoring. Ann Arbor, Michigan, USA: Health Administration Press.
- Donovan GH, Michael YL, Gatzliolis D, Mannetje A, Douwes J. (2019). Association between exposure to the natural environment, rurality, and attention-deficit hyperactivity disorder in children in New Zealand: A linkage study. *The Lancet Planetary Health*, 3(5), e226–e234. [https://doi.org/10.1016/S2542-5196\(19\)30070-1](https://doi.org/10.1016/S2542-5196(19)30070-1)
- Fiorini S, Hajati F, Barla A, Girosi F. (2019). Predicting diabetes second-line therapy initiation in the Australian population via time span-guided neural attention network. *PLOS ONE*, 14(10), e0211844. <https://doi.org/10.1371/journal.pone.0211844>
- Freed GL, Allen AR. (2018). General paediatrics outpatient consultation fees, bulk billing rates and service use patterns in Australia. *Australian and New Zealand Journal of Public Health*, 42(6), 582–587. <https://doi.org/10.1111/1753-6405.12819>

- Ganguli I, Souza J, McWilliams JM, Mehrotra A. (2019). Association Of Medicare’s Annual Wellness Visit With Cancer Screening, Referrals, Utilization, And Spending. *Health Affairs*, 38(11), 1927–1935. <https://doi.org/10.1377/hlthaff.2019.00304>
- Gaudl R. (2019). The New Zealand Health Care System. *The Commonwealth Fund*. Retrieved from: https://international.commonwealthfund.org/countries/new_zealand/ (last accessed: 18.11.2019)
- Gayle AV, Axson EL, Bloom CI, Navaratnam V, Quint JK. (2019). Changing causes of death for patients with chronic respiratory disease in England, 2005–2015. *Thorax*, 74(5), 483–491. <https://doi.org/10.1136/thoraxjnl-2018-212514>
- Gentil ML, Cuggia M, Fiquet L, Hagenbourger C, Le Berre T, Banâtre A, ... Chapron A. (2017). Factors influencing the development of primary care data collection projects from electronic health records: A systematic review of the literature. *BMC Medical Informatics and Decision Making*, 17(1). <https://doi.org/10.1186/s12911-017-0538-x>
- Gisev N, Pearson SA, Karanges EA, Larance B, Buckley NA, Larney S, ... Degenhardt L. (2018). To what extent do data from pharmaceutical claims under-estimate opioid analgesic utilisation in Australia? *Pharmacoepidemiology and Drug Safety*, 27(5), 550–555. <https://doi.org/10.1002/pds.4329>
- Glover L. (2019). The Australian Health Care System. *The Commonwealth Fund*. Retrieved from: <https://international.commonwealthfund.org/countries/australia/> last accessed: 18.11.2019)
- Gothe H. (2014). Kapitel 19: Routinedaten im Ausland. In E. Swart (Ed.), *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven* (2., vollständig überarbeitete und erweiterte Auflage, pp. 260–268). Bern: Huber.
- Greyer S, Jaunzeme J. (2014). Kapitel 15: Möglichkeiten und Grenzen von Befragungsdaten und Daten gesetzlicher Krankenversicherungen. In E. Swart (Ed.), *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven* (2., vollständig überarbeitete und erweiterte Auflage, pp. 223–233). Bern: Huber.
- Hamad AF, Alessi-Severini S, Mahmud SM, Brownell M, Kuo I. (2019). Prenatal antibiotics exposure and the risk of autism spectrum disorders: A population-based cohort study. *PLOS ONE*, 14(8), e0221921. <https://doi.org/10.1371/journal.pone.0221921>
- Harris C, Daniels B, Ward R, Pearson SA. (2017). Retrospective comparison of Australia’s Pharmaceutical Benefits Scheme claims data with prescription data in HER2-positive early breast cancer patients, 2008–2012. *Public Health Research & Practice*, 27(5). <https://doi.org/10.17061/phrp2751744>
- Häyrynen K, Saranto K, Nykänen P. (2008). Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *Int J Med Inf* 77: 291-304
- Hoffmann F, Glaeske G. (2017). Kapitel 2.9: Analyse von Routinedaten. In H. Pfaff, E. Neugebauer, G. Glaeske, M. Schrappe, S. Zeike, FW. Schwartz, M. Rothmund (Eds.), *Lehrbuch Versorgungsforschung: Systematik—Methodik—Anwendung* (2., vollständig überarbeitete Auflage, pp. 122–127). Stuttgart: Schattauer.
- HSR-Europe. (2011). Health services research into European policy and practice: Final report of the HSREPP project. *Netherlands Institute for Health Services Research*. Retrieved from: <http://www.healthservicesresearch.eu/publications> (last accessed: 18.11.2019)
- Huygens MWJ, Swinkels ICS, Verheij RA, Friele RD, van Schayck OCP, de Witte LP. (2018). Understanding the use of email consultation in primary care using a retrospective observational study with data of Dutch electronic health records. *BMJ Open*, 8(1), e019233. <https://doi.org/10.1136/bmjopen-2017-019233>

- Ihle P, Matusiewicz D, Grobe T, Selke G. (2014). Kapitel 22: Datenhaltung und -analyse aus technischer Sicher. In E. Swart (Ed.), *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven* (2., vollständig überarbeitete und erweiterte Auflage, pp. 304–313). Bern: Huber.
- ISD Scotland. (2016). Quality and Outcomes Framework: Prevalence, achievement, payment and exceptions data for Scotland, 2015/2016. *NHS Scotland*. Retrieved from: <https://www.isdscotland.org/Health-Topics/General-Practice/Quality-And-Outcomes-Framework/> (last accessed: 18.11.2019)
- ISD Scotland. (2019). EDRIS Frequently Asked Questions. Retrieved from: <https://www.isdscotland.org/Products-and-Services/eDRIS/FAQ-eDRIS/> (last accessed: 18.11.2019)
- Jansen T, Verheij RA, Schellevis FG, Kunst AE. (2019). Use of out-of-hours primary care in affluent and deprived neighbourhoods during reforms in long-term care: An observational study from 2013 to 2016. *BMJ Open*, 9(3), e026426. <https://doi.org/10.1136/bmjopen-2018-026426>
- Katz A, Chateau D, Enns JE, Valdivia J, Taylor C, Walld R, McCulloch S. (2018). Association of the Social Determinants of Health With Quality of Primary Care. *The Annals of Family Medicine*, 16(3), 217–224. <https://doi.org/10.1370/afm.2236>
- KBV. (2019, April 25). Elektronische Patientenakte (EPA). Retrieved from: <https://www.kbv.de/html/epa.php> (last accessed: 18.11.2019)
- Keskimäki I, Tynkkynen LK, Reissell E, Koivusalo M, Syrjä V, Vuorenkoski L, ... others. (2019). Finland: Health system review. *Health System Review*, 21(2), 1–166.
- Kim L, Kim JA, Kim S. (2014). A guide for the utilization of Health Insurance Review and Assessment Service National Patient Samples. *Epidemiology and Health*, e2014008. <https://doi.org/10.4178/epih/e2014008>
- Kim TJ, Lee JS, Kim JW, Oh MS, Mo H, Lee CH, ... Yoon BW (2018). Building Linked Big Data for Stroke in Korea: Linkage of Stroke Registry and National Health Insurance Claims Data. *Journal of Korean Medical Science*, 33(53). <https://doi.org/10.3346/jkms.2018.33.e343>
- Ko YJ, Kim S, Park K, Kim M, Yang BR, Kim MS, ... Park BJ. (2018). Impact of the Health Insurance Coverage Policy on Oral Anticoagulant Prescription among Patients with Atrial Fibrillation in Korea from 2014 to 2016. *Journal of Korean Medical Science*, 33(23). <https://doi.org/10.3346/jkms.2018.33.e163>
- Koch-Gromus U, Wegscheider K. (2017). Kapitel 1: Einführung in Konzept und Grundlagen der Versorgungsforschung. In H. Pfaff, E. Neugebauer, G. Glaeske, M. Schrappe, S. Zeike, F. W. Schwartz, & M. Rothmund (Eds.), *Lehrbuch Versorgungsforschung: Systematik—Methodik—Anwendung* (2., vollständig überarbeitete Auflage, pp. 134–139). Stuttgart: Schattauer.
- Kreis K, Neubauer S, Klora M, Lange A, Zeidler J. (2016). Status and perspectives of claims data analyses in Germany - A systematic review. *Health Policy*, 120(2), 213–226. <https://doi.org/10.1016/j.healthpol.2016.01.007>
- Lee J, Lee JS, Park SH, Shin SA, Kim K. (2016). Cohort Profile: The National Health Insurance Service–National Sample Cohort (NHIS-NSC), South Korea. *International Journal of Epidemiology*, dyv319. <https://doi.org/10.1093/ije/dyv319>
- Lu E, Zhu F, Zhao Y, van der Kop M, Synnes A, Dahlgren L, ... Tremlett H. (2014). Birth Outcomes In Newborns Fathered By Men With Multiple Sclerosis Exposed To Disease-Modifying Drugs (P4.155). *Neurology*, 82(10 Supplement). Retrieved from: https://n.neurology.org/content/82/10_Supplement/P4.155 (last accessed: 18.11.2019)

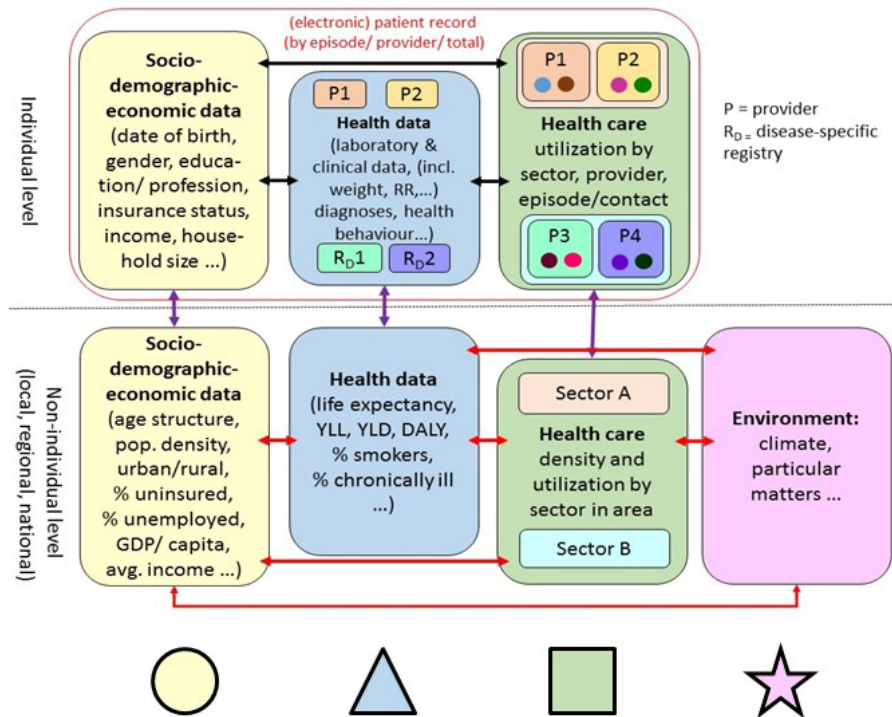
- Mansky T, Nimptsch U. (2010). Notwendigkeit eines ungehinderten Zugangs zu sozial-und krankheitsbezogenen Versichertendaten (...). Expertise im Rahmen der Förderinitiative zur Versorgungsforschung der Bundesärztekammer, Retrieved from: https://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/Datenzugang-2.pdf (last accessed: 16.3.2021)
- Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers AH. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey & Company*. Retrieved from: [https://www.mckinsey.com/~media/McKinsey/Business %20Functions/McKinsey %20Digital/Our %20Insights/Big %20data %20The %20next %20frontier %20for %20innovation/MGI_big_data_exec_summary.ashx](https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_exec_summary.ashx) (last accessed: 18.11.2019)
- March S, Andrich S, Drepper J, Horenkamp-Sonntag D, Icks A, Ihle P, ... Hoffmann F. (2019). Gute Praxis Datenlinkage (GPD). *Das Gesundheitswesen*, 81(08/09), 636–650. <https://doi.org/10.1055/a-0962-9933>
- March S, Stallmann C, Swart E. (2014). Kapitel 25: Datenlinkage. In E. Swart (Ed.), *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven* (2., vollständig überarbeitete und erweiterte Auflage, pp. 347–356). Bern: Huber.
- Maria H, Jonas M, Rose-Mharie Å, Isabella S. (2018). Timing It Right—Patients’ Online Access to Their Record Notes in Sweden. *Studies in Health Technology and Informatics*, 336–340. <https://doi.org/10.3233/978-1-61499-852-5-336>
- May T, Williams K. (2018). Brief Report: Gender and Age of Diagnosis Time Trends in Children with Autism Using Australian Medicare Data. *Journal of Autism and Developmental Disorders*, 48(12), 4056–4062. <https://doi.org/10.1007/s10803-018-3609-7>
- Milne BJ, Atkinson J, Blakely T, Day H, Douwes J, Gibb S, ... Teng A. (2019). Data Resource Profile: The New Zealand Integrated Data Infrastructure (IDI). *International Journal of Epidemiology*, 48(3), 677–677e. <https://doi.org/10.1093/ije/dyz014>
- Morgan A, Joshy G, Schaffer A, Laba TL, Litchfield M, Pearson S, Banks E. (2018). Rapid and substantial increases in anticoagulant use and expenditure in Australia following the introduction of new types of oral anticoagulants. *PLOS ONE*, 13(12), e0208824. <https://doi.org/10.1371/journal.pone.0208824>
- Mues K, Liede A, Liu J, Wetmore J, Zaha R, Bradbury BD, ... Gilbertson D. (2017). Use of the Medicare database in epidemiologic and health services research: A valuable source of real-world evidence on the older and disabled populations in the US. *Clinical Epidemiology, Volume 9*, 267–277. <https://doi.org/10.2147/CLEP.S105613>
- Nimptsch U, Bestmann A, Erhart M, Dudey S, Marx Y, Saam J, ... Swart E. (2014). Kapitel 20: Zugang zu Routinedaten. In E. Swart (Ed.), *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven* (2., vollständig überarbeitete und erweiterte Auflage, pp. 304–313). Bern: Huber.
- OECD. (2015). *Health Data Governance: Privacy, Monitoring and Research*. <https://doi.org/10.1787/9789264244566-en>
- Office of the National Coordinator for Health Information Technology. (2019, February 5). What are the differences between electronic medical records, electronic health records, and personal health records? Retrieved from: <https://www.healthit.gov/faq/what-are-differences-between-electronic-medical-records-electronic-health-records-and-personal> (last accessed: 18.11.2019)
- Padmanabhan S, Carty L, Cameron E, Ghosh RE, Williams R, Strongman H. (2019). Approach to record linkage of primary care data from Clinical Practice Research Datalink to other health-related patient data: Overview and implications. *European Journal of Epidemiology*, 34(1), 91–99. <https://doi.org/10.1007/s10654-018-0442-4>

- Park J, Han E, Lee SO, Kim DS. (2017). Antibiotic use in South Korea from 2007 to 2014: A health insurance database-generated time series analysis. *PLOS ONE*, 12(5), e0177435. <https://doi.org/10.1371/journal.pone.0177435>
- Paun C. (2019). Germany aims to become EU leader in digital health revamp. *Politico*, 10/16/19. Online erhältlich unter <https://www.politico.eu/article/germany-aims-to-become-eu-leader-in-digital-health-revamp/> (letzter Zugriff 15. November 2019).
- Pfaff H, Neugebauer E, Glaeske G, Schrappe M, Zeike S, Schwartz FW, Rothmund M (Eds.). (2017). *Lehrbuch Versorgungsforschung: Systematik - Methodik - Anwendung* (2., vollständig überarbeitete Auflage). Stuttgart: Schattauer.
- Raspe H, Pfaff H, Härter M, Hart D, Koch-Gromus U, Schwartz FW, ... Wittchen HU. (2010). Versorgungsforschung in Deutschland: Stand–Perspektiven–Förderung. *Deutsche Forschungsgemeinschaft (DFG), Bonn*.
- Rice T, Rosenau P, Unruh LY, Barnes AJ, Saltman RB, others. (2013). United States of America: Health system review. *Health Systems in Transition*, 15(3), 1–431.
- Sax Institute. (2019). SURE Users – Summary Steps. Retrieved from: https://www.saxinstitute.org.au/wp-content/uploads/SURE_flow_diagram.pdf (last accessed: 18.11.2019)
- Schrappe M, Pfaff H. (2017). Kapitel 1: Einführung in Konzept und Grundlagen der Versorgungsforschung. In H. Pfaff, E. Neugebauer, G. Glaeske, M. Schrappe, S. Zeike, F. W. Schwartz, & M. Rothmund (Eds.), *Lehrbuch Versorgungsforschung: Systematik—Methodik—Anwendung* (2., vollständig überarbeitete Auflage, pp. 1–68). Stuttgart: Schattauer.
- Schreyögg J. (2017). Plädoyer für eine effektivere Nutzung der Datenbestände in Deutschland für Wissenschaft und Patienteninformation: Expertise im Auftrag der Bertelsmann Stiftung für das Fachgespräch „Datennutzung im Gesundheitswesen“. Hamburg: *Hamburg Centre for Health Economics*. Retrieved from: https://blog.der-digitale-patient.de/wp-content/uploads/sites/9/2017/11/Plädoyer_Nutzung-Datenbestände_Schreyögg_170118.pdf (last accessed: 18.11.2019)
- Schubert I, Swart E, Ihle P, Küpper-Nybelen J, Rentzsch M, Stallmann C, ... Winkler C. (2014). Datengutachten für das Deutsche Institut für Medizinische Dokumentation und Information (DIMDI). *Köln: PMV Forschungsgruppe an Der KJP, Universität Zu Köln*. Retrieved from: <https://www.dimdi.de/static/.downloads/deutsch/dimdi-sekundaerdaten-expertise.pdf> (last accessed: 18.11.2019)
- Sellberg N, Eltes J. (2017). The Swedish Patient Portal and Its Relation to the National Reference Architecture and the Overall eHealth Infrastructure. In M. Aanestad, M. Grisot, O. Hanseth, & P. Vassilakopoulou (Eds.), *Information Infrastructures within European Health Care: Working with the Installed Base* (pp. 225–244). https://doi.org/10.1007/978-3-319-51020-0_14
- Shin S, Burnett RT, Kwong JC, Hystad P, van Donkelaar A, Brook JR, ... Chen H. (2019). Ambient Air Pollution and the Risk of Atrial Fibrillation and Stroke: A Population-Based Cohort Study. *Environmental Health Perspectives*, 127(8), 087009. <https://doi.org/10.1289/EHP4883>
- Smeets HM, Kortekaas MF, Rutten FH, Bots ML, van der Kraan W, Daggelders G, ... de Wit NJ. (2018). Routine primary care data for scientific research, quality of care programs and educational purposes: The Julius General Practitioners' Network (JGPN). *BMC Health Services Research*, 18(1). <https://doi.org/10.1186/s12913-018-3528-5>
- Song S. (2017). HIRA under fire for ‘selling personal data’ to private insurers. *Korea Biomedical Review*. Online erhältlich <http://www.koreabiomed.com/news/articleView.html?idxno=1745> (letzter Zugriff 18.11.2019)

