

Health and Care Data

Approaches to data linkage for evidence-informed policy

Dimitra Panteli, Katherine Polin,
Erin Webb, Sara Allin, Andrew Barnes,
Alexander Degelsegger-Márquez,
Saira Ghafur, Margaret Jamieson,
Yoon Kim, Yulia Litvinova,
Ulrike Nimptsch, Maari Parkkinen,
Trine Aagren Rasmussen,

Christoph Reichebner, Julia Röttger,
Juliet Rumball-Smith, Giada Scarpetti,
Anna Lene Seidler, Johanna Seppänen,
Merran Smith, Morgan Snell, Dalibor
Stanimirovic, Robert Verheij, Metka
Zaletel, Reinhard Busse

Katherine Polin (Editor), Dimitra Panteli (Editor) and Erin Webb (Editor)
were responsible for this HiT

Editorial Board

Series editors

Reinhard Busse, Berlin University of Technology, Germany
Josep Figueras, European Observatory on Health Systems and Policies
Martin McKee, London School of Hygiene & Tropical Medicine, United Kingdom
Elias Mossialos, London School of Economics and Political Science, United Kingdom
Ewout van Ginneken, European Observatory on Health Systems and Policies

Series coordinator

Anna Maresso, European Observatory on Health Systems and Policies

Editorial team

Jonathan Cylus, European Observatory on Health Systems and Policies
Cristina Hernández-Quevedo, European Observatory on Health Systems and Policies
Katherine Polin, Berlin University of Technology, Germany
Marina Karanikolos, European Observatory on Health Systems and Policies
Sherry Merkur, European Observatory on Health Systems and Policies
Dimitra Panteli, European Observatory on Health Systems and Policies
Wilm Quentin, Berlin University of Technology, Germany
Bernd Rechel, European Observatory on Health Systems and Policies
Erica Richardson, European Observatory on Health Systems and Policies
Anna Sagan, European Observatory on Health Systems and Policies
Giada Scarpetti, Berlin University of Technology, Germany
Nathan Shuftan, Berlin University of Technology, Germany
Anna Maresso, European Observatory on Health Systems and Policies

International advisory board

Tit Albreht, Institute of Public Health, Slovenia
Carlos Alvarez-Dardet Díaz, University of Alicante, Spain
Rifat Atun, Harvard University, United States
Armin Fidler, Management Center Innsbruck
Colleen Flood, University of Toronto, Canada
Péter Gaál, Semmelweis University, Hungary
Unto Häkkinen, National Institute for Health and Welfare, Finland
William Hsiao, Harvard University, United States
Allan Krasnik, University of Copenhagen, Denmark
Joseph Kutzin, World Health Organization
Soonman Kwon, Seoul National University, Republic of Korea
John Lavis, McMaster University, Canada
Vivien Lin, La Trobe University, Australia
Greg Marchildon, University of Regina, Canada
Nata Menabde, World Health Organization
Charles Normand, University of Dublin, Ireland
Robin Osborn, The Commonwealth Fund, United States
Dominique Polton, National Health Insurance Fund for Salaried Staff (CNAMTS), France
Sophia Schlette, Federal Statutory Health Insurance Physicians Association, Germany
Igor Sheiman, Higher School of Economics, Russian Federation
Peter C. Smith, Imperial College, United Kingdom
Wynand P.M.M. van de Ven, Erasmus University, The Netherlands
Witold Zatonski, Marie Skłodowska-Curie Memorial Cancer Centre, Poland

Health Systems in Transition

Health and Care **Data:** Approaches to data linkage for evidence-informed policy

Zi Zentralinstitut
kassenärztliche
Versorgung



The European Observatory on Health Systems and Policies supports and promotes evidence-based health policy-making through comprehensive and rigorous analysis of health systems in Europe. It brings together a wide range of policy-makers, academics and practitioners to analyse trends in health reform, drawing on experience from across Europe to illuminate policy issues.

The Observatory is a partnership, hosted by WHO/Europe, which includes other international organizations (the European Commission); national and regional governments (Austria, Belgium, Finland, Ireland, the Kingdom of the Netherlands, Norway, Slovenia, Spain, Sweden, Switzerland, the United Kingdom and the Veneto Region of Italy (with Agenas)); other health system organizations (the French National Union of Health Insurance Funds (UNCAM), the Health Foundation); and academia (the London School of Economics and Political Science (LSE) and the London School of Hygiene & Tropical Medicine (LSHTM)). The Observatory has a secretariat in Brussels and it has hubs in London (at LSE and LSHTM) and at the Berlin University of Technology.

KEYWORDS:

DELIVERY OF HEALTH CARE

FINANCING HEALTH

HEALTH CARE REFORM

HEALTH AND CARE DATA

© World Health Organization 2023 (acting as the host organization for, and secretariat of, the European Observatory on Health Systems and Policies).

All rights reserved. The European Observatory on Health Systems and Policies welcomes requests for permission to reproduce or translate its publications, in part or in full.

Please address requests about the publication to:

**Publications,
WHO Regional Office for Europe,
UN City,
Marmorvej 51,
DK-2100 Copenhagen Ø, Denmark**

Alternatively, complete an online request form for documentation, health information, or for permission to quote or translate, on the Regional Office website (<https://www.who.int/about/policies/publishing/permissions>).