- Spilsbury K, Rosman D, Alan J, Ferrante AM, Boyd JH, Semmens JB. (2017). Improving the Estimation of Risk-Adjusted Grouped Hospital Standardized Mortality Ratios Using Cross-Jurisdictional Linked Administrative Data: A Retrospective Cohort Study. *Frontiers in Public Health*, 5. <https://doi.org/10.3389/fpubh.2017.00013>
- SPIRE. (2019). Your questions answered. Retrieved from: <https://spire.scot/your-questions-answered/> (last accessed: 18.11.2019)
- Spittal MJ, Forsyth S, Pirkis J, Alati R, Kinner SA. (2014). Suicide in adults released from prison in Queensland, Australia: A cohort study. *Journal of Epidemiology and Community Health*, 68(10), 993–998. <https://doi.org/10.1136/jech-2014-204295>
- Statistics New Zealand. (2014). *Linking methodology used by Statistics New Zealand in the Integrated Data Infrastructure project*. Retrieved from: <http://www.stats.govt.nz/~media/Statistics/browse-categories/snapshots-of-nz/integrated-data-infrastructure/linking-methodology-IDI-project.pdf> (last accessed: 18.11.2019)
- Suissa S, Dell’Aniello S, Ernst P. (2019). Comparative effectiveness and safety of LABA-LAMA versus LABA-ICS treatment of COPD in real world clinical practice. *Chest*.
- Tanke MAC, Feyman Y, Bernal-Delgado E, Deeny SR, Imanaka Y, Jeurissen P, ... Meyer GS. (2019). A challenge to all. A primer on inter-country differences of high-need, high-cost patients. *PLoS One*, 14.
- Tew M, Dalziel KM, Petrie DJ, Clarke PM. (2017). Growth of linked hospital data use in Australia: a systematic review. *Australian Health Review* 41:394–400
- Trudeau R. (2017). Social Data Linkage Environment. *International Journal of Population Data Science*, 1(1). <https://doi.org/10.23889/ijpds.v1i1.76>
- Verheij RA, Curcin V, Delaney B C, & McGilchrist MM. (2018). Possible Sources of Bias in Primary Care Electronic Health Record Data Use and Reuse. *Journal of medical Internet research*, 20(5), e185. doi:10.2196/jmir.9134
- Vezyridis P, Timmons S. (2016). Evolution of primary care databases in UK: a scientometric analysis of research output. *BMJ Open* 6:e012785
- Von der Leyen U, Spahn J (2019). Soziale Marktwirtschaft für den europäischen Datenraum. *Frankfurter Allgemeine Zeitung* 10/10/19. Online erhältlich unter <https://www.faz.net/aktuell/politik/von-der-leyen-und-spahn-europaeischen-weg-fuer-umgang-mit-daten-finden-16426668.html> (letzter Zugriff 15.11.2019)
- Wammes J, Jeurissen P, Westert G, Tanke M. (2019). The Dutch Health Care System. *The Commonwealth Fund*. Retrieved from: <https://international.commonwealthfund.org/countries/netherlands/> (last accessed: 18.11.2019)
- Young A, Flack F. (2018). Recent trends in the use of linked data in Australia. *Australian Health Review*, 42(5), 584. <https://doi.org/10.1071/AH18014>

Anhang

Legende für die Steckbriefe pro Fallstudie: Gesundheitsdaten werden mit Dreiecken, Gesundheitsversorgungsdaten mit Quadraten, soziodemographische Daten mit Kreisen und Umweltdaten mit Sternen gekennzeichnet.



Australia

Australia (AUS)

Population	24,992,369	1
Administrative structure of country	Federal parliamentary democracy under a constitutional monarchy.	2, 3
Jurisdiction of health policy	Three levels of government are collectively responsible for providing universal health care: the federal government, the states/territories and local government. The former is responsible for regulating private health insurance, pharmaceuticals, and therapeutic goods, but has a limited role in direct service delivery.	3
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 9.4 % of GDP • 4,919.2 US\$ PPP 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	17.8 % (2018)	4
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing in AUS is a combination of public government (at federal and state levels) funding and private funding.</p> <p>Public healthcare insurance: Medicare, this regionally administered and federal government-run health scheme, is funded by earmarked income taxes, through what is called the Medicare Levy (2 % of taxpayers' taxable income), and general tax revenue from the federal and states level of government.</p> <p>Voluntary, private health insurance (VHI): There are two types of VHI in AUS: (1) complementary, which provides hospital cover and treatment cover for services excluded from Medicare and (2) supplementary, which increases choice and faster access to general nonemergency services. VHI is financed through private finances; the government encourages registration through a tax rebate and a penalty payment called the Medicare Levy surcharge for individuals above a certain income.</p> <p>OOP payments: OOP payments are increasing in AUS for services and medicines that do not fall under the Medicare Benefits Schedule (MBS) or the Pharmaceutical Benefits Schedule (PBS).</p>	3, 5
Extent of coverage by insurance type	<p>Medicare: There is universal health coverage, primarily through Medicare, for Australian and New Zealand citizens, permanent residents and people from countries with reciprocal agreements.</p> <p>VHI: Around 46 % of population has some form of private-patient hospital cover; about 55 % of population has a form of general treatment cover.</p>	3, 5, 6,
Organization of health service delivery		
Primary care - gatekeeper function?	Yes; a referral is needed from a GP for a patient to receive MBS subsidies for specialist services.	3
Specialist outpatient care	Specialists deliver outpatient care in private practice or in public hospitals; many specialists split their time between private and public practice.	3
Payment mechanisms	<p>Hospital care: There are both public and private hospitals in AUS. Public hospital payment comes through global budgets from states and case-based payment (including physician costs). States finance budgets from a combination of tax revenues. Private hospitals are paid via fee-for-service through PHI. A DRG payment system (AR-DRG) is in use to match resources and activities.</p> <p>Primary care: GPs are largely self-employed, work in joint practices and are primarily paid fee-for-service, with fees set by the federal health minister through MBS; some use of quality-of-care activity incentive payments.</p>	3, 6

Australia

	<p>Specialist care: Specialists are paid on fee-for service basis; they receive a subsidy through the MBS of up to 85 % and can set patients' OOP fees independently. Global budgets and bulk billing are also used to a degree.</p> <p>Other: Capitation is employed to a limited degree mainly to fund healthcare in under serviced areas, e.g., remote Indigenous communities.</p>	
Examples of health data protection laws	<ul style="list-style-type: none"> • The Privacy Act of 1988 • The Privacy Regulation 2013 • The Privacy Amendment Act 2017 • The Information Privacy Act 2014 (Australian Capital Territory) • The Information Act 2002 (Northern Territory) • The Privacy and Personal Information Protection Act 1998 (New South Wales) • Information Privacy Act 2009 (Queensland) • Personal Information Protection Act 2004 (Tasmania) • Privacy and Data Protection Act 2014 (Victoria) 	7, 8, 9
Example of health data strategies	National Digital Health Strategy	10
Index of technical infrastructure	N/A	
Relevant actors	<p>Australian Government Department of Health: Develops and delivers policies and programs related to health, older people's care and sport.</p> <p>Australian Digital Health Agency: Established in 2016, tasked with improving health outcomes via delivery of national digital health services and systems, with focus on engagement, innovation and clinical quality and safety; it is the system operator of the My Health Record.</p> <p>Services Australia: Responsible for delivering a range of welfare, health, child support payments and other services to AUS' population, including Medicare and the PBS.</p> <p>Australian Institute for Health and Welfare (AIHW): AUS' national agency for information and statistics on health and welfare; holds data on a wide range of health and welfare topics.</p> <p>Commonwealth Accredited Integrating Authorities and Data Linkage Units (DLUs): Agencies that are approved to undertake data linkage, including: Population Health Research Network (PHRN)</p> <ul style="list-style-type: none"> • Centre for Victorian Data Linkages (VDL) • SA-NT Datalink • Data Linkage Queensland (DLQ) • Tasmanian Data Linkage Unit (TDLU) • Western Australia Data Linkage 	7, 11

Australia

List of sources:

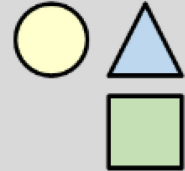
- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] Australia – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] The Commonwealth Fund. (2017) International Health System Profile – Australia. Retrieved from: <https://www.commonwealthfund.org/international-health-policy-center/countries/australia>
- [4] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [5] Department of Health. (2019). The Australian health system. Australian Government. Retrieved from: <https://www.health.gov.au/about-us/the-australian-health-system>
- [6] Australian Institute of Health and Welfare. (2018). Australia’s health 2018. Australias health series no. 16. AUS 221. Canberra. Retrieved from: <https://www.aihw.gov.au/getmedia/63fe0895-b306-4375-95ff-162149ffc34b/aihw-aus-221-chapter-2-1.pdf.aspx>
- [7] Office of the Australian Information Commissioner. (2019). Health and medical research. Australian Government. Retrieved from: <https://www.oaic.gov.au/privacy/the-privacy-act/health-and-medical-research/>
- [8] Office of the Australian Information Commissioner. (2019) Privacy Regulation 2013. Australian Government. Retrieved from: <https://www.oaic.gov.au/privacy/the-privacy-act/privacy-regulations>
- [9] Office of the Australian Information Commissioner. (2019) Privacy. Australian Government. Retrieved from: <https://www.oaic.gov.au/privacy>
- [10] Australian Government. (2018). Australia’s National Digital Health Strategy: Framework for Action – How Australia will deliver the benefits of digitally enable health and care. Retrieved from: https://www.digitalhealth.gov.au/sites/default/files/2020-11/Framework_for_Action.pdf
- [11] Population Health Research Network (PHRN). (2019) Retrieved from: <https://www.phrn.org.au/>

My Health Record

Care coordination, patient empowerment, research

Data content

My Health Record contains all information related to a patient's health care, e.g. care planning documents, emergency contact information, and reports from tests and scans



Data collection

Is the data structured?
No

Who collects the data?
Individual patient and health care providers



Data linkage and cleaning

At what level is the data aggregated?
Data aggregated at national level by the Australian Digital Health Agency

Is the data checked or cleaned?
A Data Management Committee manages operational issues

Is the data linkable?
Yes, data is linkable through a patient identifier



Data storage

Where is the data located?
Australian Institute of Health and Welfare is the data custodian

How long is the data stored?
Data stored 30 years from time of death or deleted permanently upon request

Who owns the data?

Individual patients own their data

Who performs the data linkage?

The Australian Digital Health Agency stores and links the data

Who has access to data?

Australian-based entities (except insurance agencies) can apply to access data for secondary use



Research use

What is the process for acquiring access?
Application precondition to access to My Health Record data for secondary use purposes

How long does it take to become available?
It varies how long it takes for data to become available



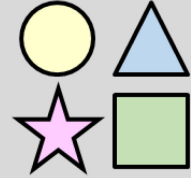
How is the data used for research purposes?
My Health Record is currently being rolled out for research purposes

PHRN

Aggregate and prepare data from around Australia for research

Data content

PHRN contains information including 1) health, 2) health care, 3) administrative, clinical, and registry data. It also has 4) de-identified demographic data and data on 5) housing, 6) justice, 7) education, and 8) some environmental data.



Data collection

Is the data structured?
Yes

Who collects the data?
Different agencies or programs are responsible for collecting data



Data linkage and cleaning

At what level is the data aggregated?
The data is aggregated on regional or national level

Is the data checked or cleaned?
PHRN supports network of regional data linkage units which service each state and territory

Is the data linkable?
Yes, data is linked through patient identifiers



Data storage

Where is the data located?
The data is located throughout Australia and is managed by the agency responsible for collection

How long is the data stored?
N/A

Who owns the data?

The various agencies and programs collecting the data serve as custodians

Who performs the data linkage?

Territorial data linkage units perform the data linkage

Who has access to data?

Researchers who meet the purpose and ethical requirements and submit a fee are granted access to the data after a period of waiting



Research use

What is the process for acquiring access?
Researchers must complete an application via PHRN Online Application System as a prerequisite for acquiring access

How long does it take to become available?
It varies according to the data collection processes

How is the data used for research purposes?
N/A



MBS

Health system governance

Data content

The MBS contains a current and historical record of all medical and hospital services that are subsidized by the government under the MBS. Information is available for all Australian residents and particular categories of visitors



Data collection

Is the data structured?
Yes

Who collects the data?
The Australian Government collects the data



Data linkage and cleaning

At what level is the data aggregated?
Data is aggregated at national level but varies slightly by jurisdiction

Is the data checked or cleaned?
N/A

Is the data linkable?
Yes, data linkable through universal national healthcare number which is then provided to end-user using a de-identified patient ID



Data storage

Where is the data located?
The data is located with the Australian Government

How long is the data stored?
N/A

Who owns the data?

The Australian Government and the Department of Human Services own the data

Who performs the data linkage?

Linkage must be facilitated either through PHRN or the Australian Institute of Health and Welfare (AIHW)

Who has access to data?

PHRN and AIHW regulate access to data



Research use

What is the process for acquiring access?
Application to the Department of Human Services' External Request Evaluation Committee (EREC) and AIHW Ethics

How long does it take to become available?
Data ≤ 5 years old is easily accessible.
Data ≥ 5 years is archived and the retrieval process more lengthy and costly



How is the data used for research purposes?
The information is used for essential analysis of the MBS, for example investigating service use and costs

Denmark

Denmark (DNK)

Population	5,797,446	1
Administrative structure of country	Parliamentary constitutional democracy	2
Jurisdiction of health policy	With Denmark's decentralized health system, health policy is the joint responsibility of the national government and the regions. The national government sets the regulatory framework for health services and is responsible for general planning, quality monitoring care quality, and licensing as well as collecting taxes to be allocated to local governments. Five regions are owners of hospitals and responsible for the planning and delivery of specialized health care services and coordination of specialized social care.	3
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 10.0 % of GDP • 5,477.6 US\$PPP 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	14.2 %	4
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing is mostly a combination of public financing and private sources.</p> <p>National health care system: the compulsory, national health care system in DNK is financed through taxes set at the national and local level. Revenues from the national level are allocated by the state to regions and municipalities, mostly as block grants. The municipalities finance around 20 % of regional health care services, in order to encourage the municipalities to prioritize prevention of disease. Municipal taxation is determined by each municipality individually up to a ceiling of 51.5 % of income. A very small part of the regions' funding (1 %) is contingent on the regions' fulfillment of criteria indicating that they cooperate well with the municipalities (Nærhedsfinansiering), e.g. reduction in hospital contacts per citizen, reduction in number of readmissions and increasing use of virtual solutions.</p> <p>Voluntary, private health insurance (VHI): There are two types of PHI, purchased by individuals: (1) complementary voluntary insurance, through almost exclusively, the not-for-profit Danmark, covers statutory cost-sharing, benefits excluded in the statutory package and OOP payments and (2) supplemental insurance to gain access to private providers.</p> <p>OOP payments: There is no cost-sharing for hospital and primary services, but for dental care, outpatient prescriptions and glasses. There is no cap on OOP costs, except for drug OOP spending.</p>	3, 5, 6, 7
Extent of coverage by insurance type	<p>National healthcare system: the national healthcare system provides universal healthcare to all registered residents.</p> <p>VHI: 39 % have complementary coverage for cost-sharing, etc., while around 26 % have supplementary coverage for access to private providers.</p>	3, 4, 6
Organization of health service delivery		
Primary care - gatekeeper function?	Yes; citizens are free to choose between two coverage options in the national health system (1) to register with a GP of their choice within 15km of their home (98 % of covered population) and (2) to not register and be free to consult any GP and any specialist without referral.	3, 6
Specialist outpatient care	After referral from a GP, office-based specialists and hospitals-based ambulatory clinics provide specialist outpatient care.	3, 6
Payment mechanisms	<p>Hospital care: Hospitals in DNK are almost all public and payment mechanisms include global budgets from regions.</p> <p>Primary care: Most GPs are self-employed and are mainly paid through fee-for-service, with some capitation. Regions are responsible for payment, though rates are set nationally in agreement with doctors' associations.</p>	3, 6

Denmark

	<p>Specialist care: Services by private specialists are paid by the regions via fee-for-service for referred patients. These fees are set in negotiation with regions. Private specialists and hospitals can also get OOP payments or funding through VHI. Specialists can set their own fees for patients coming without referral. Specific agreements with regions determine fees for services rendered at private institutions for patients referred from public sector. Specialist outpatient care delivered in hospitals is paid via the general payment mechanisms for hospital care (above).</p>	
Examples of health data protection laws	<ul style="list-style-type: none"> • EU Regulation 2016/679 (General Data Protection Regulation) • Danish Data Protection Act 2018 (<i>Databeskyttelsesloven</i>) 	8, 9, 10
Example of health data strategies	With DNK's latest national "Digital Health Strategy 2018-2022", elaborated by the Danish Ministry of Health, the Danish Ministry of Finance, Danish Regions, and Local Government Denmark, the country attempts to create a coherent and trustworthy health network for all.	11, 12
Index of technical infrastructure	DNK has a strong health data information system, with high health data availability and coverage. It reports lower use of data for research/monitoring purposes (OECD's Key national dataset availability, maturity and use score: 5.70/7.0).	13
Relevant actors	<p>Ministry of Health (SUM; Sundhedsministeriet): Provides the basis for political decisions in healthcare and elderly care; contributes to formulating agendas, developing solutions and ensuring coherence.</p> <p>Danish Health Authority (SST; Sundhedsstyrelsen): Responsible for supervising health personnel and institutions and for advising different ministries, regions and municipalities on health issues. Recent reforms gave the DHA increased power with the responsibility for planning the distribution of medical specialties among hospitals.</p> <p>Regional authorities and municipalities: The five regions responsible for GPs, specialists, physiotherapists, dentists, chiropractors and pharmacists represented by their interest organizations. Both collect data on healthcare services as well as demographic and economic data.</p> <p>Statistics Denmark (DST; Danmarks Statistik): A state institution under the Ministry of Economic Affairs and the Interior and the central producer of statistics in Denmark analyzing social and economic conditions.</p> <p>The Danish Clinical Quality Program – National Clinical Registries (RKKP; Regionernes kliniske kvalitetsudviklingsprogram): Responsible for ensuring continuous improvement in the utilization of registries for clinical, managerial and research. RKKP facilitates the use of available data, designs new registers and works on optimizing existing ones.</p> <p>Sundhed.dk: Provides quality assured health information, access to medical records and medication, and an overview of Danish health care. It offers a shared communication platform; aims for patient empowerment through transparency and quality improvement by bringing together all relevant information; and offers easy access to individual patients' clinical information for healthcare providers.</p>	3, 6, 14, 15, 16, 17

List of Sources:

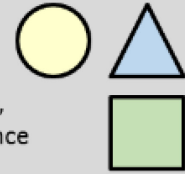
- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] Denmark – The World Factbook (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Vrangbæk, K. (2017). The Danish health care system. In Mossialos, E., Djordjevic, A., Osborn, R. & Sarnak, D. (Eds.), International profiles of health care systems. The Commonwealth Fund. Retrieved from: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf
- [4] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [5] Oresunddirekt. (2019). Danish income tax – a short introduction. Retrieved from: <https://www.oresunddirekt.se/en/working-in-denmark/taxes/danish-income-tax-a-short-introduction>
- [6] Eds. Mossialos, E., Djordjevic, A., Osborn, R., Sarnak, R. (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Retrieved from: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf
- [7] Ministry of Health. (2017). Healthcare in Denmark. Retrieved from: <https://www.sum.dk/English/Healthcare-in-Denmark-An-Overview/>
- [8] Federspiel, G. (2018). Data protected – Denmark. Linklaters. Retrieved from: <https://www.linklaters.com/de-de/insights/data-protected/data-protected---denmark>
- [9] General Data Protection Regulation (GDPR). (2018). General Data Protection Regulation (GDPR) – Final text neatly arranged. Retrieved from: <https://gdpr-info.eu>
- [10] Government of Denmark. (2018). Datatilsynet Retrieved from: <https://www.datatilsynet.dk/english/legislation/>
- [11] The Danish Ministry of Health, the Danish Ministry of Finance, Danish Regions, and Local Government Denmark. (2018). Danish Digital Health Strategy 2018-2022. Retrieved from: <https://www.healthcaredenmark.dk/news/danish-digital-health-strategy-2018-2022-now-available-in-english.aspx>
- [12] Healthcare Denmark. (2018). A Coherent and Trustworthy Health Network for All: Digital Health Strategy 2018-2022. Retrieved from: <https://www.healthcaredenmark.dk/news/listnews/danish-digital-health-strategy-2018-2022-now-available-in-english/>
- [13] OECD. (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en>
- [14] Sundhed.dk. (2016). Background. Retrieved from: <https://www.sundhed.dk/borger/service/om-sundheddk/ehealth-in-denmark/background/>
- [15] Danish Health Authority (Sundhedsstyrelsen). (2019). About us. Retrieved from: <https://www.sst.dk/en/English>
- [16] Statistics Denmark (Danmarks Statistik). (2019) About us. Retrieved from: <https://www.dst.dk/en>
- [17] The Danish Clinical Quality Program – National Clinical Registries (Regionernes kliniske kvalitetsudviklingsprogram). (2019). In English. Retrieved from: <https://www.rkkp.dk/in-english/>

Sundhed.dk

Facilitates data flow and communication between patients, health professionals and improves transparency

Data content

Sundhed.dk includes individual level health data such as 1) electronic patient record (clinical data, medical history, medications, lab results, etc.) as well as 2) sociodemographic data and 3) healthcare utilization data (appointments, treatments, prescriptions). Aggregated data are also contained, such as 4) health system performance (benchmarking service and quality of health services, waiting lists for operations)



Data collection

Is the data structured?
Varies; yes and no

Who collects the data?
The GP or specialist's office, in the hospital patient administrative system (PAS), and hospital electronic health record system



Data linkage and cleaning

At what level is the data aggregated?
Not done within sundhed.dk, but some included sources are aggregated (e.g., registries)

Is the data checked or cleaned?
The data is displayed as close to source data as possible, often without correction

Is the data linkable?
Yes, data can be linked to Central Person Register



Data storage

Where is the data located?
Stored locally at the health care providers. Access is managed via The Danish Healthcare Data Network operated by MedCom

How long is the data stored?
No restrictions on data storage periods

Who owns the data?

Depends on the type of data. Systemic health data is usually owned by the organization generating the data. Data from GPs are submitted to the Danish Quality Unit of General Practice, who makes it accessible to patients via Sundhed.dk. Hospital data is transferred to the E-Record (repository) database

Who performs the data linkage?
N/A

Who has access to data?
All patients and health care providers



Research use

What is the process for acquiring access?
Sundhed.dk is not accessible for research directly through the portal, but the data can be used for research by permission from the relevant organization

How long does it take to become available?
N/A

How is the data used for research purposes?
Sundhed.dk is not used for research purposes

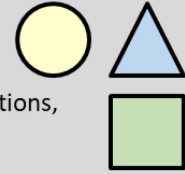


RKKP

Improving quality and research utilization of Danish clinical registers

Data content

RKKP includes 90 different databases. Coverage of each database varies and they use different indicators. The databases hold information and variables that among other things cover 1) clinical data, 2) assessments of diagnosis, 3) treatment, 4) medications, 5) laboratory results, 6) rehabilitation, 7) patient-reported outcomes (PRO) data, and 8) sociodemographic data



Data collection

Is the data structured?
Yes

Who collects the data?
Clinicians collect data. Other data is collected through other registries (e.g. the national clinical registries)



Data linkage and cleaning

At what level is the data aggregated?

Individual level data is aggregated on department, hospital, regional and national level

Is the data checked or cleaned?

Yes, RKKP is obliged to supply feedback monthly/quarterly, and produces an annual analytic report with recommendations per database

Is the data linkable?

Yes, data is linkable to the Central Person Register



Data storage

Where is the data located?

Region Central Jutland (Midtjylland) is responsible for data storage

How long is the data stored?

Central Jutland will delete personal information that is not needed for further processing

Who owns the data?

RKKP is exempt from patient consent to data collection. Central Jutland is the responsible institution

Who performs the data linkage?

RKKP

Who has access to data?

Researchers and health care providers have general access. Clinicians have access to individual-level data for patients treated in their unit. Others have access to aggregated data on sundhed.dk



Research use

What is the process for acquiring access?

Researchers must submit an application to RKKP to access the registers. For the access to anonymized individual-level data, researchers need to be affiliated with an authorized Danish research environment

How long does it take to become available?

The data is available to researchers no later than 3 month after a completed application via remote online databases. RKKP also provides access to monthly updated data



How is the data used for research purposes?

Data is used to develop and analyze quality indicators. In regional and national audits, clinicians determine indicators and standards for good clinical quality, and report recommendations back to clinical personnel

Finland

Finland (FIN)

Population	5,518,050	1
Administrative structure of country	Parliamentary republic; unitary state organized on a decentralized basis, with central, regional and local levels of governance.	2
Jurisdiction of health policy	The health system has a highly decentralized administration, multiple funding sources, and three separate channels for delivering services in first-contact care. The role of the state is to oversee and steer the system's functioning through legislation, decrees and the provision of information and the core health system is organized by the municipalities.	3
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 9.2 % of GDP • 4,558.5 US\$ PPP 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	17.4 %	4
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing in FIN is a combination of public financing, statutory health insurance (SHI) and private sources.</p> <p>Municipal healthcare and statutory national health insurance (NHI): Funding for healthcare comes through state-raised and municipal taxes, as well as the statutory NHI scheme. The NHI scheme is financed by state transfers and compulsory income-based insurance fees from employers and employees.</p> <p>Voluntary, private health insurance (VHI): There are two types of VHI in FIN: (1) complementary, which covers NHI user charges and (2) supplementary, which ensures more choice and faster access to municipal healthcare.</p> <p>Occupational healthcare: Employers are responsible for organizing health services for employees, and NHI covers about half of employers' costs.</p> <p>OOP payments: OOP costs come from inpatient hospital care cost-sharing and some user fees for municipal services and private services as NHI reimbursement levels have not kept pace with actual service fees.</p>	3
Extent of coverage by insurance type	<p>NHI: The NHI provides universal healthcare for all permanent residents.</p> <p>Municipal health services: These services are available to all residents in a respective municipality.</p> <p>VHI: Altogether about 15 % of the population has some type of VHI.</p> <p>Occupational healthcare: Employees are covered by occupational health care, the scope of which varies (about 33 % of total population).</p>	3
Organization of health service delivery		
Primary care - gate-keeper function?	Yes; patients can choose where to receive specialist services with a referral from a GP.	3
Specialist outpatient care	Hospital districts (20) are responsible for specialized care. Specialized care, both out- and inpatient, is provided by five different organizations (four are owned by hospital districts and one by one or more municipality): (1) Five university hospitals; (2) central hospitals; (3) university clinic departments; (4) psychiatric inpatient facilities; and (5) specialist-run health center hospitals (14). Private clinics can also offer services.	3
Payment mechanisms	Hospital and specialist care: Hospital districts providing inpatient and specialist care are mainly funded by the municipalities they are made of. Funding most comes from a fixed part based on population and a part based on services used . Hospital districts use DRGs for invoicing and have a funding pool to cover high individual patient expenses.	3

Finland

	<p>Physicians in hospital districts are salaried employees, though many may work in private clinics on a fee-for-service basis.</p> <p>Primary care: Municipalities prospectively fund budget for health centers; costs usually determined by previous years' budgets. In centers run by joint-municipal health authorities, budgets are developed similarly, but costs are determined by volume of services provided. GPs working in municipal centers have variable salary scheme, including monthly salary with some extra fee-for-service for selective services. In centers where a "personal doctor system" has been introduced, payment is via basic salary, capitation and fee-for-service for visits.</p> <p>Occupation health services: These are paid by employers using a range of methods, including partial reimbursement by NHI (50-60 %).</p>	
Examples of health data protection laws	<ul style="list-style-type: none"> • The Act on Secondary Use of Health and Social Data 201 • The Data Protection Act 2018 	5, 6, 7
Example of health data strategies	<ul style="list-style-type: none"> • The Health-Sector Growth Strategy for Research • eHealth and eSocial Strategy 2020 	8, 9, 10
Index of technical infrastructure	<p>FIN has a strong health information system, with high data availability, though it is less frequently used for reporting on system performance and care quality (OECD's Key national dataset availability, maturity and use score: 5.47/7.0).</p>	11
Relevant actors	<p>Finnish Institute for Health and Welfare (THL; Terveystieteiden tutkimuskeskus ja hyvinvointivirasto): Responsible for studying, monitoring and developing measures of health and well-being, as well as providing information for various actors based on research and register data.</p> <p>The Social Insurance Institution (Kela; Kansaneläkelaitos): a government agency responsible for settling benefits under national social security programs, including retirement pay, child benefits, unemployment benefits, sickness benefits, health insurance and student financial aid. As such, Kela's statistics provide information about the Finnish social security system. The statistics aid the development of and research on the social security system</p> <p>Kanta Services: Finland's e-health records system serving as one repository for all health-related data. In the future, Findata will be able to sample data directly from Kanta Services rather than going to each individual data controller. Kanta Services is maintained by Kela.</p> <p>Statistics Finland (Tilastokeskus): A large amount of socioeconomic data available through this institution.</p> <p>Findata: The new data permit authority overseeing data regulation. It is the single-point contact for permits to access several different databases. It safeguards secure and efficient procedures in utilizing the data in research, development and innovation and works in collaboration with the THL, but separately from the Institute's other activities.</p> <p>Additional controllers providing access to data specified through Act on Secondary Use of Health and Social Data include:</p> <ul style="list-style-type: none"> • Social and health care operating units • Finnish Centre for Pensions (ETK; Eläketurvakeskus) • The National Supervisory Authority for Welfare and Health (Valvira; Sosiaali- ja terveystieteiden tutkimuskeskus ja valvontavirasto) • Finnish Medicines Agency (Fimea) • Regional State administrative agencies • Väestörekisterikeskus (VRK; Population Register Centre) • Työterveyslaitos (TTL; Finnish Institute of Occupational Health) 	2, 8, 9, 10, 12

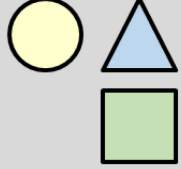
List of sources:

- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] Finland – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Keskimäki, I., Tynkkynen, L-K., Reissell, E., Koivusalo, M., Syrjä, V., Vuorenkoski, L., Rechel, B. & Karanikolos, M. (2019). Finland health system review. Health systems in transition, 21(2). Retrieved from: <https://apps.who.int/iris/bitstream/handle/10665/327538/18176127-eng.pdf?sequence=1&isAllowed=y>
- [4] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [5] Seppänen, J. (2019). Findata – Secondary use of Finnish social and health data – a new act and Data permit authority. [Presentation] Retrieved from: https://www.himss-europe-conference.eu/sites/himss-europe-conference/files/sponsors/himss-eu-19/presentations/seppanen_secondary_use_of_finnish_social_and_health_data_-_a_new_act_and_data_permit_authority.pdf
- [6] Dammert, J. (2019). Finland: Parliament approves new act on the secondary use of social and health care personal data. Privacy matters – DLA Piper’s global privacy and data protection resource. Retrieved from: <https://blogs.dlapiper.com/privacymatters/finland-parliament-approves-new-act-on-the-secondary-use-of-social-and-health-care-personal-data/>
- [7] Bird & Bird. (2018). Finland supplements the GDPR: the Parliament approves the new Data Protection Act. Retrieved from: <https://www.twobirds.com/en/news/articles/2018/finland/eduskunta-hyvaksyi-uuden-tietosuojalain>
- [8] Ministry of Employment and the Economy. (2016). Innovating together – Health sector growth strategy for research and innovation activities. Roadmap for 2016-2018. Retrieved from: https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/75145/MEE_guidelines_8_2016_Health_sector_growth_strategy_17062016_web.pdf
- [9] Ministry of Social Affairs and Health. (2015). Information to support well-being and service renewal – eHealth and eSocial strategy. Retrieved from: http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/74459/URN_ISBN_978-952-00-3575-4.pdf?sequence=1&isAllowed=y
- [10] Malkamäki, S. & Hämäläinen, H. (2019). New legislation will speed up the use of Finnish health data. Retrieved from: <https://www.sitra.fi/en/blogs/new-legislation-will-speed-use-finnish-health-data/>
- [11] OECD. (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en>
- [12] Finnish Institute for Health and Welfare. (2019). Data permit authority Findata. Retrieved from: <https://thl.fi/en/web/thlfi-en/statistics/data-and-services/data-permit-authority-findata>

Findata

New data permit authority oversees data regulation under the Act on the Secondary Use of Health and Social Data. It is the single-point of contact for permits to access multiple databases and ensures secure and efficient procedures for data utilization in research, development and innovation.

Data content
National social and healthcare registers, data from operational client and patient systems in primary care, specialist care and social services



Data collection

Is the data structured?
Yes (presumed)

Who collects the data?
Government agencies and collaborative health system services

Data linkage and cleaning

At what level is the data aggregated?
Will be aggregated at national level, but also available for multiple levels of jurisdictions

Is the data checked or cleaned?
N/A

Is the data linkable?
Datasets will be made available without individual's direct identifying data. Unclear what methods will be used.

Data storage


Where is the data located?
At the new national Social and Health Data Permit Authority Findata

How long is the data stored?
N/A

Who owns the data?
Patients ultimately own their data, but until 2020 each data owner controls access to data for secondary use. From 2020, Findata will control access and release of data for secondary purposes

Who performs the data linkage?
Findata will see to anonymisation centrally


Who has access to data?
Access to aggregated data can be provided for all purposes specified in the Act on the Secondary Use of Health and Social Data



Research use

What is the process for acquiring access?
Free access to aggregate data will be subject to an information request. A data utilisation plan is required for access to data sets. Users can be granted access to anonymised or pseudonymised data via secured remote access system for a prefixed period. Access via remote access system will be possible for all other purposes referred to in the Act, except for development and innovation.

How long does it take to become available?
N/A



How is the data used for research purposes?
Datasets from data owners, including Kanta Services and registries, is already used extensively for research.