The views expressed by authors or editors do not necessarily represent the decisions or the stated policies of the European Observatory on Health Systems and Policies or any of its partners.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the European Observatory on Health Systems and Policies or any of its partners concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Where the designation "country or area" appears in the headings of tables, it covers countries, territories, cities, or areas. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the European Observatory on Health Systems and Policies in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

The European Observatory on Health Systems and Policies does not warrant that the information contained in this publication is complete and correct and shall not be liable for any damages incurred as a result of its use. Printed and bound in the United Kingdom.

Suggested citation:

Panteli D, Polin K, Webb E, Allin S, Barnes A, Degelsegger-Márquez A, Ghafur S, Jamieson M, Kim Y, Litvinova Y, Nimptsch U, Parkkinen M, Rasmussen TA, Reichebner C, Röttger J, Rumball-Smith J, Scarpetti G, Seidler AL, Seppänen J, Smith M, Snell M, Stanimirovic D, Verheij R, Zaletel M, Busse R. Health and Care Data: Approaches to data linkage for evidence-informed policy. *Health Systems in Transition*, 2023; 25(2): pp. i–213.

CONTENTS

Preface	vii
Acknowledgements	ix
List of abbreviations	xiii
List of key terms	xvii
List of tables, figures and boxes	xix
Abstract	xxi
Executive summary	xxiii
1 Introduction	1
1.1 Methodological approach	2
1.2 Overview of the review	9
2 Overview of included countries and case studies	11
2.1 Australia	11
2.2 Austria	13
2.3 Canada	15
2.4 Denmark	18
2.5 Finland	20
2.6 France	21
2.7 Netherlands (Kingdom of the)	23
2.8 New Zealand	24
2.9 Republic of Korea	26
2.10 Slovenia	27
2.11 Sweden	31
2.12 United Kingdom	32
2.13 United States	35
2.14 Summary of case studies presented in this chapter	38

3	Expanding the possibilities of routinely collected data	43
3.1	<i>Datasets that bring together data from different sectors to fulfill their primary purpose</i>	43
3.2	<i>Merging of routinely generated data into a new database for secondary use</i>	47
3.3	<i>Linking of independently created datasets</i>	48
4	Differences in the linking process	57
5	Access to health-related data for researchers	63
6	Considerations for electronic patient records	69
7	Considerations for privacy and governance of digital health data	73
7.1	<i>Rights to privacy and data protection regarding secondary use of digital health data</i>	74
7.2	<i>Risks to privacy and data security when using linked health data</i>	75
7.3	<i>Principles and regulatory considerations for digital health data protection and security</i>	76
7.4	<i>Overview of data governance approaches</i>	81
7.5	<i>The General Data Protection Regulation of the European Union</i>	84
8	International data sharing and the EU health data space	87
9	Data linkage in the COVID-19 pandemic	93
9.1	<i>New sources of data developed for the pandemic response</i>	94
9.2	<i>National examples of COVID-related data linkage activities and innovations</i>	97
9.3	<i>Building on existing databanks and data linkage platforms</i>	100
9.4	<i>Innovations in data access and sharing</i>	101
9.5	<i>Integration of COVID-19 data with electronic health records</i>	102
9.6	<i>Examples of international data linkage and sharing</i>	102
9.7	<i>Data protection and privacy considerations</i>	104
10	Conclusions	109
10.1	<i>What criteria should be met to enable meaningful healthcare research based on linked data?</i>	111
10.2	<i>How can coordinating bodies facilitating data linkage support policy-informing research?</i>	112
10.3	<i>How can health information from electronic records be leveraged?</i>	114
10.4	<i>Outlook and future work</i>	116

11	References	117
12	HiT methodology and production process	131

APPENDIX

I	Appendix I – Country profiles	135
	<i>AUSTRALIA</i>	136
	<i>AUSTRIA</i>	143
	<i>CANADA</i>	150
	<i>DENMARK</i>	157
	<i>FINLAND</i>	162
	<i>FRANCE</i>	166
	<i>NETHERLANDS (KINGDOM OF THE)</i>	171
	<i>NEW ZEALAND</i>	176
	<i>REPUBLIC OF KOREA</i>	180
	<i>SLOVENIA</i>	185
	<i>SWEDEN</i>	191
	<i>UNITED KINGDOM</i>	194
	<i>UNITED STATES OF AMERICA (US)</i>	200
II	Appendix II – Examples of privacy and data protection policies related to secondary use of health data in study countries	209

PREFACE

The Health Systems in Transition (HiT) series consists of two lines of studies:

- a) HiTs, which are country-based reviews that provide a detailed description of a health system and of reform and policy initiatives in progress or under development in a specific country. Each review is produced by country experts in collaboration with the Observatory's staff. In order to facilitate comparisons between countries, reviews are based on a template, which is revised periodically. The template provides detailed guidelines and specific questions, definitions and examples needed to compile a report; and
- b) special issues, which are comparative, cross-country studies on a specific topic of importance to policy-makers.

HiTs seek to provide relevant information to support policy-makers and analysts in the development of health systems in Europe. They are building blocks that can be used:

- to learn in detail about different approaches to the organization, financing and delivery of health services and the role of the main actors in health systems;
- to describe the institutional framework, the process, content and implementation of health care reform programmes;
- to highlight challenges and areas that require more in-depth analysis;
- to assist other researchers in more in-depth comparative health policy analysis; and
- to provide a tool for the dissemination of information on health systems and the exchange of experiences of reform strategies between policymakers and analysts in different countries;
- to draw out experiences in different countries and flag up the similarities and divergences between them.

Special issues build on existing knowledge from the country-based reviews; they synthesize and expand it using additional data sources, peer-reviewed and grey literature as well as the input of relevant country experts.

Compiling the HiT studies poses a number of methodological problems. In many countries, there is relatively little information available on the health system and the impact of reforms. Due to the lack of a uniform data source, quantitative data on health services are based on a number of different sources, including the World Health Organization (WHO) Regional Office for Europe's European Health for All database, data from national statistical offices, Eurostat, the Organisation for Economic Co-operation and Development (OECD) Health Data, data from the International Monetary Fund (IMF), the World Bank's World Development Indicators and any other relevant sources considered useful by the authors. Data collection methods and definitions sometimes vary but typically are consistent within each separate review.

A standardized review has certain disadvantages because the financing and delivery of health care differ across countries. However, it also offers advantages because it raises similar issues and questions. HiTs can be used to inform policy-makers about experiences in other countries that may be relevant to their own national situation. They can also be used to inform comparative analysis of health systems. This series is an ongoing initiative and material is updated at regular intervals.

Comments and suggestions for the further development and improvement of the HiT series are most welcome and can be sent to contact@obs.who.int.

HiTs and HiT summaries are available on the Observatory's website (www.healthobservatory.eu).

ACKNOWLEDGEMENTS

Editors: Katherine Polin, Dimitra Panteli, Erin Webb.

Authors: Dimitra Panteli, Katherine Polin, Erin Webb, Sara Allin, Andrew Barnes, Alexander Degelsegger-Márquez, Saira Ghafur, Margaret Jamieson, Yoon Kim, Yulia Litvinova, Ulrike Nimptsch, Maari Parkkinen, Trine Aagren Rasmussen, Christoph Reichebner, Julia Röttger, Juliet Rumball-Smith, Giada Scarpetti, Anna Lene Seidler, Johanna Seppänen, Merran Smith, Morgan Snell, Dalibor Stanimirovic, Robert Verheij, Metka Zaletel, Reinhard Busse.

Acknowledgements (in alphabetical order):

Emmanuel Bacry, Hans Okkels Birk, Tobias Dreischulte, Andina Feldman, Jochen Gensichen, Anna Gref, Anne-Marie Sigsgaard Hansen, Falk Hoffmann, Allan Krasnik, Lars Kroll, Bernd Mühlbauer, Allie Peckham, Georgios Raptis, Yeun Sook Rho, Charlotte von Rosenstiel, David Rudoler, Nanna Skovgaard, Melissa Spoden, Dominik von Stillfried, Louisa Stuewe, Maximilian Teitscheid, Michel Wensing, Imane Zidi, PHARMO Team, Stats NZ Microdata Customer Liaison Team.

This Health Systems in Transition (HiT) report on data linkage in 13 countries was produced by the European Observatory on Health Systems and Policies (the Observatory) with support from the Zentralinstitut für die kassenärztliche Versorgung (ZI) [German Central Institute for Statutory Health Insurance] and the Health Systems and Policy Monitor (HSPM) network.

The HSPM is an international network that works with the Observatory on Country Monitoring. It is made up of national counterparts that are highly regarded at national and international level and have particular strengths in the areas of health systems, health services, public health and health management research. They draw on their own extensive networks in the health field and their track record of successful collaboration with the Observatory to develop and update the HiT.

Writing for this edition was led by Dimitra Panteli (European Observatory on Health Systems and Policies), Katherine Polin (Technische Universität Berlin/European Observatory on Health Systems and Policies), Erin Webb

(Technische Universität Berlin/European Observatory on Health Systems and Policies), and Reinhard Busse (Technische Universität Berlin/European Observatory on Health Systems and Policies). It builds and expands on a report written by Dimitra Panteli, Julia Röttger (Technische Universität Berlin), Ulrike Nimptsch (Technische Universität Berlin), and Reinhard Busse (Technische Universität Berlin/European Observatory on Health Systems and Policies) in 2019, funded by the ZI.

This current report was edited by Katherine Polin, Dimitra Panteli and Erin Webb. Final responsibility for creating this document based on input from review contributors and other expert feedback rests with the editors.

The Observatory is extremely grateful to Sara Allin, Andrew Barnes, Alexander Degelsegger-Márquez, Saira Ghafur, Margaret Jamieson, Yoon Kim, Yulia Litvinova, Ulrike Nimptsch, Maari Parkkinen, Trine Aagren Rasmussen, Christoph Reichebner, Julia Röttger, Juliet Rumball-Smith, Giada Scarpetti, Anna Lene Seidler, Johanna Seppänen, L. Merran Smith, Morgan Snell, Dalibor Stanimirovic, Robert Verheij and Metka Zaletel for their contributions and review of the report.