Canada (CAN)

Population	36,290,000	1
Administrative structure of country	Federal parliamentary democracy (Parliament of Canada) under a constitutional monarchy, where authorities and responsibilities are divided between federal government and provincial governments	2, 3
Jurisdiction of health policy	Jurisdiction over healthcare governance and provision lies primarily with the individual provinces and territories.	3, 4
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 10,8 % of GDP • 5,370.4 US\$ PPP 	5
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	14,9 %	5
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing in CAN is a mostly a combination of taxation and private sources.</p> <p>Public insurance program: CAN has a regionally administered universal public insurance program, called ‘medicare’. Provinces/territories (P/T) are most responsible for financing largely private provision; federal government transfers funding. The main source of financing is general taxation at all government levels, e.g., income taxes, consumption taxes, corporation taxes and some resource royalties. General tax revenues are supplemented by health premiums in three provinces: British Columbia (BC), Ontario and Quebec.</p> <p>Private health insurance (PHI): Most private health insurance is employment-based and compulsory. Employees are required to take it on as part of their remuneration and benefits package. PHI covers non-medicare sectors and is paid by employers, unions or other organizations under group- or uninsured contract.</p> <p>OOP payments: OOP costs form the main source of funding for vision care, over-the-counter medicines and complementary and alternative medicines.</p> <p>Other: Other funding sources include voluntary and charitable donations; social insurance comprises the largest share, paid by employers, to cover workplace injuries and ailments.</p>	3, 4
Extent of coverage by insurance type	<p>Medicare: Medicare is a decentralized, universal healthcare system, providing 100 % of residents with “medically necessary” hospital, diagnostics and medical services.</p> <p>PHI: About 66 % of Canadians purchase PHI to cover services excluded from public reimbursement.</p>	3, 4
Organization of health service delivery		
Primary care - gate-keeper function	Yes; many provinces pay lower fees to specialists for non-referred consultations.	3, 4
Specialist outpatient care	Hospitals provide the majority of specialist care; private, non-hospital facilities are also involved in specialist outpatient care provision.	3, 4
Payment mechanisms	<p>Primary care: Most physicians are self-employed and paid fee-for-service across Canada’s provinces/territories; however, there has been movement towards alternative payment methods, e.g., capitation, blended (salary and fee) payments, incentive-based bonuses, pay-for-performance, group-based profit sharing and fundholding systems in various jurisdictions. P/T ministries of Health negotiate fee schedules.</p> <p>Specialist care: Specialists are mostly financed through fee-for-service, though there is some use of global budgets. Specialists have the same fee schedule as primary care physicians.</p>	3, 4, 6

	<p>Hospital care: Hospital ownership is a mix of public and private, mostly managed by regional or community authorities. Hospitals are mostly financed by global budgets made on the basis of the previous year's allocation adjusted for inflation and budget growth. Physicians working in hospitals are paid fee-for-service.</p> <p>Other: Mental health care occurs in specialty and general hospitals as well as in outpatient settings. Psychologists are paid OOP or through PHI, or under salary with publicly funded organizations.</p>	
Examples of health data protection laws	<ul style="list-style-type: none"> • Privacy Act 1985 • Personal Information Protection and Electronic Documents Act (PIPEDA) (2000) • Health information Act (Alberta) • E-Health Personal Health Information Access and Protection of Privacy Act (British Columbia) • Personal Health Information Act (PHIA) (Manitoba) • Personal Health Information Privacy and Access Act (New Brunswick) • Personal Health Information Act and Pharmacy Network Regulations (Newfoundland and Labrador) • Health Information Act (Northwest Territories) • Personal Health Information Act (Nova Scotia) • Access to Information and Protection of Privacy Act (Nunavut) • Personal Health Information Act (Ontario) • Health Information Act (Prince Edward Island) • An Act to amend the Act respecting health services and social services (Quebec) • Health Information Protection Act (Saskatchewan) • Health Information Privacy and Management Act (Yukon) 	7, 8
Example of health data strategies	N/A	
Index of technical infrastructure	CAN has a limited health information system, with medium coverage (60 % of datasets cover 80 % or more of population) and strong use of data to report on care quality and system performance (OECD's Key national dataset availability, maturity and use score: 5.14/7.0)	9
Relevant actors	<p>Federal and provincial government agencies</p> <p>Manitoba Health Information Management Unit: Ensures that data infrastructure, standards and policies are in place to support the appropriate collection, management, use and disclosure of health information in accordance with PHIA.</p> <p>Population Data BC (PopData BC): A multi-university, data and education resource facilitating research on the human health, well-being, and development via access to individual-level, de-identified longitudinal data on BC's residents.</p> <p>Statistics Canada: The federal government agency responsible for producing statistics on CAN's population, including on the economy, society and health.</p> <p>Statistics Canada Social Data Linkage Environment (SDLE): At Statistics Canada; supports the innovative use of administrative and survey data, e.g., record linkage, for research and to inform socio-economic policy.</p> <p>ICES Linked Data (ICES): A not-for-profit research institute that comprises a community of research, data and clinical experts and a secure and accessible array of health-related data in Ontario.</p>	10, 11, 12, 13

List of Sources:

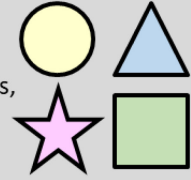
- [1] WHO. (2019). Canada. Retrieved from: <https://www.who.int/countries/can/en/>
- [2] Canada – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Marchildon, G. P. (2013). Canada health system review. *Health systems in transition*, 15(1). Retrieved from: http://www.euro.who.int/__data/assets/pdf_file/0011/181955/e96759.pdf
- [4] Eds. Mossialos, E., Djordjevic, A., Osborn, R., Sarnak, R. (2017). *International Profiles of Health Care Systems*. The Commonwealth Fund. Retrieved from: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf
- [5] OECD Statistics. (2019). *Health Expenditure and Financing*. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [6] PricewaterhouseCoopers. (2018). *Funding for value*. Retrieved from: <https://www.pwc.com.au/publications/pdf/funding-thought-leadership-18apr18.pdf>
- [7] ICES. (2019). *ICES Data*. Retrieved from: <https://www.ices.on.ca/Data-and-Privacy/ICES-data>
- [8] Office of the Privacy Commissioner of Canada. (2019). *Provincial and territorial privacy laws and oversight*. Retrieved at: <https://www.priv.gc.ca/en/about-the-opc/what-we-do/provincial-and-territorial-collaboration/provincial-and-territorial-privacy-laws-and-oversight/>
- [9] OECD. (2015). *Health data governance: Privacy, monitoring and research*. OECD Health policy studies, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en>
- [10] Ishiguro, L., Saskin, R., Vermeulen, M. J., Yates, E., Gunraj, N. & Victor, J. C. (2016). Increasing access to health administrative data with ICES Data & Analytic Services. *Healthcare Quarterly*, 19(1).
- [11] Population Data BC. (2019). *About PopData*. Retrieved from: <https://www.popdata.bc.ca/aboutus>
- [12] ICES. (2019). *Mission, vision & values*. Retrieved from: <https://www.ices.on.ca/About-ICES/Mission-vision-and-values>
- [13] Statistics Canada. (2017). *Social Data Linkage Environment*. Retrieved from: <https://www.statcan.gc.ca/eng/sdle/index>

SDLE

SDLE facilitates research across Canada by: filling data gaps related financial, social, economic and general activities and conditions; reducing the burden of having to collect new data for different projects; and improving efficiency in data linkage and processing

Data content

The SDLE has 2 major components: the Derived Record Depository (DRD) and the Key Registry. The DRD is a national-level database with personal identifiers. Data content includes 1) tax files, 2) birth and death records, 3) social insurance records, 4) socio-demographic-economic information (including immigration). Linkable datasets include data about 5) individual health, 6) health care use, 7) health care encounters and ICD codes, 8) environmental data, and 9) provincial datasets



Data collection

Is the data structured?
Yes

Who collects the data?
The DRD database is sourced from Statistics Canada index files. Linked datasets may be collected by other agencies

Data linkage and cleaning

At what level is the data aggregated?
The SDLE links existing datasets, so the data is not aggregated until used by a researcher

Is the data checked or cleaned?
Statistics Canada researchers are responsible for minimizing false positives in data linkages

Is the data linkable?
Yes, the source index files are linked by both unique and non-unique personal identifiers

Data storage

Where is the data located?
A secure environment at Statistics Canada

How long is the data stored?
Data is stored indefinitely

Who owns the data?
Statistics Canada owns some of the datasets, with linked SDLE datasets from external research

Who performs the data linkage?
Statistics Canada employees perform linkage

Who has access to data?
Some employees cannot access the source data files, and have access to only the personal identifiers required for linkage. Other employees only have access to the source data files and cannot see personal identifiers

Research use

What is the process for acquiring access?
Researchers can apply to access already linked records through Statistics Canada's (StatsCan) Research Data Centers (RDCs) with the steps: (1) draft a project proposal, (2) complete an online form, (3) StatsCan evaluation of the proposal, (4) complete a security screening form, (5) sign a microdata research contract, and (6) submit the research output to StatsCan

How long does it take to become available?
It depends on the dataset

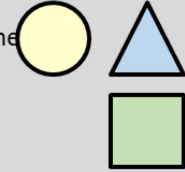
How is the data used for research purposes?
The SDLE has been used to link many surveys and datasets. StatsCan does not provide a full list of projects that have used the SDLE, but they do have a complete record of all projects that have been approved for microdata linkages

PopData BC

Facilitates research into the determinants of health, well-being, and development from an inter-disciplinary perspective

Data content

PopData BC manages 21 datasets from federal and provincial sources. External researchers also have the option of linking their own datasets to PopData BC. Datasets include Information about 1) health care usage, 2) health service usage, 3) vital statistics, 4) socio-economic-demographic data, 5) income data, 6) occupational safety data, 7) educational data



Data collection

Is the data structured?
Yes

Who collects the data?
Data is collected from federal and provincial sources



Data linkage and cleaning

At what level is the data aggregated?
Some of the data is aggregated (e.g., regional income data), but most stored at individual level

Is the data checked or cleaned?
Not for projects, but PopData BC validates and cleans new datasets to prepare them for linkage

Is the data linkable?
Yes, it is facilitated through a "population directory" table with personal identifiers



Data storage

Where is the data located?
A PopData BC Secure Research Environment (SRE), a central server accessible by an encrypted VPN firewall

How long is the data stored?
Archived in accordance with a project's protocol once the project is complete

Who owns the data?

The data providers, or trustees, who choose to provide their data to PopData BC to facilitate access to researchers

Who performs the data linkage?

PopData BC performs the linkage through a series of deterministic and probabilistic matching techniques. In cases where a match is uncertain, a PopData BC programmer will make final resolutions

Who has access to data?

Eligible public researchers based in Canada may apply to access specific datasets



Research use

What is the process for acquiring access?
Acquiring access follows a three-step process:
1. Submit data access and request form to PopData BC's Data Access Unit (DAU)
2. The DAU and the researcher work to achieve approval from the data stewards
3. Researchers sign agreements, undergo privacy training and set up their working environment

How long does it take to become available?
Generally, 4-10 months before year end, then it takes up to 3 months to link and validate



How is the data used for research purposes?
PopData BC does not have any research activity of its own – most completed by University researchers to publish in academic journals or inform policy development. Many data date back to 1985

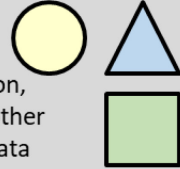
Manitoba Population Research Data Repository

(a project at the Manitoba Centre for Health Policy (MCHP) at the University of Manitoba)

MCHP conducts population-based health research with the end goal of informing health and social policy, resource allocation decisions, and improving health equity in Manitoba

Data content

Data in the repository comes from a variety of federal and provincial government department administrative datasets, including data on: 1) health care usage, 2) education, 3) justice, 4) social and family data, and 5) various provincial registries. There are also other datasets from regional authorities with 6) health services usage and 7) federal census data



Data collection

Is the data structured?
Yes

Who collects the data?
A variety of sources including federal and government agencies, local health authorities, and indigenous groups



Data linkage and cleaning

At what level is the data aggregated?
Data may be aggregated when sharing results with un-accredited analysts and researchers

Is the data checked or cleaned?
Yes, MCHP is responsible for conducting checks before datasets are made available

Is the data linkable?
Yes, it is linkable using crosswalk files



Data storage

Where is the data located?
The data is housed at MCHP facilities at the University of Manitoba

How long is the data stored?
7 years following the completion of a program

Who owns the data?
The data is not owned by the MCHP, but rather MCHP is able to access the data through a collection of data sharing agreements

Who performs the data linkage?
N/A

Who has access to data?
Accredited individuals who must renew accreditation every year. Once a project has been identified and approved, analysis can be completed either by an MCHP or external analyst



Research use

What is the process for acquiring access?
Acquiring access follows a three-step process:
1. Complete research proposal and associated forms,
2. Obtain research approval
3. Draw up a MCHP-researcher agreement and initiate project

How long does it take to become available?
Some of the records available on the data registry date back to 1970. All data are uploaded after year-end files have closed



How is the data used for research purposes?
Extensively used by the MCHP for deliverables and reports, having a significant impact on Manitoba's health policy. Over 500 papers have been published by MCHP and external researchers using this linked data

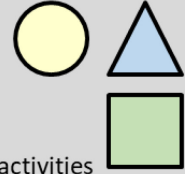
ICES

ICES linked data is a way to provide information about continuity of care in order to support health system policy, planning, and evaluation in Ontario

Data content

The ICES data repository houses information about 1) individual health care, 2) health care use, 3) socio-demographic-economic information, 4) registry data (e.g., cancer), with the assertion that the repository is “linkable across all data”.

ICES produces a yearly linkability report in order to demonstrate greater transparency to researchers who wish to know more about the quality and reliability of their linkage activities



Data collection

Is the data structured?
Yes

Who collects the data?
Various groups; ICES repositories hold data from a variety of different administrative levels

Data linkage and cleaning

At what level is the data aggregated?
Depends on the level of access given to researchers

Is the data checked or cleaned?
Statistics Canada researchers are responsible for minimizing false positives in data linkages

Is the data linkable?
Yes, each person is assigned an ICES key number, so their data linkable across the other datasets

Data storage

Where is the data located?
Servers at the Sunnybrook Health Sciences Center in Toronto within a closed-computing system

How long is the data stored?
Data is stored indefinitely

Who owns the data?
ICES is a “prescribed entity”, which grants it rights as a data custodian. Data ownership considered case-by-case

Who performs the data linkage?
ICES researchers manually implement the probabilistic matching methods used to link datasets, which allows for the possibility of human error in linkage efforts

Who has access to data?
ICES researchers, public sector researchers, who apply through ICES, and private sector researchers (only analytic services, not datasets)

Research use

What is the process for acquiring access?
Non-ICES researchers must apply for access and provide information about the intended use of the data. The datasets delivered to researchers are linked, with identifying information removed

How long does it take to become available?
Datasets are updated with varying frequencies and there may be a lag before a researcher is able to access a dataset due to costs and approval. Many of the datasets date back to 1991

How is the data used for research purposes?
Various projects have already been completed through internal ICES, private and public sector researchers. ICES researchers routinely publish reports using the linked datasets to inform health policy planning

New Zealand

New Zealand (NZL)

Population	4,885,500	1
Administrative structure of country	Parliamentary democracy under a constitutional monarchy	2
Jurisdiction of health policy	The national government plays a central role in determining the healthcare policy agenda, service requirements and budget. The government dominates all aspects of health care as the primary funder and supplier of healthcare; it also sets regulations and monitors compliance. District health boards are responsible for planning, purchasing, and providing health services and disability support for those over age 65.	3
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 9.1 % of GDP • 4,211.9 US\$ PPP 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	12.9 % (2018)	4
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing in NZL primarily comes from public, government funding, though PHI and OOP payments also play a role.</p> <p>National health care system: The national healthcare system in NZL is financed through general tax revenue.</p> <p>Voluntary, private health insurance (VHI): VHI is offered by both not-for and for-profit organizations and is financed by individuals. It is mainly purchased to cover cost-sharing, specialist fees and elective surgery in private hospitals (complementary). There is also supplementary coverage for faster access to non-urgent treatments.</p> <p>OOP payments: OOP costs include cost-sharing and other direct costs, such as copayments for general practitioner (GP) visits and pharmaceuticals. There are reduced fees after 12 doctor visits and no drug copayments after 20 prescriptions annually.</p>	3
Extent of coverage by insurance type	<p>Public health care system: NZL's public health care system provides universal health care for all permanent residents.</p> <p>VHI: About 33 % of the population purchases some form of VHI.</p>	3
Organization of health service delivery		
Primary care - gate-keeper function?	Yes	3
Specialist outpatient care	Most specialists are employed by district health boards (DHB) and work in public hospitals; however, special care is also provided in private clinics and hospitals.	3
Payment mechanisms	<p>Hospital care: There is a mix of private and public hospitals in NZL. Public hospitals receive global budgets from the DHBs (the owners), which is derived from past utilization patterns, projection of population needs and government goals. Budget includes staffing costs. A case-mix funding system is used to allocate budget within a DHB hospital.</p> <p>Primary care: GPs are mostly self-employed and independent. Half their income is based on capitation through a government subsidy paid via primary health organizations (PHOs); the other half from patient copayments for fee-for-service set by individual GPs. PHOs themselves also receive per-capita funding to improve access and promote health and care coordination. They may receive and provide additional funding to GPs who reach particular quality/performance targets for chronic care.</p> <p>Specialist care: Specialists employed by DHBs are salaried for work in public hospitals; when specialists work privately or treat patients in private hospitals payment is fee-for-service.</p>	3

Examples of health data protection laws	<ul style="list-style-type: none"> • Privacy Act 1993 • Privacy Act 2020 • Health Information Privacy Code 1994 • Health Information Privacy Code 2020 • Health Record Standard NZS 8153:2002 • Health (Retention of Health Information) Regulations 1996 • Public Records Act 2005 • Health Information Security Framework 2015 • Te Tiriti o Waitangi (Treaty of Waitangi) 	5, 6
Example of health data strategies	<ul style="list-style-type: none"> • NZ Digital Health Strategy 	7
Index of technical infrastructure	NZL has a strong health information system and reports great data sharing and accessibility (OECD's Key national dataset availability, maturity and use score: 5.35/7.0).	8
Relevant actors	<p>NZ Ministry of Health: leads NZL's health and disability system and is responsible for the management and development of the system, and for collecting health and disability information.</p> <p>Stats NZ Tatauranga Aotearoa (Stats NZ): the official data agency; operates independently of government and is responsible for the collection of statistics related to the economy, population and society of NZL. Stats NZ curates the Integrated Data Infrastructure (IDI), a large research database.</p>	3, 9

List of Sources:

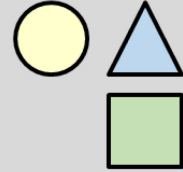
- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] New Zealand – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Gauld, R. (2017). The New Zealand health care system. In Mossialos, E., Djordjevic, A., Osborn, R. & Sarnak, D. (Eds.), International profiles of health care systems. The Commonwealth Fund. Retrieved from: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf
- [4] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [5] Ministry of Health. (2009). Data protection and privacy. New Zealand Government. Retrieved from: <https://www.health.govt.nz/nz-health-statistics/access-and-use/data-protection-and-privacy>
- [6] Waitangi Tribunal. (2019). The Treaty of Waitangi/Te Tiriti O Waitangi. Retrieved at: <https://waitangitribunal.govt.nz/treaty-of-waitangi/>
- [7] Ministry of Health. (2019). Digital health portfolio – February 2019. New Zealand Government. Retrieved from: <https://www.health.govt.nz/system/files/documents/pages/digital-health-portfolio-report-feb-2019.pdf>
- [8] OECD. (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en>
- [9] Stats NZ Tatauranga Aotearoa. (2019). About us. Retrieved from: <https://www.stats.govt.nz/about-us/>

IDI

The IDI is a large research database providing evidence-based insights into the society, economy, and effectiveness of government initiatives in New Zealand

Data content

The IDI contains micro data about people and households in New Zealand. Data is longitudinal, probabilistically linked, and de-identified. The data includes accident compensation, crime, education, health, medical, social welfare, tax data, and others



Data collection

Is the data structured?
Yes

Who collects the data?
Government agencies, Stats NZ surveys, and non-governmental organizations (NGOs)



Data linkage and cleaning

At what level is the data aggregated?
Many datasets in the IDI have national coverage

Is the data checked or cleaned?
Yes, both 'probabilistic' and 'deterministic' linking methods are used, assisted by linking software

Is the data linkable?
There is no common linking variable to match individuals across different data sources in the IDI



Data storage

Where is the data located?
Microdata from IDI can only be accessed in a secure Data Lab environment

How long is the data stored?
The IDI contains over 166 billion data points, and is growing. Some birth records date back to the 1840's

Who owns the data?