Gratitude is also extended to Emmanuel Bacry, Hans Okkels Birk, Tobias Dreischulte, Andina Feldman, Jochen Gensichen, Anna Gref, Anne-Marie Sigsgaard Hansen, Falk Hoffmann, Allan Krasnik, Lars Kroll, Bernd Mühlbauer, Allie Peckham, Georgios Raptis, Yeun Sook Rho, Charlotte von Rosenstiel, David Rudoler, Nanna Skovgaard, Melissa Spoden, Dominik von Stillfried, Louisa Stuewe, Maximilian Teitscheid, Michel Wensing, Imane Zidi, PHARMO Team, Stats NZ Microdata Customer Liaison Team for their input.

The Observatory is a partnership, hosted by the WHO Regional Office for Europe, that includes the Governments of Austria, Belgium, Finland, Ireland, the Kingdom of the Netherlands, Norway, Slovenia, Spain, Sweden, Switzerland, the United Kingdom and the Veneto Region of Italy (with Agenas); the French National Union of Health Insurance Funds (UNCAM); the Health Foundation; the European Commission; the London School of Economics and Political Science (LSE); and the London School of Hygiene & Tropical Medicine (LSHTM). The Observatory is composed of a Steering Committee, core management team, research policy group and staff. Its Secretariat is based in Brussels, and it has offices in London at LSE and LSHTM and in Germany at the Berlin University of Technology. The Observatory team working on HiTs is led by Josep Figueras, Director; Elias Mossialos, Martin McKee, Reinhard Busse (Co-directors); Ewout van Ginneken and Suszy Lessof. The Country Monitoring Programme of the Observatory and the HiT series are coordinated by Anna Maresso. The production and copyediting process was coordinated by Jonathan North, with the support of Lucie Jackson, Sarah Cook (copy-editing) and Steven Still (typesetting).

LIST OF ABBREVIATIONS

Abbreviation	Name
AIHW	Australian Institute for Health and Welfare
AMDC*	Austrian Micro Data Centre
ASCO	American Society of Clinical Oncology (US)
CCGs	Clinical Commissioning Groups (NHS England)
CDM	Common Data Model (PCORnet US)
CDPCA	Contagious Disease Prevention and Control Act (Republic of Korea)
CépiDC	Centre for the Epidemiology of Causes of Death [<i>Centre d'épidémiologie des causes de décès</i>] (France)
CESREES	Ethical and Scientific Committee for Research, Studies and Evaluations in the Field of Health [<i>Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé</i>] (France)
CHESS	COVID-19 hospitalisation in England surveillance system
CMS	Centers for Medicare & Medicaid Services (US)
CNAMTS	National Health Insurance Fund for Salaried Workers [<i>Caisse Nationale d'Assurance Maladie des Travailleurs Salariés</i>] (France)
CNIL	National Commission for Data Protection and Liberties [<i>Commission nationale de l'informatique et des libertés</i>] (France)
CNSA	National Solidarity Fund for Autonomy [<i>Caisse nationale de solidarité pour l'autonomie</i>] (France)
CPES	Cancer Patient Experience Survey (UK)
CPR	Central Person Register [<i>Centrale Personregister</i>] (Denmark)
CPRD	Clinical Practice Research Datalink (UK)
CRPP	Central Registry of Patient Data [<i>Centralni register podatkov o pacientu</i>] (Slovenia)
CVDL	Centre for Victorian Data Linkage (Australia)
DALY	Disability-adjusted life years
DAU	Data Access Unit (Canada)
DRD	Derived Record Depository (Canada)
DRGs	Diagnosis Related Groups
ECFR	European Charter of Fundamental Rights (EU)
ECHR	European Convention of Human Rights (EU)
eDRIS	Electronic Data Research and Innovation Service (NHS Scotland)
EHDS	European Health Data Space
EHR	Electronic health records
ELGA	Electronic health files [<i>Elektronische Gesundheitsakte</i>] (Austria)

Abbreviation	Name
EOSC	European Open Science Cloud
ERNs	European Reference Networks
EU	European Union
FFS	Fee-for-service
GDPR	General Data Protection Regulation
GIS	Geographic information system
GIS&T	Geographic Information Science and Technology
GP	General Practitioner
HDH*	Health Data Hub (France)
HDR UK	Health Data Research UK
HES	Hospital Episode Statistics (UK)
HiM*	Health in the Municipality [<i>Zdravje v občini</i>] (Slovenia)
HIRA	Health Insurance Review and Assessment Service (Republic of Korea)
ICES	Institute for Clinical Evaluative Sciences (Canada)
ICGS	Integrated care systems (NHS England)
ICNARC	Intensive Care National Audit And Research Centre
IDI	Integrated Data Infrastructure (New Zealand)
IMD	Index of Multiple Deprivation (UK)
INN	International non-proprietary name
KCDC	Korea Centre for Disease Control and Prevention
Kela	Social Insurance Institution of Finland [<i>Kansaneläkelaitos</i>]
LDS	Limited Data Set (US)
MBS	Medicare Benefit Schedule (Australia)
MCHP	Manitoba Centre for Health Policy (host for Manitoba Population Research Repository, Manitoba, Canada)
MHRA	Medicines and Healthcare Products Regulatory Agency (UK)
MHSDS	Mental Health Services Data Set (UK)
MiBa	Danish Microbiology Database [<i>Den danske mikrobiologidatabase</i>]
MoH	Ministry of Health
NHI	National Health Insurance (Republic of Korea)
NHIDB	National Health Insurance Database (Republic of Korea)
NHIS	National Health Insurance Service (Republic of Korea)
NHS	National Health Service (UK)
NIHR	National Institute for Health Research (UK)
NIJZ	National Institute of Public Health [<i>Nacionalni inštitut za javno zdravje</i>] (Slovenia)
NIR	Registration number in the register of natural persons (social security number) [<i>Numéro d'inscription au répertoire des personnes physiques</i>] (France)
Nivel	Netherlands Institute for Health Services Research