Stats NZ

Who performs the data linkage?

Stats NZ

Who has access to data?

Researchers can apply to access micro data for research, but Stats NZ only provides access to the datasets that are relevant to the research project



Research use

What is the process for acquiring access?
Researcher submits request -> Stats NZ decides on applications with inputs from internal teams -> If approved, researchers can access the secure data lab environment where they can perform and submit their analysis outputs for approval -> Only after final approval the data gets out of secure environment

How long does it take to become available?
Estimated to take up to 6 weeks



How is the data used for research purposes?

Used by analysts and researchers in the public sector, universities and NGOs to improve outcomes in New Zealand. For example, by using health and social data to produce evidence to improve health outcomes

The Netherlands

The Netherlands (NLD)

Population	17,231,017	1
Administrative structure of country	Parliamentary constitutional democracy	2
Jurisdiction of health policy	A tradition of self-regulation, private provision of services and financing via a system of social health insurance created a healthcare sector that is dominated by many mutually dependent actors with different backgrounds. The government still has an important role in health policy development and implementation, but there are many actors involved, with only limited opportunities for the government to act autonomously.	3, 4
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 10.2 % of GDP • 5,739.2 US\$PPP 	5
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	10.6 %	5
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing in NLD is mixed and includes government funding, statutory contributions and a combination of private and other sources.</p> <p>Statutory health insurance (SHI): SHI is financed through a nationally defined, income-related contribution, government grant for the insured below age 18, and community-rated insurance premiums. The NLD government partially finances SHI through subsidies from general taxation and reallocation of payroll levies among insurers. It is provided through private insurers.</p> <p>Voluntary, private health insurance (VHI): VHI provide statutory and non-statutory benefits and complementary VHI covers cost-sharing for medicines and services excluded from the statutory package, e.g., dental care, alternative medicine, eyeglasses, physiotherapy. Premiums for VHI are not regulated.</p> <p>OOP payments: Since 2016, every insured must pay an annual deductible for healthcare costs. Besides this, OOP costs include cost-sharing of some services.</p>	3, 4
Extent of coverage by insurance type	<p>SHI: In the publicly financed health insurance system there is near universal health coverage, with all residents and nonresidents who pay income tax mandated to purchase SHI (from private insurers). 99.8 % of population is insured.</p> <p>VHI: 84 % of the population purchases some combination of complementary VHI.</p>	3, 4
Organization of health service delivery		
Primary care - gatekeeper function?	Yes; referral by a GP is required for hospital and specialist care.	3, 4
Specialist outpatient care	Almost all specialists are hospital-based; there is a growing trend to work in multidisciplinary ambulatory centers. NLD has six types of institutions that provide specialist care: General hospitals; academic (university) hospitals; specialized hospitals; independent treatment centers (day care only); top clinical centers (both general hospital care and complex care); and trauma centers.	3, 4
Payment mechanisms	<p>Hospital care: Most hospitals are private, nonprofit and are paid through case-based per diem payments, negotiated with insurer over price, quality and volume, using the case-based diagnosis-treatment combination system (DBC). Fee-for-service is also used.</p> <p>Primary care: GPs mainly work independently or in a self-employed partnership. There are three paths for GP financing: (1) "Core primary services" are</p>	3, 4

The Netherlands

	<p>paid through capitation fees per registered patient and consultation fees for GPs and ambulatory mental health care; the Dutch Health Care Authority sets fees; (2) Programmatic, multidisciplinary care is funded for certain chronic diseases via bundled payments; prices are negotiated with insurers; (3) additional contracts with insurers based on pay-for-performance.</p> <p>Specialist care: Within hospitals, specialists are either in group practice, paid fee-for-service, or salaried. Specialist fees are freely negotiable as a part of hospital payment. Specialists in ambulatory centers are paid fee-for-service; the fee schedule is insurer negotiated.</p> <p>Other: Long-term residential care is paid through care intensity packages.</p> <p>Of note, insurers in NLD are expected to engage in strategic purchasing, and contracted providers are expected to compete on both quality and cost.</p>	
Examples of health data protection laws	<ul style="list-style-type: none"> • EU Regulation 2016/679 (General Data Protection Regulation) • Personal Data Protection Act 2000 (<i>wet Bescherming persoonsgegevens</i>) 	3, 6, 7
Example of health data strategies	N/A	
Index of technical infrastructure	NLD has limited health information systems, reporting higher availability and coding standardization and lower coverage and data linking for research/monitoring (OECD's Key national dataset availability, maturity and use score: 4.54/7.0).	8
Relevant actors	<p>Statistics Netherlands (CDS; Centraal Bureau voor de Statistiek): If government invests public funds in a dataset, Statistics Netherlands can access it. It conducts data linkage project using own data.</p> <p>Health Care TTP: This agency leads on other data linkage projects and provides access to de-identified data.</p> <p>Nivel Primary Care Database (Nivel PC Database; Nivel Zorgregistraties Eerste Lijn): uses routinely recorded data from healthcare providers to monitor health and usage of services in a population representative sample.</p> <p>PHARMO Database Network: seeks to gain greater insights of healthcare with the collection and analysis of patient-centric data. It provides information of more than 4 million (25 %) residents of a well-defined population in the Netherlands for an average of ten years, creating disease-specific cohorts which are linkable to perinatal registry and various healthcare databases.</p> <p>Other: Datasets that are not financed by the government belong to different owners and access to data can be approved on a project-by-project basis.</p>	2, 9, 10

List of sources:

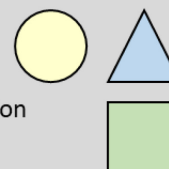
- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] The Netherlands – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Kroneman, M., Boerma, W., van den Berg, M., Groenewegen, P., de Jong, J., & van Ginneken, E. (2016). Netherlands health system review. *Health systems in transition*, 18(2).
- [4] Eds. Mossialos, E., Djordjevic, A., Osborn, R., Sarnak, R. (2017). *International Profiles of Health Care Systems*. The Commonwealth Fund. Retrieved from: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf
- [5] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [6] General Data Protection Regulation (GDPR). (2018). General Data Protection Regulation (GDPR) – Final text neatly arranged. Retrieved from: <https://gdpr-info.eu>
- [7] UCI – Donald Bren School of Information & Computer Sciences. (2019). Personal Data Protection Act. Unofficial Translation. Retrieved at: <https://www.ics.uci.edu/~kobsa/privacy/Netherlands.htm>
- [8] OECD. (2015a). Health data governance: Privacy, monitoring and research. *OECD Health policy studies*, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en>
- [9] Nivel, Research for better care. (2019). Nivel Primary Care Database. Retrieved from: <https://www.nivel.nl/en/nivel-zorgregistraties-eerste-lijn/nivel-primary-care-database>
- [10] PHARMO. (2019). PHARMO Database Network. Retrieved from: <https://pharmo.nl/what-we-have/pharmo-database-network/>

Nivel Primary Care Database

'Research for better care' is Nivel's mission statement and also for Nivel's Primary Care Database. This implies research on developments in public health, quality of care and health care utilization

Data content

Nivel Primary Care Database uses routinely recorded data, such as health problem and treatment data, to monitor health and use of health services among the Dutch population through a representative sample of up to 10% of the Dutch population



Data collection

Is the data structured?
Yes; no free text fields

Who collects the data?
Nivel collects the data. Recording of the data takes place in the primary care practices



Data linkage and cleaning

At what level is the data aggregated?
Could be at national, regional, practice, individual level and individual consultation

Is the data checked or cleaned?
Yes. Shared responsibility between Nivel and applicant

Is the data linkable?
Yes, through pseudonyms of a national identifier. Linkage with other databases is possible



Data storage

Where is the data located?
Data is stored in data warehouse located at Nivel

How long is the data stored?
Data are stored 'forever'. If the funding of the project would stop, data would be safely stored

Who owns the data?

No organization owns the data, but Nivel has access to it for research purposes

Who performs the data linkage?

A trusted third party performs pseudonymization. Nivel performs linkages between data from different disciplines. Linkage with external data sources is often via Statistics Netherlands

Who has access to data?

Access to data is dependent on conditions that have to be fulfilled



Research use

What is the process for acquiring access?
Researchers contact Nivel using a form available on the website. Application has to be approved according to governance code of Nivel Primary Care Database, in which umbrella organizations of primary care providers and a privacy committee play a central role

How long does it take to become available?
Depends on complexity of the application and possible privacy issues involved



How is the data used for research purposes?

Many research projects on health care use and population health status have been completed using Nivel data. Applicants can be Nivel researchers, but also researchers from other institutes

PHARMO Database Network

Rapid access to primary and secondary healthcare data for research purposes

Data content

The PHARMO database provides rich and detailed information of more than 4 million residents in the Netherlands (25%). It is based on patient information retrieved from 1) GPs, 2) hospitals, 3) in-outpatient pharmacies, 4) laboratories and 5) patient reported outcomes



Data collection

Is the data structured?
Varies; yes and no

Who collects the data?
PHARMO links different data sources from GPs, hospitals, registries etc.



Data linkage and cleaning

At what level is the data aggregated?
At the national and regional level

Is the data checked or cleaned?
Yes. It is the responsibility of STIZON, but is also done on project level by PHARMO

Is the data linkable?
Some data is linked using the unique social security number



Data storage

Where is the data located?
Data is stored on secured servers in the Netherlands

How long is the data stored?
N/A

Who owns the data?

Healthcare providers and patients

Who performs the data linkage?

STIZON

Who has access to data?

Data access is granted to PHARMO and INSZO by STIZON. Access can be requested by researchers from universities



Research use

What is the process for acquiring access?

1. Researchers have to complete a data request form which is available on the website of PHARMO
2. PHARMO request approval from the compliance committee
3. Upon approval, the data is readily available

How long does it take to become available?
A few weeks



How is the data used for research purposes?

PHARMO itself conducts analyses for insights into the performance of medicines. Pharmaco-epidemiological studies are conducted with different stakeholders, including the pharmaceutical industry

Austria (AUT)

Population	8,847,037	1
Administrative structure of country	Federal parliamentary republic	2
Jurisdiction of health policy	Responsibilities are shared between the federal and state levels. Some have been delegated to self-governing bodies. The federal level is primarily responsible for regulating social insurance and healthcare provision, except hospital care. Here, the federal level gives basic guidance while the states define the specifics of legislation and implementation.	3, 4, 5
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 10.4 % of GDP • 5,705.1 US\$PPP 	6
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	17.7 %	6
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing in AUT is mainly financed through a mix of income-related statutory health insurance (SHI) contributions, general taxation and private sources.</p> <p>SHI: Contributions are wage-based. Contributions for unemployed and pensioners are financed through general tax revenues; federal states cover contributions for those on social welfare benefits.</p> <p>Voluntary, private health insurance (VHI): VHI is offered by the SHI system and is mainly supplementary for (1) cover for extra amenities in the hospital sector and (2) cover for more provider choice and medicines in ambulatory care. Premiums are risk-based.</p> <p>OOP payments: direct payments, user chargers (cost-sharing) and informal payments. These have increased in the last decade in terms of absolute spending.</p>	3, 4
Extent of coverage by insurance type	<p>SHI: there is near universal health coverage, with 99.9 % of the population covered.</p> <p>VHI: More than 33 % of population is covered by some form of VHI.</p> <p>Uninsured: a limited number of people are not insured due to employment situations; they can register for VHI.</p>	3, 4
Organization of health service delivery		3, 4
Primary care - gate-keeper function	No	3, 4
Specialist outpatient care	Specialist outpatient care is provided by office-based doctors, outpatient clinics and in hospitals.	3, 4
Payment mechanisms	<p>Hospital care (inpatient): Payment for acute services are regulated by the national <i>Leistungsorientierte Krankenanstaltenfinanzierung</i> (LKF) framework by: (1) DRG-based payment system and (2) state-specific steering, which includes a fixed-budget. Public hospitals sometimes receive a per diem fee for patients with VHI. VHI also covers fee-for-service payments for hospital physicians on top of their salaries, with a share kept by the hospital. Some specialist departments are paid on a per diem basis.</p> <p>Hospital care (outpatient): Outpatient hospital services paid via state health funds (LGFs). 50 % of payment is based on a DRG-outpatient system; the other 50 % is covered by a fixed-budget component of the LKF system.</p> <p>Primary care: SHI GPs generate revenue by fee-for-service and capita-tions. Non-contracted GPs are paid fee-for-service, setting own fees.</p> <p>Specialist care: Ambulatory services are provided by GPs, specialist physicians, dentists and outpatient clinics, etc. Payment mechanisms include</p>	3, 4, 5, 7

Austria

	<p>fee-for-service and capitation, based on whether services are included in SHI catalogue of reimbursable services and kind of contract provider has with SHI. Otherwise, direct payment is required, reimbursable up to 80 %. PHI may cover the difference.</p> <p>Other: Long-term care institutions are financed by states and local authorities.</p>	
Examples of health data protection laws	<ul style="list-style-type: none"> • EU Regulation 2016/679 (General Data Protection Regulation) • The Austrian Data Protection Act • The Austrian Health Telematics Act • The Electronic Health File Act • The Hospitals and Sanatoriums Act • The Patient Charter • The Research Organization Act • The Privacy Amendment Act 	8, 9, 10, 11, 12
Example of health data strategies	<ul style="list-style-type: none"> • Austria Health Targets (since 2011) 	13
Index of technical infrastructure	N/A	
Relevant actors	<p>Austrian National Public Health Institute (GÖG; Gesundheit Österreich GmbH): Responsible for Austrian health information system (ÖGIS; Österreichisches Gesundheitsinformationssystem); the regional health information system (REGIS; Regionale Gesundheitsinformationssystem); and the health data warehouse (DIAG; Dokumentations- und Informationssystem für Analysen im Gesundheitswesen).</p> <p>Regional social insurance authorities: Submits collected data from healthcare providers to the main social insurance federation after recoding it.</p> <p>Main social insurance federation (HVB; Hauptverband der österreichischen Sozialversicherungsträger): Data received is pseudonymized by a designated unit; geographic and structural data is added to individual datasets. HVB submits the data to the Ministry of Labor, Social Affairs, Health and Consumer Protection (Sozialministerium; Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz).</p> <p>DIAG: Stores data from the Sozialministerium.</p> <p>ÖGIS: main instrument of health reporting and health planning.</p> <p>Electronic Health File (ELGA; Elektronische Gesundheitsakte): supports treatment with improved availability of information.</p> <p>REGIS: part of ÖGIS; provides regional health-related comparisons according to selected indicators.</p> <p>Statistics Austria (Statistik Austria): Provides additional data (e.g., basic demographic data) from various sub-sources.</p> <p>Federal Ministry for Digital and Economic Affairs (BMDW; Bundesministerium für Digitalisierung und Wirtschaftsstandort): Responsible for legal, business and technical guidance.</p> <p>The Austrian Federal Computing Center (BRZ; Bundesrechenzentrum): Provides assistance with software development and integration, as well as data security and protection.</p> <p>Federal Ministry of Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK; Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie): Provides distance and travel time data.</p>	3, 4, 5, 14, 15

List of sources:

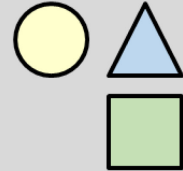
- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] Austria – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Bachner, F., Bobek, J., Habimana, K., Ladurner, J., Lepuschütz, L., Ostermann, H., Rainer, L., Schmidt, A. E., Zuba, M., Quentin, W. & Winkelmann, J. (2018). Austria health system review. Health systems in transition, 20(3). Retrieved from: http://www.euro.who.int/__data/assets/pdf_file/0009/382167/hit-austria-eng.pdf?ua=1
- [4] OECD/European Observatory on Health Systems and Policies. (2017). Austria: Country health profile 2017, State of health in the EU. OECD Publishing, Paris/European Observatory on Health Systems and Policies, Brussels. <http://dx.doi.org/10.1787/9789264283268-en>
- [5] Gesundheit.gv.at. (2019). Das Gesundheitswesen im Überblick. Retrieved from: <https://www.gesundheit.gv.at/gesundheitsleistungen/gesundheitswesen/gesundheitswesen>
- [6] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [7] Kobel, C. & Pfeiffer, K-P. (2011). Austria: Inpatient care and the LKF framework. In Busse, R., Geissler, A., Quentin, W., Wiley, M. M. Diagnosis-related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals. Retrieved from: http://eurodrp.projects.tu-berlin.de/publications/DRGbook/Ch11_Kobel.pdf
- [8] European Commission. (2018). EU data protection rules. Retrieved from: https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en
- [9] General Data Protection Regulation (GDPR). (2018). General Data Protection Regulation (GDPR) – Final text neatly arranged. Retrieved from: <https://gdpr-info.eu>
- [10] Gesundheit.gv.at. (2018). Patientencharta & Rechtsgrundlagen. Retrieved from: <https://www.gesundheit.gv.at/gesundheitsleistungen/patientenrechte/patientencharta>
- [11] Bundesministerium für Digitalisierung und Wirtschaftsstandort. (2019). Bundesgesetz über die Dokumentation im Gesundheitswesen. Rechtsinformationssystem des Bundes. Retrieved from: <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10011011>
- [12] Bundesministerium für Gesundheit. (2013). Wirkungsorientierte Folgenabschätzung. Verordnung des Bundesministeriums für Gesundheit zur ambulanten Dokumentation. Rechtsinformationssystem des Bundes. Retrieved from: https://www.ris.bka.gv.at/Dokumente/Begut/BEGUT_COO_2026_100_2_886724/COO_2026_100_2_892059.pdf
- [13] Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz. (2019). Austrian health targets. Retrieved from: <https://gesundheitsziele-oesterreich.at/english-summary/>
- [14] Gesundheit Österreich. (N.D.) Daten & Register. Retrieved from: https://goeg.at/index.php/Daten_und_Register
- [15] ELGA. (N.D). Gesetzliche Grundlagen von ELGA. Retrieved from: <https://www.elga.gv.at/faq/gesetzliche-grundlagen-von-elga/index.html>
- [16] Fülöp, G. (2017). Patientenorientierte Bedarfsplanung in Österreich. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen, 125: 60-69. <https://doi.org/10.1016/j.zefq.2017.07.003>

ELGA

ELGA supports medical treatment through improved availability of information

Data content

ELGA includes individual-level health information such as 1) medication data (for prescription as well as non-prescription medications relevant to interaction), 2) examination results, 3) laboratory tests, and 4) imaging output (further planned: patient discharge, precautionary powers of attorney, statutory medical registers)



Data collection

Is the data structured?
Yes

Who collects the data?
GPs, hospitals, care centers other medical specialists, dentists and pharmacies



Data linkage and cleaning

At what level is the data aggregated?
On demand when an individual or a health service provider accesses patient-level data

Is the data checked or cleaned?
No, there is no authority that checks the data for its accuracy and quality

Is the data linkable?
Yes, data is linked to central patient index (Zentraler Patientenindex)



Data storage

Where is the data located?
Data is kept decentralized in 13 ELGA storage units/areas

How long is the data stored?
The legal period for storing medical data is 30 years for inpatient data and 10 years for ambulatory data

Who owns the data?
Data ownership lies with the site of data generation (clinic, laboratory, etc.)