Abbreviation	Name
OECD	Organisation for Economic Co-operation and Development
ÖGIS	Austrian Health Information System [<i>Österreichisches Gesundheitsinformationssystem</i>]
OHDSI	Observational Health Data Studies and Informatics
OOP	Out of pocket
PBS	Pharmaceutical Benefits Scheme (Australia)
PCORI	Patient-Centered Outcomes Research Institute (US)
PCORnet	National Patient-Centered Clinical Research Network (US)
PHARMO	PHARMO Institute for Drug Outcomes Research (The Kingdom of the Netherlands)
PHE	Public Health England (UK)
PHI	Private health insurance
PHIN	Personal health information number (Canada)
PHRN	Population Health Research Network (Australia)
PIPA	Personal Information Protection Act (Republic of Korea)
PMSI	Programme for the Medicalization of Information Systems [<i>Le Programme de médicalisation des systèmes d'information</i>] (France)
PN	Personal identification number [<i>Personnummer</i>] (Sweden)
PopData BC	Population Data British Columbia (Canada)
PP	Health Data Portal [<i>Podatkovni portal</i>] (Slovenia)
PROMs	Patient Reported Outcomes Measures
PUFs	Public Use Files (US)
QOLC	Quality of Life of Colorectal Cancer Survivors in England (UK)
QOLP	Quality of Life of Cancer Survivors in England: Pilot Survey (UK)
REGIS	Regional Health Information System [<i>Regionales Gesundheitsinformationssystem</i>] (Austria)
ResDAC	Research Data Assistance Center (US)
RIFs	Research Identifiable Files
KKP	Regional Clinical Quality Programme [<i>Regionernes kliniske kvalitetsudviklingsprogram</i>] (Denmark)
RTDS	National Radiotherapy Dataset (UK)
RUT	Register Utilizer Tool [<i>Register Utiliser Tool</i>] (Sweden)
RWD	Real-world data
SACT	Systemic Anti-Cancer Therapy Dataset (UK)
SAIL	Secure Anonymised Information Linkage (SAIL) Databank (Wales, UK)
SDLE	Social Data Linkage Environment (Canada)
SEER	Surveillance, Epidemiology, and End Results (US)
SGSS	Second generation surveillance system
SHI	Social health insurance
SNDS	French National Health Data System [<i>Système national des données de santé français</i>]
SNIIRAM	National Inter-Plan Health Insurance Information System (Outpatient claims dataset) [<i>Système national d'information inter-régimes de l'Assurance maladie</i>] (France)
SPIRE	Scottish Primary Care Information Resource (Scotland, UK)

Abbreviation	Name
TEHDAS	Towards the European Health Data Space (Joint Action funded by the European Union)
THL	Finnish Institute for Health and Welfare [<i>Terveyden ja hyvinvoinnin laitos</i>]
UDHR	Universal Declaration of Human Rights (UN)
UPI	Unique patient identifier
VHI	Voluntary health insurance
WMC	Welsh Multimorbidity e-Cohort
zVEM	Patient Data Portal [<i>Zdravje Vse na Enem Mestu</i>] (Slovenia)
YLD	Years lost to disability
YLL	Years of life lost
ZZS	Health Insurance Institute of Slovenia [<i>Zavod za zdravstveno zavarovanje Slovenije</i>]

* Most abbreviations are derived from the original language name; however, for some, which are indicated with an asterisk (*), the used abbreviation is in English, despite the country and language of origin.

LIST OF KEY TERMS

Term	Definition	Source*
Aggregated data	Data of several individuals that have been combined to show general trends or values	OECD Glossary of Statistical Terms
Anonymization	Process of rendering personal data anonymous	GDPR
Biometric data	Personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data	GDPR
Confidentiality	Ensuring that information is accessible only to those authorized to have access	OECD Glossary of Statistical Terms
Consent	Any freely given, specific, informed, and unambiguous indication of the data subject's wishes by which they, by a statement or by a clear affirmative action, signify agreement to the processing of personal data relating to them	GDPR
Cross-border processing	Processing of personal data which takes place in the context of the activities of establishments in more than one Member State of a controller or processor in the EU where the controller or processor is established in more than one Member State; or processing of personal data which takes place in the context of the activities of a single establishment of a controller or processor in the EU but which substantially affects or is likely to substantially affect data subjects in more than one Member State	GDPR
Data access	Authorized, on-demand ability to access, modify or edit selected data, regardless of location. Data Access is one of the main aspects of establishing successful data governance systems	OECD Glossary of Statistical Terms
Data concerning health	Personal data related to the physical or mental health of a natural person, including the provision of healthcare services, which reveal information about their health status	GDPR
Data governance	The exercise of decision-making and authority for data-related matters	Digital Health Europe, Glossary
Data linkage	Technique that involves bringing together and analysing records of data from a variety of sources, typically data that relate to the same individual	Digital Health Europe, Glossary, and OECD Glossary of Statistical Terms
De-identification	General term for any process of removing the association between a set of identifying data and the data subject	Digital Health Europe, Glossary

Term	Definition	Source*
eHealth	According to the World Health Organization (WHO), eHealth (or ehealth, or e-Health) is the cost-effective and secure use of information and communication technologies (ICT) in support of health and health-related fields. It encompasses multiple interventions, including telehealth, telemedicine, mobile health (mHealth), electronic medical or health records (eMR/eHR), big data, wearables, and even artificial intelligence	WHO, 2005
EHR	An Electronic Health Record (EHR) is a comprehensive medical and cross-institutional record or similar documentation of the past and present physical and mental state of health of an individual in electronic form	Digital Health Europe, Glossary
Electronic patient identifier	This commonly refers to a unique number or chip card used to electronically identify the patient. Patient identification is necessary to correctly match a patient to an intended treatment and prevent harm due to potential mistreatment	Digital Health Europe, Glossary
Genetic data	Personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or the health of that natural person	GDPR
Interoperability	The ability of organizations to interact towards mutually beneficial goals, involving the sharing of information and knowledge between these organizations, through the business processes they support, by means of the exchange of data between their ICT systems	Digital Health Europe, Glossary
Microdata	Individual level data	—
Personal data	Any information relating to an identified or identifiable natural person ('data subject')	GDPR
Processing	Any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means	GDPR
Pseudonymization	The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information	GDPR
Real world data	Real world data are big data, referring specifically to any type of data not collected in a randomized clinical trial	Digital Health Europe, Glossary
Recipient	A natural or legal person, public authority, agency or another body, to which the personal data are disclosed, whether a third party or not	GDPR
Routine data	Routine data, or routinely collected data, in healthcare are a subset of real world data (see above) and refer to data that are generated in administrative processes or clinical care and whose primary purpose is to support these processes. They can include hospital discharge data, prescription dispensing, emergency department attendances, insurance claims and death certificates	OECD, 2019
Secondary use of data (or data re-use)	Any legitimate use of a healthcare record other than for the purpose of supporting the direct delivery of healthcare services to the subject of care	Digital Health Europe, Glossary
Unique patient identifier	A method for standardizing patient identification. Individuals are assigned a unique code, and that code is used by healthcare organizations to identify and manage patient information	CMS

* Full citations for sources listed here can be found in the References section.

LIST OF TABLES, FIGURES AND BOXES

Tables

TABLE 1.1	Selection of countries and reasons for inclusion	6
TABLE 1.2	Survey instrument for describing case studies	8
TABLE 2.1	Overview of case studies and their contents	40
TABLE 3.1	Scope of datasets bringing together information from different sectors and examples of studies using their data	46
TABLE 3.2	Merging of data in primary care	49
TABLE 3.3	Models with a central location linking distributed datasets (different data holders)	51
TABLE 4.1	Examples of linking methods used	62
TABLE 5.1	Data access processes for selected case studies	66
TABLE 6.1	Access options and patient rights for electronic patient records	71
TABLE 8.1	Patient summary administrative and clinical data	89
TABLE 8.2	Expanded example of patient summary data	90

Figures

FIGURE 1.1	Conceptual framework for the review	3
FIGURE 1.2	Possible content of different data categories	4
FIGURE 2.1	Sundhed.dk: Data flow and actors	19
FIGURE 2.2	Core processes of the Danish Clinical Registries	20
FIGURE 2.3	Indicators included in HiM	30
FIGURE 3.1	Overview of CMS data formats according to ResDAC	47

FIGURE 4.1	Linking approach in Scotland (eDRIS)	58
FIGURE 4.2	Linking approach in Manitoba	60
FIGURE 5.1	Requesting and using data via SURE	65
FIGURE 7.1	Data protection considerations for data content and processing	79
FIGURE 8.1	European Health Data Space	92

Boxes

BOX 2.1	Health data in the IDI	25
BOX 2.2	Databases that can be linked to CPRD data	33
BOX 6.1	Definitions according to the US Department of Health IT	69
BOX 7.1	The rights to privacy in international consensus documents	75
BOX 9.1	The ZOE COVID app in Scotland	96
BOX 9.2	The use of geographic data during COVID-19	96
BOX 9.3	Two examples of international health data research cooperation during COVID-19	103
BOX 9.4	Three country examples of new health data protection guidance adopted during the COVID-19 pandemic	105