Who performs the data linkage?
N/A

Who has access to data?
Only the patient and the medical specialist who is currently treating the patient have access to the data. The authorization system (Berechtigungssystem) in the ELGA architecture checks every access inquiry of health data



Research use

What is the process for acquiring access?
Access is restricted to the patient and the health service provider that is currently treating the patient

How long does it take to become available?
Ambulatory physicians have access to the digital medical file for 28 days after treatment was registered, inpatient physicians have access for 28 days after hospital discharge, pharmacists can access information about medication up to two hours



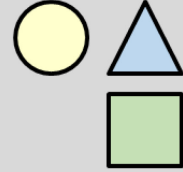
How is the data used for research purposes?
Research use under certain conditions is legally allowed. However the ELGA infrastructure is technically not able to provide data for research so far

ÖGIS

ÖGIS is the main instrument of health reporting and health planning

Data content

ÖGIS stores individual level health data such as 1) diagnosis and performance of Austrian hospitals ("LKF data"), 2) documentation for the entire outpatient area, 3) information on patient/mobile care of elderly patients, and 4) long-term care data. It also contains aggregated data like 5) causes of death, 6) cancer statistics, 7) hospital statistics 8) medical statistics, 9) basic demographic data and small-scale population forecasts, 10) accessibility data and 11) basic socio-economic data



Data collection

Is the data structured?
Yes

Who collects the data?
Gesundheit Österreich GmbH (GÖG) collects data from several sources and stores data



Data linkage and cleaning

At what level is the data aggregated?

Data is aggregated at the national, regional and community level

Is the data checked or cleaned?

Yes, the system software includes plausibility checks. Testing is performed quarterly

Is the data linkable?

Yes, data is linked to bPK (unique healthcare identifier)



Data storage

Where is the data located?
Data is stored at the DIAG

How long is the data stored?
The legal period for storing medical data is 30 years for inpatient data and 10 years for ambulatory data

Who owns the data?

Gesundheit Österreich GmbH (GÖG) as an operating body is responsible for data storage after receiving data from the various sub-sources. Pseudonymized data, performance and diagnostic data are linked in the Federal Ministry of Health and stored in the DIAG

Who performs the data linkage?

N/A

Who has access to data?

Access to data is granted after formal request and payment of fees



Research use

What is the process for acquiring access?

After sending a formal request, ÖGIS charges the researcher for its evaluation (at least 180 Euro) and carries it out as customized as possible according to the wishes of the client and taking into account the data protection requirements

How long does it take to become available?

ÖGIS is updated on yearly basis.
The time lag for data to be available is nearly one year



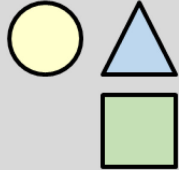
How is the data used for research purposes?

Data is used for health reporting and systems planning by national institutions, as well as European cooperations. Detailed information on individual requesting persons or institutions is not available

REGIS

REGIS is part of ÖGIS and provides regional health-related comparisons according to selected indicators

Data content
 REGIS provides aggregated data such as 1) population statistics, 2) causes of death statistics, 3) diagnosis and performance documentation of Austrian hospitals, 3) Austrian Health Survey 2006/07 and 4) road accident statistics



Data collection

Is the data structured?
Yes

Who collects the data?
Data presented in REGIS is collected by other institutions (e.g. Statistics Austria)

Data linkage and cleaning

At what level is the data aggregated?
Evaluations are available at two regional levels: 94 political districts and 32 supply regions

Is the data checked or cleaned?
The system software includes plausibility checks. Testing is performed quarterly

Is the data linkable?
Yes, data is linked to bPK (unique healthcare identifier)

Data storage


Where is the data located?
Data is stored at the DIAG

How long is the data stored?
The legal period for storing medical data is 30 years for inpatient data and 10 years for ambulatory data

Who owns the data?
Gesundheit Österreich GmbH (GÖG) as an operating body is responsible for data storage after receiving data from the various sub-sources. Pseudonymized data, performance and diagnostic data are linked in the Federal Ministry of Health and stored in the DIAG

Who performs the data linkage?
N/A


Who has access to data?
Anyone can access REGIS



Research use

What is the process for acquiring access?
Everyone can access respective representations under <https://regis.goeg.at/>. Data and images cannot be downloaded or used for further analysis

How long does it take to become available?
Data is updated irregularly, last queries were performed in autumn 2015. Individual information on available data years are provided with each respective graphic



How is the data used for research purposes?
Data is used for health reporting and systems planning by national institutions, as well as European cooperations. Detailed information on individual requesting persons or institutions is not available

e-prescription ("e-Rezept")

(pilot-phase April 2020)

e-prescription is an e-card application for social security administration system

Data content

Statutory health insurance prescriptions are stored as e-prescription. Focus is the reduction of paper-based processes by electronically recording, redeeming and settling cash receipts. In addition, the redeemed recipe fees will be charged daily to the recipe fee account (REGO) of the insured



Data collection

Is the data structured?

Yes

Who collects the data?

Every attending prescribing physician collects data



Data linkage and cleaning

At what level is the data aggregated?

Data is not aggregated

Is the data checked or cleaned?

Data is checked for completeness. Plausibility checks are not planned

Is the data linkable?

Yes, the e-card system processes sensitive, personal health data



Data storage

Where is the data located?

Data is stored in two e-card data centers and is pulled together on demand

How long is the data stored?

The legal period for storing medical data is 30 years for inpatient data and 10 years for ambulatory data

Who owns the data?

Data ownership lies with the site of data generation (clinic, laboratory, etc.)

Who performs the data linkage?

N/A

Who has access to data?

Every patient has access to their individual e-prescription data under www.meinesv.at. Patients grant access to attending physicians



Research use

What is the process for acquiring access?

Access is depending on two-factor principle. The e-card is the key card of the patient. This "key" gives the authorized physician access to the claim data. The admin card is the healthcare provider's entitlement card. Only with the e-card of the patient and the admin card can data be accessed in the e-card system

How long does it take to become available?

N/A



How is the data used for research purposes?

e-prescription starts its pilot phase in 2020 and nationwide roll-out is planned for 2022. So far e-prescription data is not used for research purposes yet

Sweden

Sweden (SWE)

Population	10,183,175	1
Administrative structure of country	Parliamentary constitutional monarchy	2
Jurisdiction of health policy	The healthcare system is organized into three levels: the national, regional and local. The Health and Medical Services Act of 1982 specifies that the responsibility for ensuring access to good health care is that of the county councils/regions and municipalities.	3
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 10.9 % of GDP • 5,551.9 US\$ PPP 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	13.9 %	4
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing in SWE is derived mainly from government funding and some private sources.</p> <p>National healthcare system: the national health insurance (NHI) is financed via general taxes raised by county councils, along with some national tax revenue. The system is delivered at the regional level.</p> <p>Voluntary, private health insurance (VHI): VHI is mainly supplementary coverage from employers for enhanced access to specialists and elective treatments.</p> <p>OOP payments: OOP costs are contained so a degree through caps on cost-sharing for health services (USD123) and drugs (USD246); most OOP spending goes to pharmaceuticals. County councils set copayment rates.</p>	3, 5
Extent of coverage by insurance type	<p>NHI: Within the national healthcare system, coverage is universal and automatic for all legal residents.</p> <p>VHI: Around 10 % of all employed individuals ages 15-74 have private insurance.</p>	3, 5
Organization of health service delivery		
Primary care - gatekeeper function?	No	3
Specialist outpatient care	Specialist outpatient care is provided by both public and private providers, including university hospitals, county council hospitals and in private clinics.	3, 5
Payment mechanisms	<p>Hospitals: Hospitals in SWE are mostly public, with some private for-profit and not-for-profit organizations. Payment mechanisms include global budgets and case-based payment, with limited pay-for-performance.</p> <p>Primary care: SWE's primary care is delivered by both public and private providers, who are paid through a mix of, largely, capitation and fee-for-service, with even more limited use of performance-related payments.</p> <p>Specialist care: Independent of service delivery location, outpatient specialist care is paid through fixed, prospective, per-case payments. These are based on DRGs. In addition, there is some use of price/ volume ceilings and quality components to compensation.</p>	3, 5
Examples of health data protection laws	<ul style="list-style-type: none"> • EU Regulation 2016/679 (GDPR; General Data Protection Regulation) • The Public Access to Information and Secrecy Act • The Personal Data Act • The Patient Data Act • The Data Protection Act 	6, 7
Example of health data strategies	National eHealth Strategy 2010	8

Sweden

Index of technical infrastructure	SWE has very high overall data availability, sharing and accessibility, but reports less use of using data to report on quality/performance or for research/monitoring purposes (OECD's Key national dataset availability, maturity and use score: 5.70/7.0).	9
Relevant actors	<p>National Board of Health and Welfare (Socialstyrelsen): central national authority for social services, public health, infectious diseases prevention and health services. It also maintains health data registers and official statistics.</p> <p>Swedish National Patient Portal (Journalen): the Swedish national patient accessible EHR system, which is accessible online via the national patient portal with an electronic identifier or Swedish personal identity number. Journalen was originally piloted as Sustains in the Uppsala region with an "internal health account",</p> <p>Registerforskning.se: operated by the Swedish Research Council, provides researchers with information on existing registers, as well as support during the process of register-based research, e.g., information on each part of the process of identifying, requesting and using register data.</p> <p>Register Utiliser Tool (RUT): The Swedish Research Council is developing a tool for registerforskning.se to facilitate and enable efficient searching and matching of information on metadata in registers.</p>	3, 5, 10, 11

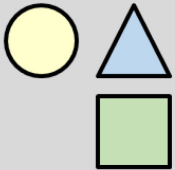
List of sources:

- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] Sweden – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Anell, A., Glengård, A. H. & Merkur, S. (2012). Sweden health system review. Health systems in transition, 14(5). Retrieved from: http://www.euro.who.int/__data/assets/pdf_file/0008/164096/e96455.pdf
- [4] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [5] Eds. Mossialos, E., Djordjevic, A., Osborn, R., Sarnak, R. (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Retrieved from: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf
- [6] Kirchberger, C. (2014). Overview of the national laws on electronic health records in the EU Member States. National report for Sweden. time.lex/Milieu Ltd. Retrieved from: https://ec.europa.eu/health/sites/health/files/chealth/docs/laws_sweden_en.pdf
- [7] General Data Protection Regulation (GDPR). (2018). General Data Protection Regulation (GDPR) – Final text neatly arranged. Retrieved from: <https://gdpr-info.eu>
- [8] International Society for Telemedicine & eHealth. Sweden National eHealth Strategy 2010. Retrieved from: https://www.isfteh.org/media/sweden_national_ehealth_strategy_2010
- [9] OECD. (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en>
- [10] Registerforskning.se. (2019). Easier to find register data with the Register Utiliser Tool (RUT). Retrieved from: <https://www.registerforskning.se/en/registers-in-sweden/easier-to-find-register-data-with-the-register-utiliser-tool/>
- [11] Armstrong, S. (2017). Patient access to health records: striving for the Swedish ideal. The BMJ, 357:j2069. Retrieved from: <https://www.bmj.com/content/357/bmj.j2069>

Journalen

Makes health data available for patients, increases quality of care and cost-effectiveness in the health care system

Data content
Journalen contains information on patients' medical records



Data collection

Is the data structured?
Yes

Who collects the data?
The health care provider

Data linkage and cleaning

At what level is the data aggregated?
N/A

Is the data checked or cleaned?
N/A

Is the data linkable?
No

Data storage


Where is the data located?
Patient records are kept electronically. Most GPs use electronic patient records for diagnostic data

How long is the data stored?
Data is not stored. Information is retrieved every time a provider or patient logs in

Who owns the data?
Each health care provider owns their respective journal information

Who performs the data linkage?
N/A


Who has access to data?
Health care providers and individual patients



Research use

What is the process for acquiring access?
When a health care provider connects to display journal information, a technical connection to a service platform occurs. This service platform allows the sending of structured data, regardless of journal systems, to display the same way for the provider

How long does it take to become available?
N/A



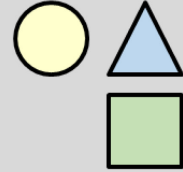
How is the data used for research purposes?
Data is not used for research purposes

Registerforskning / RUT

Provides researchers with information on existing registers. RUT is being developed to facilitate searching and matching of information among registers

Data content

Registerforskning includes several registers available through Swedish national public authorities with data covering for example 1) health care, 2) taxes, 3) residency, 4) civil status, 5) criminal convictions and 6) education. The tool RUT includes only metadata, without any micro-data



Data collection

Is the data structured?

Varies; RUT metadata structured

Who collects the data?

Government agencies or other organizations, using various methods



Data linkage and cleaning

At what level is the data aggregated?

Data is at individual level and is generally linked before being released in a de-identified format to the researchers

Is the data checked or cleaned?

N/A

Is the data linkable?

Yes, unique personal identity numbers allow for data to be linkable at the source



Data storage

Where is the data located?

Different register holders are responsible for different registers

How long is the data stored?

As a general rule it should be stored for at least 10 years after the conclusion of a project, but variable

Who owns the data?

Data ownership depends on the type of organization that collected the data and is administering it

Who performs the data linkage?

Normally the register holder

Who has access to data?

Different registry holders use different systems to make data available. The metadata in RUT is modelled by the Swedish Research Council in a collaborative process with the register holder before being published



Research use

What is the process for acquiring access?

Several regulations define how data from registers can be processed for research purposes and when researchers can obtain data from a register holder. To request register data, researchers need to contact the register holder responsible for the particular register of interest, with different processes to disclose data.

How long does it take to become available?

Most recent data is from 2017, 2018 and 2019



How is the data used for research purposes?

RUT helps researchers access information on existing registers, and offers support through the research process. Several publications have used register data. RUT supports researchers in how to link data from different sources

Republic of Korea

Republic of Korea (KOR)

Population	51,635,256	1
Administrative structure of country	Presidential republic	2
Jurisdiction of health policy	Central government and centralized, single insurer with six regional headquarters.	3
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 8.2 % of GDP • 3,406.3 US\$ PPP 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	30.2 %	4
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare in KOR is financed mainly by a social (national) health insurance system (NHIS), though other sources, including private, are also in play.</p> <p>NHIS: The NHIS is financed through employment-based social contributions, based on a uniform rate, or other measures of ability to pay. The poorest of the population do not pay contributions to NHIS and are managed through the Medical Aid Program, financed by the general revenue of the central and local governments but administered (including payments to providers) through the health insurance system.</p> <p>Voluntary, private health insurance (VHI): VHI is available to cover the co-payments for insured services and payments for uninsured services.</p> <p>OOP payments: OOP spending is high, including for cost-sharing (co-payments for covered services) and full payment for services excluded from the benefits package.</p>	3
Extent of coverage by insurance type	<p>NHIS: There is near universal health coverage through NHIS (97 % of population).</p> <p>Medical Aid Program: The Medical Aid Program covers remaining 3 % of the population.</p> <p>VHI: About 75 % of the population in KOR purchases VHI.</p>	3
Organization of health service delivery		
Primary care - gate-keeper function?	No	3
Specialist outpatient care	All general hospitals provide outpatient specialist services; many physicians have specialist certifications and practice from their offices as single operators.	3
Payment mechanisms	<p>Hospital care: Hospitals are non-for-profit; payment methods from NHIS insurance funds include fee-for-service, DRG-based prospective payments to acute care providers (for seven disease categories) and per-diem payments (for 17 disease categories) to long-term care hospitals. Individual physicians are salaried; hospitals sometimes give financial incentive systems, based on number and profits.</p> <p>Primary care: Healthcare facilities are paid fee-for-service through health insurance funds; individual physicians are salaried. Fee-for-service is based on the Resource-based Relative Value (RBRVI) system; fees are set by insurers.</p> <p>Specialist care: Healthcare providers are paid fee-for-service, based on fees set by insurers based on the RBRVI system; individual physicians are salaried.</p>	3
Examples of health data protection laws	<ul style="list-style-type: none"> • The Basic Law on Health and Healthcare • The Bioethics and Safety Act 2005 • The Personal Information Protection Act (PIPA) 2011 	3, 5, 6

Republic of Korea

	<ul style="list-style-type: none"> • The Rare Disease Management Act 2015 • The Regulation on the Operation and the Management of the National Biobank of Korea 2013 • The Cancer Control Act 2003 	
Example of health data strategies	<ul style="list-style-type: none"> • The Regulatory “Sand Box” Initiative • EMR Certification Initiative • Big data initiatives such as information linkage between National Health Insurance agencies, Centers for Disease Control and Prevention and hospitals 	7
Index of technical infrastructure	KOR rates among the highest for health data availability, though less so for accessibility. Data is also not often used for research/monitoring purposes (OECD’s Key national dataset availability, maturity and use score: 5.92/7.0).	8
Relevant actors	<p>Ministry for Health and Welfare: The main tasks of this ministry are oversight of healthcare and quarantine, compulsory administration, pharmacist administration, health insurance, basic living insurance, welfare support, social security and social service policies, and population policy to cope with low birth rate, aging and child welfare.</p> <p>Ministry of Economy and Finance: This ministry oversees the financial policies of KOR, including taxation system, management of public institutions and planning and coordinating the mid- and long-term economic development goals of the country.</p> <p>National Health Insurance (NHIS) and National Health Insurance Database (NHIDB): NHIS includes data from all three types of health insurances in the Republic of Korea – NHIS, Medical Aid Program, and long-term care as well as from several sub-databases.</p> <p>Health Insurance Review and Agency (HIRA): The HIRA database contains medical claims data for the entire Korean population as a result of NHIS.</p>	3, 5, 7

List of sources

- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] South Korea – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Kwon, S., Lee, T. & Kim, C. (2015). Republic of Korea health system review. Asia Pacific observatory on public health systems and policies, Health systems in transition series, 5 (4). Retrieved from: https://apps.who.int/iris/bitstream/handle/10665/208215/9789290617105_eng.pdf?sequence=1&isAllowed=y
- [4] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [5] Republic of South Korea. (2011). Personal Information Protection Act. Retrieved from: <http://koreanlii.or.kr/w/images/0/0e/KoreanDPAct2011.pdf>
- [6] Lee, D. Park, M., Chang, S., & Haksoo K. (2019). Protecting and Utilizing Health and Medical Big Data: Policy Perspectives from Korea. Healthcare Informatics Research 2019; 25(4). Retrieved from: <https://synapse.koreamed.org/articles/1136692>
- [7] Seong, S., Kim, Y., Park, S, Khan, Y., Kim, H., Park, J., Kang, H., Do, C., Song, J., Lee, E., Ha, S., Shin, S., & Jeong, S. (2017). Cohort profile: the National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea. BMJ Open 2017; 7: e016640. Retrieved from: <https://bmjopen.bmj.com/content/7/9/e016640>
- [8] OECD. (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Retrieved from: <https://doi.org/10.1787/9789264244566-en>

NHIS & NHIDB

A database which gathers beneficiary health and health care data

Data content

NHIS includes data from all three types of health insurances in the Republic of Korea – NHIS, Medical Aid Program, and LTIC. It includes data from several sub-databases. This includes data on: 1) patient sociodemographics, insurance, and health check-up Info, 2) diagnosis, 3) encounters, and 4) medical utilization, including prescriptions



Data collection

Is the data structured?
Yes

Who collects the data?
The NHIS itself collects the data, supervised by the Ministry of Health and Welfare



Data linkage and cleaning

At what level is the data aggregated?
At the regional and national level

Is the data checked or cleaned?
N/A

Is the data linkable?
Yes, using a unique ID on the NHID platform



Data storage

Where is the data located?
NHIS has two big data analyses centers for analysis and provision of public data, located in its headquarter Seoul office

How long is the data stored?
N/A

Who owns the data?

The NHI

Who performs the data linkage?

The NHI, on the NHID platform, can link NHIS data to HIRA data, government-owned national surveys and registries, and/or hospital lab data

Who has access to data?

NHIS data is available to those with academic or public policy affiliations



Research use

What is the process for acquiring access?

The NHI Corporation operates a review committee to review and decide on the provision and use of NHI data. There are 2 ways to access data:

1. Access raw data by visiting 1/2 research centers through customized DB provided in “data analysis room” in the NHIC;
2. Through sample cohort datasets in sharable format

How long does it take to become available?
N/A



How is the data used for research purposes?

Research purposes include analysis of adherence and persistence, prescribing patterns, healthcare utilization, burden of disease, outcomes and adverse events, as well as for policy evaluation

HIRA

The HIRA database is primarily a reimbursement claims database for providers within the Korean health system

Data content

The HIRA database contains medical claims data for the entire Korean population as a result of the NHI system. This includes: information on diagnosis and status of outpatients and inpatients; information related to drugs, e.g., name, dosage, prescription date and periods, and method of administration; information on the use of laboratory and imaging tests. There are also data on patient age, gender, diagnosis, and lists of prescribed medicines



Data collection

Is the data structured?
Yes

Who collects the data?
HIRA

Data linkage and cleaning

At what level is the data aggregated?
At the regional and national level

Is the data checked or cleaned?
N/A

Is the data linkable?
Yes, linkable to other data sources, including NHIS data, national surveys and registries, etc.

Data storage

Where is the data located?
In a data warehouse (DW)

How long is the data stored?
A 5-year period, after which claims are removed. Plans to expand data storage to 10 years are underway

Who owns the data?
HIRA and NHIS

Who performs the data linkage?
The NHIS

Who has access to data?
Raw data available for academics, government agencies, and the private sector. The scope of raw data provided is more limited to those in the private sector



Research use

What is the process for acquiring access?
There are 3 ways to access data:
1. Access raw data by visiting 1/7 research centers or remotely. Review process at HIRA required
2. HIRA offers 4 types of patient samples. No review process required
3. Summary statistics related to healthcare services (e.g., expenditures, utilization) are publically available via big data open system

How long does it take to become available?
N/A



How is the data used for research purposes?
Research purposes include analysis of adherence and persistence, prescribing patterns, healthcare utilization, burden of disease, outcomes and adverse events, as well as for policy evaluation

United Kingdom

United Kingdom (GBR)

Population	66,488,991	1
Administrative structure of country	Parliamentary constitutional monarchy	2
Jurisdiction of health policy	GBR is mainly a devolved matter, in that Scotland, Wales and Northern Ireland make their own decisions about the way in which health services are organized. The devolved administrations set health policy in these jurisdictions. Health policy for NHS England is decided by the United Kingdom government directly.	3, 4
Current health expenditure (CHE) and per capita health expenditure	<ul style="list-style-type: none"> • 10.2 % of GDP • 4,5000.1 US\$ PPP 	5
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	15.9 %	5
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>Healthcare financing in the UK is primarily derived from government funding, with some private sources.</p> <p>National Health Scheme (NHS): The national health system of GBR is financed mostly through general tax revenue, with a smaller portion coming from an automatic payroll tax for insurance contributions. NHS also receives income from copayments.</p> <p>Voluntary, private health insurance (PHI): Private insurance in GBR varies from coverage for specific conditions to broader packages. It is primarily employer-based private medical insurance, but individuals also purchase plans.</p> <p>OOP payments: OOP costs are mainly due to cost-sharing of prescriptions medicines and medical appliances or direct payments for services not covered by the NHS or for private treatment.</p>	3, 4
Extent of coverage by insurance type	<p>NHS: Coverage is universal through NHS for all legal residents.</p> <p>VHI: Approximately 11 % of the population has VHI for quicker access to services.</p>	3, 4
Organization of health service delivery		
Primary care - gatekeeper function?	Yes	3
Specialist outpatient care	Most outpatient specialist consultations are performed in hospitals; some take place in general practices.	3
Payment mechanisms	<p>Hospital care: Hospitals in GBR are mostly public and payment is made mostly through case-based payments. Budgets are used to cover mental health, education, and research and training; these also include physician and drug costs.</p> <p>Primary care: Primary care is delivered primarily via private GPs, with some NHS-owned practices. With the latter, physicians are salaried. With the former, care is paid through a mix of capitation for essential services, fee-for-service and pay-for-performance. An increasing number of GPs playing a substitute role in private practices are salaried as well.</p> <p>Specialist care: Most specialists are salaried at NHS hospitals. Hospitals are paid for outpatient consultations at nationally determined rates. GPs “with specialist interests” in private practice offer specialist consultations on a per-session or fee-for-service basis.</p>	3, 4
Examples of health data protection laws	<ul style="list-style-type: none"> • The Data Protection Act 2018 • The Health Service (Control of Patient Information) Regulations 2002 	6, 7

United Kingdom

<p>Example of health data strategies</p>	<ul style="list-style-type: none"> • The future of healthcare: our vision for digital, data and technology in health and care (2018) • NHS Digital Data and Information Strategy (England, 2016) • A Health and Biomedical Informatics Research Strategy (Scotland, 2015) • Joined up data for better decisions: A strategy for improving data access and analysis (Scotland, 2012) 	<p>8, 9, 10, 11, 12</p>
<p>Index of technical infrastructure</p>	<p>The UK overall has good health data availability and accessibility:</p> <p>England has high data availability but reports lower population coverage in data and use of data for system quality/performance reporting (OECD's Key national dataset availability, maturity and use score: 5.03/7.0).</p> <p>Scotland has strong health information systems and reports high use of data for quality/performance reporting and for research; population coverage, however, is limited as is accessibility of some data (OECD's Key national dataset availability, maturity and use score: 5.58/7.0).</p>	<p>13</p>
<p>Relevant actors</p>	<p>Department of Health and Social Care (DHSC): Stewards the system and sets the overall strategic direction for health, public health and adult social care in England.</p> <p>NHS Digital: Formerly, Health and Care Information Centre (HSCIC) established with new powers and responsibilities in relation to data in 2005, NHS Digital it is the national provider of information, data and IT systems for commissioners, analysts and clinicians in health and social care in England; it provides digital services for the NHS and social care.</p> <p>Health and Social Care Transparency Panel (HSCTP): Is the transparency board for the DHSC; gives advice on making data more transparent and available.</p> <p>Clinical Practice Research Datalink (CPRD): a research service supporting retrospective and prospective public health and clinical studies. CPRD is sponsored by the Medicines and Healthcare products Regulatory Agency and the National Institute for Health Research (NIHR), as part of the DHSC. It collects anonymised patient data from a network of GP practices across the UK, which are then linked to other health related data to provide a longitudinal, representative UK population health dataset.</p> <p>electronic Data Research and Innovation Service (eDRIS): a single point of contact associated with the Edinburgh International Data Facility in Scotland for those in the research community who want to link datasets on Scottish healthcare with other public sector administrative datasets</p> <p>Scottish Primary Care Information Resource (SPIRE): a service which enables information to be requested from GP practice records and collected centrally to produce statistics for Scotland.</p>	<p>3, 4, 14, 15, 16, 17</p>

List of sources

- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] United Kingdom – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Cylus J., Richardson, E., Findley, L., Longley, M., O’Neill, C. & Steel, D. (2015). United Kingdom health system review. Health systems in transition 17(5). Retrieved from: http://www.euro.who.int/__data/assets/pdf_file/0006/302001/UK-HiT.pdf?ua=1
- [4] Eds. Mossialos, E., Djordjevic, A., Osborn, R., Sarnak, R. (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Retrieved from: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf
- [5] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [6] legislation.gov.uk. (N.D.). The Health Service (Control of Patient Information) Regulations 2002. Retrieved from: <http://www.legislation.gov.uk/uksi/2002/1438/contents/made>
- [7] NHS Digital. (N.D.). Codes of practice for handling information in health and care. Retrieved from: <https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/codes-of-practice-for-handling-information-in-health-and-care>
- [8] Department of Health and Social Care. (2018). The future of healthcare: our vision for digital, data and technology in health and care. Retrieved from: <https://www.gov.uk/government/publications/the-future-of-healthcare-our-vision-for-digital-data-and-technology-in-health-and-care/the-future-of-healthcare-our-vision-for-digital-data-and-technology-in-health-and-care>
- [9] NHS Digital. (2018). NHS Digital data and information strategy. Retrieved from: <https://digital.nhs.uk/data-and-information/nhs-digital-data-and-information-strategy>
- [10] Scottish Government. (2015). A health and biomedical informatics research strategy for Scotland. Retrieved from: <https://www.gov.scot/publications/health-biomedical-informatics-research-strategy-scotland/>
- [11] Scottish Government. (2012a). Joined up data for better decisions: A strategy for improving data access and analysis. Retrieved from: <https://www.gov.scot/publications/joined-up-data-better-decisions-strategy-improving-data-access-analysis/>
- [12] Scottish Government. (2012b). Joined-up data for better decisions: Guiding principles for data linkage. Retrieved from: <https://www.gov.scot/publications/joined-up-data-better-decisions-guiding-principles-data-linkage/>
- [13] OECD. (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en>
- [14] UK Government. (2016). HSCIC changing its name to NHS Digital. Retrieved from: <https://www.gov.uk/government/news/hscic-changing-its-name-to-nhs-digital>
- [15] CPRD. (2019). Clinical Practice Research Datalink. Retrieved from: <https://www.cprd.com/>
- [16] Public Health Scotland. (2019). Data and intelligence- eDRIS. Retrieved from: <https://www.isdscotland.org/Products-and-Services/eDRIS>.
- [17] Scottish Primary Care Information Resource. (2019). NHISScotland is improving the way it uses information from GP patient records. Retrieved from: <https://spire.scot/>

CPRD

CPRD provides anonymized UK Electronic Health Records (EHRs) to researchers within academic, regulatory, and pharmaceutical organizations worldwide to support observational public health research

Data content

Primary care data includes files on the 1) Patient, 2) Practice, 3) Staff, 4) Consultation, 5) Clinical details, 6) Referrals, 7) Immunizations, 8) Tests, 9) Prescriptions and devices

Linked datasets includes 1) Hospital Episode Statistics (HES) Admitted Patient Care, 2) HES Outpatient data, 3) HES Accident & Emergency data, 4) HES Diagnostic Imaging dataset, 5) Death registration data, 6) Cancer data, 7) Mental health dataset, 8) Small area level data



Data collection

Is the data structured?

Yes

Who collects the data?

Primary care data is collected by GPs. Afterwards, software suppliers de-identify and pseudonymize data and submit it to CPRD



Data linkage and cleaning

At what level is the data aggregated?

N/A

Is the data checked or cleaned?

Yes, NHS Digital generates flags for linked data to indicate the validity of data and enable cleaning

Is the data linkable?

Yes, linkable by NHS number for English GP practices who have consented to participate



Data storage

Where is the data located?

CPRD. Following linkage, de-identified data flows directly from NHS Digital to CPRD

How long is the data stored?

Data collection began in 1987 and data continues to grow

Who owns the data?

N/A

Who performs the data linkage?

NHS Digital's Health and Social Care Information Centre (HSCIC) performs the data linkage

Who has access to data?

The CPRD provides access to patient level data for health research depending on approval by Independent Scientific Advisory Committee (ISAC). In general, only researchers carrying out public health studies can receive data



Research use

What is the process for acquiring access?

All requests for research use must be submitted to ISAC Secretariat using the Protocol Application Form

How long does it take to become available?

Data are collected by practices and usually uploaded to the CPRD secure servers on a monthly basis. The date of last data collection corresponds to the date of the last data upload from each practice. Monthly builds of the primary care dataset are made available for researchers



How is the data used for research purposes?

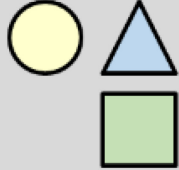
CPRD data are used worldwide for observational and EHR-enabled public health studies. Over 2,000 peer-reviewed publications informing clinical and drug safety guidelines and best practice have been published to date

eDRIS

eDRIS aims to support researchers and policy analysts by advising which health and administrative datasets are available, their data controllers and locations, strengths and weakness (in terms of content and data quality) and the processes required to gain access

Data content

Any publicly held data in Scotland can be requested for research, planning and evaluation projects. The Information Services Division (ISD) is a division of NHS National Services Scotland (NHS NSS)



Data collection

Is the data structured?
Yes

Who collects the data?
Different actors collect the data depending on the dataset

Data linkage and cleaning

At what level is the data aggregated?
N/A; varies by dataset

Is the data checked or cleaned?
N/A; varies by dataset

Is the data linkable?
Yes, data is linkable through the Community Health Index

Data storage


Where is the data located?
ISD holds most source datasets centrally

How long is the data stored?
Data is stored as per study protocol, then kept in the eDRIS archive for some time, after which it is destroyed

Who owns the data?
Shared between the NHS Boards and the ISD

Who performs the data linkage?
A trusted third party. A 'Population Spine' is used as an intermediary linkage tool and contains the personal identifiers of all individuals in Scotland who have been in contact with NHS Scotland


Who has access to data?
Approved applicants/researchers have to undergo information governance training and have the necessary approvals



Research use

What is the process for acquiring access?
eDRIS advises on process on a case-by-case basis. Some require applications and special permissions. Ethical approval is required unless exempted. Once all approvals are granted, the eDRIS Research Coordinator arranges for data to be extracted and linked

How long does it take to become available?
ISD holds information centrally that is updated on a monthly basis and provides NHS Boards with local extracts

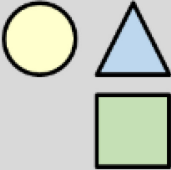


How is the data used for research purposes?
N/A

SPIRE

An approach to bring together primary care data from separate practices, which can be then linked to other datasets to better understand the health of the Scottish population in order to manage health and social care services

Data content
 SPIRE is not strictly a linkage initiative, but has the potential for linkage at the Information Services Division (ISD – see eDRIS).
 Data includes 1) Patient details, 2) Findings and procedures (Read V2 Codes), 3) Medicines, 4) Utilization, and 5) Diagnoses



Data collection

Is the data structured?
 Varies; yes and no. Only structured data are made available for research purposes

Who collects the data?
 GPs collect the data

Data linkage and cleaning

At what level is the data aggregated?
 Multiple levels, including GP, practice, cluster, locality and national

Is the data checked or cleaned?
 N/A

Is the data linkable?
 Yes, through the Community Health Index. Personally-identifiable data is pseudonymized before it leaves the practice

Data storage


Where is the data located?
 GPs store the data, and SPIRE uses software to extract data from each practice

How long is the data stored?
 Normally for six months after extraction or for the appropriate duration to process requests

Who owns the data?
 GPs

Who performs the data linkage?
 Software supplied by MSDi is used to extract data at each practice. Practices have information about the purpose and content of each data extract before deciding whether to opt in. Once extraction occurs, data is transferred securely to NSS using secure communication links


Who has access to data?
 NHS workers and researchers



Research use

What is the process for acquiring access?
 Local reports and queries are sent to the practice via eLinks and are run within the GP practice only. All requests for data extraction are scrutinized, agreed and prioritized by the SPIRE Strategy & Oversight Group. No data are extracted until the practice opts in

How long does it take to become available?
 An NSS analyst will contact the requestor for next steps after receipt of the form, normally within two weeks



How is the data used for research purposes?
 N/A

United States

United States (USA)