ABSTRACT

An indispensable prerequisite for answering research questions in health services research is the availability and accessibility of comprehensive, high-quality data. It can be assumed that health services research in the coming years will be increasingly based on data linkage, i.e., the linking, or connecting, of several data sources based on suitable common key variables. A range of approaches to data collection, storage, linkage and availability exists across countries, particularly for secondary research purposes (i.e., the use of data initially collected for other purposes), such as health systems research. The main goal of this review is to develop an overview of, and gain insights into, current approaches to linking data sources in the context of health services research, with the view to inform policy, based on existing practices in high-income countries in Europe and beyond. In doing so, another objective is to provide lessons for countries looking for possible or alternative approaches to data linkage. Thirteen country case studies of data linkage approaches were selected and analysed. Rather than being comprehensive, this review aimed to identify varied and potentially useful case studies to showcase different approaches to data linkage worldwide. A conceptual framework was developed to guide the selection and description of case studies. Information was first identified and collected from publicly available sources and a profile was then created for each country and each case study; these profiles were forwarded to appropriate country experts for validation and completion. The report presents an overview of the included countries and their case studies (Chapter 2), with key data per country and case study in the appendices. This is followed by a closer look at the possibilities of using routine data (Chapter 3); the different approaches to linkage (Chapter 4); the different access routes for researchers (Chapter 5); the use of data for research from electronic patient or health records (Chapter 6); foundational considerations related to data safety, privacy and governance (Chapter 7); recent developments in cross-border data sharing and the European Health Data Space (Chapter 8); and considerations of

changes and responses catalysed by the COVID-19 pandemic as related to the generation and secondary use of data (Chapter 9). The review ends with overall conclusions on the necessary characteristics of data to inform research relevant for policy and highlights some insights to inspire possible future solutions – less or more disruptive – for countries looking to expand their use of data (Chapter 10). It emphasises that investing in data linkage for secondary use will not only contribute to the strengthening of national health systems, but also promote international cooperation and contribute to the international visibility of scientific excellence.

EXECUTIVE SUMMARY

Health services research is an indispensable tool for evidence-informed decision-making in health policy. It is a multifaceted discipline with a complex, transdisciplinary character. In addition to descriptive and inference-oriented analysis along the three quality-of-care dimensions (structures, processes and outcomes), the questions posed by health services research include healthcare design and healthcare management, and the implementation and evaluation of interventions. Indispensable prerequisites for providing meaningful answers to all these questions, especially against the background of “double complexity” (i.e., complex interventions with complex contextual conditions), are the existence and accessibility of comprehensive high-quality data.

Given the dynamic landscape of different data sources that could be useful for health services research and the increasing importance of questions that require a wide range of information to answer, it can be assumed that health services research in the coming years will largely rely on data linkage, i.e., the linking of multiple data sources based on appropriate common key variables. Such data sources include routine healthcare data, linked with – or linkable to – clinical data, and further, partly unstructured and contextual data.

In an international comparison, it becomes clear that countries have different ways of approaching data collection, storage, linkage and availability for secondary research purposes (i.e., the use of data collected for other purposes), such as health systems research. Though varied, many countries strive to organize their data landscapes and infrastructures to enable diverse, valid and generalizable findings. The two main goals of this review are (1) to develop an overview and gain insights into current approaches to linking data sources in the context of health services research, with the view to inform policy, based on existing practices in high-income countries in Europe and beyond, and (2) to provide lessons for countries looking for possible or alternative approaches to data linkage.

Case studies of interesting linkage approaches in thirteen countries were identified and analysed. Three main criteria were used to select comparator countries: a mix of health system structures; a spectrum in terms of data availability and use; and known interesting examples of relevant initiatives. Case studies were prioritized that allow for the consolidation of intersectoral routine data, data from electronic patient or health records, and registers. They were identified following desk research and/or recommendations from country experts. For the analysis of the identified case studies, a conceptual framework was developed that distinguishes between datasets with health care-relevant data at the individual level and (1) data at the non-individual level and (2) from four major content blocks (health data, healthcare data, sociodemographic or economic data, environmental data). Information was first identified for all comparator countries (for health system and regulatory context) and case studies based on publicly available sources. A profile was then created for each country and case study and forwarded to appropriate country experts for validation and completion.

The report presents an overview of the included countries and their case studies (Chapter 2), with key data per country and case study in the appendices. This is followed by a closer look at the possibilities of using routine data (Chapter 3); the different approaches to linkage (Chapter 4); the different access routes for researchers (Chapter 5); the use of data for research from electronic patient or health records (Chapter 6); foundational considerations related to data safety, privacy and governance (Chapter 7); recent development in cross-border data sharing and the European Health Data Space (Chapter 8); and considerations of changes and responses catalysed by the COVID-19 pandemic as related to the generation and secondary use of data (Chapter 9). The review ends with overall conclusions on the necessary characteristics of data to inform research relevant for policy (Chapter 10). It summarizes patterns in data linkage from the case studies and highlights some insights to inspire possible future solutions – less or more disruptive – for countries looking to expand their use of data.

The case studies presented in this review can be roughly divided into five categories:

- i. **large administrative datasets**, which, due to a central administration point, make it possible to map questions about the care process of individual patients across sectors and offer possibilities for linking to other data;
- ii. **centralized locations that merge existing datasets** of different data holders and with variable content; here, two possibilities stand out: the data remain with the data holders and are linked on request or are stored in a centralized location;
- iii. **databases which combine routinely recorded data** and make it available to clinicians and researchers;
- iv. **patient-centric electronic platforms** that allow (inter alia) patients to gain access to their own information and which could potentially be considered as a possible basis for clinical data for research; and
- v. **tools that facilitate research** across multiple datasets.