Population	327,167,434	1
Administrative structure of country	Constitutional federal republic	2
Jurisdiction of health policy	The health sector is characterized by relatively weak planning and regulatory systems and power is divided between the federal, state and local governments and a myriad of private organizations.	3
Current health expenditure (CHE) and per capita health expenditure	16.8 % of GDP (2016) 19,948.5 US\$ PPP (2016)	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	11.3 % (2016)	4
Type of health insurance and main sources of financing (taxation; statutory contributions; private/voluntary health insurance premiums; OOP; etc.)	<p>The USA is composed of a mixed financing system, including public and private insurers.</p> <p>Medicare: One main public insurer is Medicare, which covers individuals aged 65 and older and some disabled individuals. It is financed through a payroll tax, premiums and federal tax revenue.</p> <p>Medicaid: Another public insurer is Medicaid, which provides insurance to some low-income individuals. It is financed by state and federal tax revenue.</p> <p>Other public insurances in the USA: There are also additional smaller public insurers for military, veteran, and Native Americans.</p> <p>Commercial, private insurance: This is a primary form of insurance coverage in the US and is financed by employer and/or individual. For those without employment coverage, state-level exchanges provide income-based subsidies and incentives (e.g., tax breaks). Supplementary insurance for Medicare financed by individual also exists.</p> <p>OOP payments: Cost-sharing varies from plan to plan. Copayments are required for the most part for physician visits, hospital services and prescription drugs. There is a cap on OOP spending at USD6,600 annually for individuals with private insurance.</p>	3, 5
Extent of coverage by insurance type	<p>Public insurers: In total, public insurers in the USA cover 37.7 % of total insured population, with Medicare covering 17.2 % and Medicaid 19.3 % population.</p> <p>Private insurers: 67.2 % of total insured population is covered by private insurances. Of this, 56 % come from employer-based insurance and 16 % from direct-purchase coverage.</p> <p>Uninsured: 8.8 % of the population of the USA</p>	3, 5
Organization of health service delivery		
Primary care - gatekeeper function?	No, though some insurance programs require general practitioners (GPs) to play a gatekeeping role. Private plans use narrow networks of providers, however, which limit patient choice.	3, 5
Specialist outpatient care	Specialist outpatient care is provided in both hospitals and private practices.	5
Payment mechanisms	<p>Hospital care: Hospitals are mostly not-for-profit, with some public and for-profit facilities. Outpatient hospital care is paid on a fee-for-service basis.</p> <p>Primary care: Delivered by private GPs and other physicians. It is paid primarily via fee-for-service. Capitation is used in some private plans. Limited use of quality-based payments in the public and private sector. Further financial incentives are available (e.g., shared savings), and patients often pay copayments for doctor's visits depending on insurance plan.</p>	3

United States

	Specialist care: Specialists in both hospitals and private practices are mostly paid fee-for-service .	
Examples of health data protection laws	The Health Insurance Portability and Accountability Act (HIPAA)	6
Example of health data strategies	<ul style="list-style-type: none"> • The Health Data Initiative • The 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act • The Promoting Interoperability proposed rule 	6
Index of technical infrastructure	The USA has somewhat limited health data information systems, reporting limited accessibility and population coverage (only 13 % cover 80 % or more of the population) (OECD's Key national dataset availability, maturity and use score: 4.11/7.0).	7
Relevant actors	<p>US Department of Health and Human Services (HHS) and associated agencies (e.g., Centers for Medicare and Medicaid Services (CMS), Centers for Disease Control and Prevention (CDC), National Institutes of Health (NIH), US Food and Drug Administration (US FDA))</p> <p>Providers and commercial insurers also have a strong influence within the USA's health system</p>	3

List of sources

- [1] The World Bank. (2019). Total population. Retrieved from: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>
- [2] United States – The World Factbook. (2019). Retrieved from: <https://www.cia.gov/the-world-factbook/countries/>
- [3] Eds. Mossialos, E., Djordjevic, A., Osborn, R., Sarnak, R. (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Retrieved from: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf
- [4] OECD Statistics. (2019). Health Expenditure and Financing. Retrieved from: https://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT#
- [5] Berchick, E. R., Hood, E. & Barnett, J. C. (2018). Health insurance coverage in the United States: 2017. United States Census Bureau, report no. 60-264. Retrieved from: <https://www.census.gov/library/publications/2018/demo/p60-264.html>
- [6] U.S. Department of Health & Human Services. (N.D.). Laws & regulations. Retrieved from: <https://www.hhs.gov/regulations/index.html>
- [7] OECD. (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en>

CancerLinQ

A health IT platform that connects and analyzes real-world cancer care data from many data sources

Data content:

CancerLinQ operates three databases and includes data about the 1) Patient, 2) Practice / provider, 3) Encounters, 4) Diagnosis, 5) Pathology, 6) Staging, 7) Vitals, 8) Assessments, 9) Laboratory tests, 10) Imaging / radiology reports, 11) Treatment plans, 12) Medications ordered, 13) Medications administered, 14) Radiation therapy, 15) Surgery, 16) Outcomes, 17) Clinical trial enrollment



Data collection

Is the data structured?
Varies; yes and no

Who collects the data?
A community of practices that treat cancer patients and have a partnership with CancerLinQ



Data linkage and cleaning

At what level is the data aggregated?
Based on cancer type

Is the data checked or cleaned?
No, CancerLinQ Discovery (CLD) is provided "as is", with no guarantee for accuracy/completeness

Is the data linkable?
Only for subscribers, who can access identifiable patient information in the clinical database



Data storage

Where is the data located?
CancerLinQ hosts three databases – two for subscribers (identified and de-identified) and one for researchers (de-identified)

How long is the data stored?
N/A

Who owns the data?

Once collected, the data is owned by CancerLinQ. Patients must opt out if they do not want their information added to the database

Who performs the data linkage?
N/A

Who has access to data?
CancerLinQ subscribers. Non-commercially funded researchers who request access are limited to the data needed to answer the approved research question



Research use

What is the process for acquiring access?

1. Researcher submits request
2. Initial review of data sufficiency (if data is available in CancerLinQ and fit for purpose)
3. Full review of proposal from Research and Publications Committee
4. Decision on request made

How long does it take to become available?
6-8 weeks. Data access is terminated once a project is complete



How is the data used for research purposes?

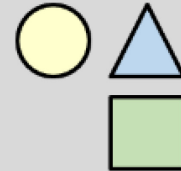
1. Custom analytics reports
2. Controlled, cloud-based access to curated, aggregated, de-identified data on specific diseases or patient populations

CMS Data

The Centers for Medicare & Medicaid (CMS) hosts over 400 databases with information related to the beneficiaries of and providers for CMS programs

Data content

CMS databases can contain socio-economic demographic, health and health care data. Many of these databases are available as public use files (PUFs) as aggregated data. Others are de-identified individual level data with a limited data set (LDS). The last category is research identifiable files (RIFs), which are customizable and contain individual level data



Data collection

Is the data structured?
Varies; yes and no

Who collects the data?
Varies based on the dataset; ultimately collected by CMS

Data linkage and cleaning

At what level is the data aggregated?
Public use files are nationally aggregated, individual level data are not aggregated

Is the data checked or cleaned?
N/A; varies depending on dataset

Is the data linkable?
Yes, CMS databases contain several linkable identifiers (e.g., patient, geography, provider)

Data storage


Where is the data located?
Varies based on database. CMS has central and local database administration

How long is the data stored?
Varies based on database

Who owns the data?
N/A; varies depending on dataset

Who performs the data linkage?
Generally the researcher performs the data linkage


Who has access to data?
Varies based on dataset. PUFs are available to anyone, other datasets with identifiable information require an application.



Research use

What is the process for acquiring access?
Many databases are available to anyone: data.cms.gov. For identifiable data, CMS has a contract with ResDAC to facilitate data access requests. Researchers submit application, which CMS privacy board reviews. After approval, researcher pays for data and CMS prepares it

How long does it take to become available?
Some PUF data are available relatively quickly (e.g., 20XX Medicare Beneficiary Summary file available in 20XX+1). ResDAC data requests take around 3-5 months



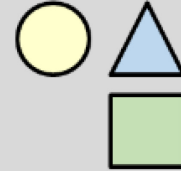
How is the data used for research purposes?
CMS data is widely used in health systems research to understand information about enrollment, payment, quality, program, and more for Medicare and Medicaid

Kaiser Health Connect

The Kaiser Health Connect system, built on an Epic product, integrates the clinical EMR with appointments, registration, and billing. Kaiser Permanente is the largest integrated delivery system and Kaiser Health Connect is the world's largest private clinical information system

Data content

Kaiser Health Connect includes two main items: MyChart (the **patient portal**) and EpicCare (**clinical data**). Clinical data collected includes 1) scheduling, 2) registration, 3) clinicals, and 4) billing data for both outpatient and inpatient care and 5) admission, discharge, and transfer, 6) pharmacy, 7) emergency department, and 8) operating room data for inpatient care



Data collection

Is the data structured?
Varies; yes and no

Who collects the data?
Kaiser Permanente staff and members



Data linkage and cleaning

At what level is the data aggregated?
18 instances of Epic and 8 regions aggregate to 1 fully integrated system

Is the data checked or cleaned?
Yes, some quality checks include profiling tools such as SAS Dataflux

Is the data linkable?
Yes, data is linkable using a patient identifier.



Data storage

Where is the data located?
Kaiser Permanente data warehouses

How long is the data stored?
Use of electronic health records began over 40 years ago; data storage continues

Who owns the data?

Kaiser Permanente staff and members

Who performs the data linkage?

Linkage is done automatically using the patient identifier. Clinicians can review a patient's chart from their home region, allowing a measure of linkage across all Kaiser regions to facilitate care delivery. Encounters outside of Kaiser can also be integrated into the patient record

Who has access to data?

Kaiser Permanente staff and members



Research use

What is the process for acquiring access?
The data is available for Kaiser Permanente Researchers and a subset of data is available for collaborating partners. A committee reviews requests from entities as collaborating partners, in areas of "mutual interest." Collaborating partners can access a subset of data to answer research questions

How long does it take to become available?
N/A



How is the data used for research purposes?

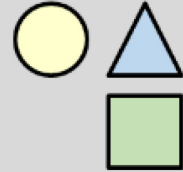
Kaiser Permanente Researchers can use the data to conduct studies. Otherwise, only organizations that require PHI for purposes of treatment or insurance can request information

PCORnet

PCORnet, the National Patient-Centered Clinical Research Network, supports clinical research through large amounts of health data and patient partnerships. Its “Front Door” provides an external online access point to researchers to use this data

Data content:

Data includes 1) demographic, 2) enrollment, 3) encounter, 4) diagnosis, 5) procedures, 6) vital signs, 7) prescriptions filled, 8) lab results, 9) condition, 10) responses to patient-reported outcome questionnaires, 11) provider orders for medication, 12) patients enrolled in PCORnet trials, 13) death, 14) medication administration by providers, 15) provider, 16) clinical trials data



Data collection

Is the data structured?
Varies; yes and no

Who collects the data?
Partner networks, including providers, patients, and payers. PCORnet uses a Common Data Model (CDM)



Data linkage and cleaning

At what level is the data aggregated?
At the national level based on the CDM

Is the data checked or cleaned?
Yes, by administrators. Foundational data quality is also checked through data curation

Is the data linkable?
Yes, the network uses a CDM without real identifiers so each site creates pseudo-identifiers



Data storage

Where is the data located?
Data remain local and secure behind the firewalls of each partner

How long is the data stored?
N/A; data storage responsibility of each partner

Who owns the data?

PCORnet uses a Distributed Research Network (DRN). Each partner network securely collects and stores data in a standardized way within their own institutions. The Patient-Centered Outcomes Research Institute oversees PCORnet

Who performs the data linkage?

The partner networks enter data into a CDM so the linkage will work

Who has access to data?

Each partner network



Research use

What is the process for acquiring access?
When a researcher submits a research question through an online access point called the “Front Door,” that question is reviewed by the Coordinating Center, which taps data of the individual Partner Networks through a specialized query format. A response is generated, allowing the researcher to ask the same question to many all at the same time

How long does it take to become available?
About 12-15 weeks



How is the data used for research purposes?

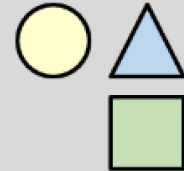
Potential investigators, patient groups, healthcare organizations, clinicians, government, industry scientists, and sponsors can use PCORnet for the conduct of more efficient clinical research

SEER-Medicare

SEER-Medicare links two datasets: the SEER dataset of the National Cancer Institute (NCI) and Medicare data to provide detailed information about Medicare recipients with cancer

Data content

The SEER dataset has data from 18 registries across the US for persons with cancer and Medicare data contains enrollment and claims data for Medicare recipients (population aged 65 and older). Combined, the two linked datasets include 1) demographic information, 2) information about the cancer, 3) Medicare enrollment data, and 4) data derived from bills submitted by providers and processed by Medicare



Data collection

Is the data structured?
Yes

Who collects the data?
CMS collects Medicare data, 18 cancer registries collect cancer data



Data linkage and cleaning

At what level is the data aggregated?
The data is aggregated from the cancer registries in defined geographic areas

Is the data checked or cleaned?
Yes. Quality is maintained through contractual agreements with regional registries and editing

Is the data linkable?
Yes, using unique personal identifiers



Data storage

Where is the data located?
SEER registries and CMS. Once the data is requested for use, a contractor creates the data files

How long is the data stored?
Earliest data available from 1973. Standard data retention time period for researchers is 5 years

Who owns the data?
SEER registry PIs and CMS. They hold NCI responsible for tracking the use and location of all released SEER-Medicare datasets

Who performs the data linkage?
IMS, a programming contractor of NCI

Who has access to data?
Release of SEER-Medicare data is project specific. Researchers must have an approved application in order to access these files



Research use

What is the process for acquiring access?
Complete an application, submit a data use agreement, receive Institutional Review Board approval, and request restricted variables. The researcher must pay costs based on data request

How long does it take to become available?
About 7-15 weeks, depending on variables. Linkages are done every 2 years, most recently in 2017, with cancer incidence data through 2015 and Medicare claims data through 2016



How is the data used for research purposes?
It is used for an array of epidemiological and health services research about cancer care, including costs of care and use of tests and treatment

Bisher erschienene Bände der Schriftenreihe

Working papers in health policy and management

ISSN 2197-8123 (print)

ISSN 1869-6287 (online)

Band 01

Sundmacher, Leonie; Scheller-Kreinsen, David:

Revisiting health inequalities in Germany. - 2009

ISBN (online) 978-3-7983-2167-0

Band 02

Sundmacher, Leonie; Jones, Andrew; Price, Nigel:

The role of health shocks in quitting smoking. - 2009

ISBN (online) 978-3-7983-2182-3

Band 03

Street, Andrew; Scheller-Kreinsen, David; Geissler, Alexander; Busse, Reinhard:

Determinants of hospital costs and performance variation: Methods, models and variables for the EuroDRG project. - 2010

ISBN (online) 978-3-7983-2124-3

Band 04

Ginneken, Ewout van:

Implications of future EU policy on the provision of medicines and on actors in the European pharmaceutical sector. - 2010

ISBN (online) 978-3-7983-2235-6

Band 05

Weid, Sabrina; Geissler, Alexander:

Möglichkeiten leistungsorientierter Vergütung von nicht angestellten Ärzten im stationären Sektor. Eine Prozesskostenrechnung am Beispiel von Varizenpatienten. - 2011

ISBN (online) 978-3-7983-2279-0

Band 06

Zentner, Annette; Velasco Garrido, Marcel; Busse, Reinhard:

Aktuelle internationale Initiativen der evidenzbasierten Bewertung von Arzneimitteln. - 2011

ISBN (online) 978-3-7983-2297-4

ISBN (print) 978-3-7983-2298-1

Band 07

Biermann, Alexandra; Geissler, Alexander:

Beatmungsfälle und Beatmungsdauer in deutschen Krankenhäusern. Analyse von DRG-Anreizen und Entwicklungen in der Beatmungsmedizin. - 2013

ISBN (online) 978-3-7983-2631-6

ISBN (print) 978-3-7983-2630-9

Band 08

Grosinska, Daria:

Transposition of the Patients' Rights Directive 2011/24/EU. A discourse analysis in Germany, Poland and Austria. - 2014

ISBN (online) 978-3-7983-2674-3

ISBN (print) 978-3-7983-2673-6

Band 09

Baier, Natalie; Struckmann, Verena:

Review zu den Perspektiven der ländlichen Versorgung. Ein Überblick international bestehender Ansätze. - 2014

ISBN (online) 978-3-7983-2719-1

ISBN (print) 978-3-7983-2718-4

Band 10

Stephani, Victor; Quentin, Wilm; Opoku, Daniel:

Does mHealth contribute to improved care for people with non-communicable diseases? A systematic review of randomized controlled trials. - 2015

ISBN (online) 978-3-7983-2753-5

ISBN (print) 978-3-7983-2752-8

Band 11

Busse, Reinhard; Panteli, Dimitra; Henschke, Cornelia:

Arzneimittelversorgung in der GKV und 15 anderen europäischen Gesundheitssystemen. Ein systematischer Vergleich. - 2015

ISBN (online) 978-3-7983-2767-2

ISBN (print) 978-3-7983-2766-5

Band 12

Busse, Reinhard; Eckhardt, Helene; Grössmann, Nicole; Wild, Claudia; Fuchs, Sabine:

PET/PET-CT Evidenz zum Bedarf und zur Planung in Deutschland und Österreich: Update 2018. - 2019

ISBN (online) 978-3-7983-3039-9

ISBN (print) 978-3-7983- 3038-2

Band 13

Victor Stephani

Effective and needed, but not used: Why do mobile phone-based health interventions in Africa not move beyond the project status? – 2019

ISBN (online) 978-3-7983-3053-5

ISBN (print) 978-3-7983- 3052-8

Internationale Datengrundlagen für die Versorgungsforschung – Impulse für Deutschland

Eine unabdingbare Voraussetzung für die Beantwortung der Forschungsfragen der Versorgungsforschung sind das Vorhandensein und die Zugänglichkeit von umfassenden Daten hoher Qualität. Es ist davon auszugehen, dass die Versorgungsforschung in den kommenden Jahren auf der Verlinkung (Datenlinkage, „data linkage“), also der Verknüpfung mehrerer Datenquellen auf Basis von geeigneten gemeinsamen Schlüsselvariablen beruhen wird. In Deutschland werden umfassende Daten hoher Qualität routinemäßig gesammelt, allerdings sind diese für Forschungszwecke nur bedingt geeignet und eingeschränkt verfügbar. Daher liegt der Ursprung dieses Gutachtens in den Grundfragen, welche Fragestellungen der Versorgungsforschung im Ausland auf Basis besserer Daten besser beantwortet werden können, und auf welchen Wegen eine solche Verbesserung der Datengrundlage in Deutschland erzielt werden könnte.

Fallbeispiele interessanter Verlinkungsansätze aus zwölf Ländern wurden identifiziert und anhand eines vordefinierten Konzeptes analysiert. Informationen wurden für alle Vergleichsländer bzw. Fallstudien zunächst auf Basis von öffentlich zugänglichen Quellen identifiziert. Anschließend wurde pro Land ein Profil erstellt und geeigneten Länderexperten zur Validierung und Ergänzung weitergeleitet.

ISBN 978-3-7983-3217-1 (print)

ISBN 978-3-7983-3218-8 (online)



ISBN 978-3-7983-3217-1



<https://verlag.tu-berlin.de>