The review captures cases studies that highlight the potential of using existing, integrated routine data (i.e. datasets that have already been created for the primary purpose of use in an integrated form across several service areas) to answer research questions. These are at the insured person level, are available over several years and cover various service areas. Examples from Australia, the Republic of Korea and the United States of America (USA) provide insights on the possibilities for the use of integrated routine data. In all countries these are datasets that are predominantly composed of billing data (claims data).

Some case studies in this review support the notion that a good data basis for research purposes exists above all when all information on the interaction of individual patients with the healthcare system is held in a central location; in this way, care processes involving service providers and, if necessary, across sectors can be mapped, and, in the optimal case, also over time. This condition is largely fulfilled in health systems that have established a “gatekeeping” system in primary care; among the case studies included, the most typical example of a database that brings together data collected by GPs is the Clinical Practice Research Datalink (CPRD) in the UK. Data held by CPRD may be requested for observational studies or used to support experimental studies. In addition, they can be linked to other NHS Digital datasets to provide richer, cross-sector insights.

The review identified several case studies in which diverse data sources with differing content are linked by central agencies via one or more linking variables to provide better data access for research. For example, data on healthcare can be linked with data from the education sector and crime statistics. These case studies clearly highlight the opportunities for combining independently generated datasets for health services research on the basis of such central bodies and different options about how to realise this operation. For example, the Australian Population Health Research Network (PHRN) and the Canadian Social Data Linkage Environment (SDLE) link data on a project-specific basis, while the source datasets remain distinct. In contrast, New Zealand’s Integrated Data Infrastructure (IDI) merges incoming data using a “spine”.

Successful linking requires the possibility of a one-to-one assignment of information from separate data records. This requires suitable key variables that are present in all data records to be linked. The linking initiatives in the countries included in this review show that even without the presence of a unique identifier, linking based on other identifying characteristics is possible and has been used successfully for several years.

This review also showcases the various pathways to data access for researchers. Since data in the healthcare system are fundamentally sensitive, there are generally three possible access routes: (1) direct data transmission, which requires an appropriate data protection and data security concept; (2) in-house evaluation, which takes place at suitable host workstations of the institution holding the data; and (3) remote query, which is possible on separate servers with “remote access” via VPN connection.

Among the countries in this review there is no single common policy approach for regulating personal digital health data. While some countries have designated comprehensive legislation on health data privacy, others have a national data governance framework, where health is incorporated – or both. In some countries additional legislative initiatives may have clauses that influence the practice of data protection, especially as it pertains to vulnerable populations or specific health areas. Depending on the administrative structure, data governance occurs at the national, regional or local levels of government or a combination. However, across countries there are similar data principles underpinning most regulatory approaches, informed by international guidance and regulations; these include accessibility, accountability, accuracy, confidentiality, data minimization, integrity, purpose limitation, storage limitation and transparency. Entered into force in 2016 and replacing the 1995 Data Protection Directive, the EU's GDPR is the core of Europe's digital privacy legislation and sets guidelines for the protection of personal information from residents of the EU.

Several key initiatives have been introduced at EU level that have implications for health services research, including the cross-border healthcare directive, European Reference Networks (ERNs) to the European Health Data Space (EHDS). Furthermore, the COVID-19 pandemic has spurred considerable innovation in health and healthcare data generation, collection, sharing and linkage. The review briefly presents new sources of data developed for the pandemic response, as well as examples of COVID-related data linkage activities and innovations along with innovations in data access and sharing and the integration of COVID-19 data with electronic health records, against the backdrop of additional data protection and privacy considerations.

The review concludes by presenting criteria that need to be met to enable meaningful healthcare research based on linked data (including data content and coverage in terms of population, healthcare sectors and timeframe) and the main desirable tenets of relevant regulatory frameworks. It showcases the strengths and limitations of coordination and linkage hubs as well as the need to continue investing in the potential of routine data on top of new modalities, e.g., on the basis of electronic patient records. It finally recognises the importance of ongoing initiatives, such as the EHDS, for the evolution of these questions in the future, and underlines the need for expanding data linkage options with environmental variables.

Introduction

Evidence-informed decisions in healthcare policy and practice rely on the timely availability of comprehensive, good quality data. With increasing shares of populations living longer and with (multiple) chronic conditions as well as both emerging and persistent viral threats putting pressure on already limited healthcare resources, health policy-makers place a premium on new approaches and solutions to improve health system governance and the quality and efficiency of care.

Scientific efforts to identify, describe and analyse health system challenges, develop, test, and evaluate different interventions, and provide insights on policy options have often been subsumed under different areas of study depending on their thematic focus and methodology, including health services research, health systems research, health policy research and public health research. A common requirement is the need for comprehensive datasets, including patient care data across care sectors, health outcomes data and cost data. Technological advances in data collection as well as the proliferation and increasing availability of electronic data from a range of sources are expanding the possibilities of scientific research and evidence-informed decision-making.

Against the dynamic landscape of different data types and sources and the increasing importance of complex questions that require a broad spectrum of information to be answered (e.g., on the influence of climate change on health-related outcomes or the impact of prioritizing COVID-19-related healthcare on different population groups), it can be assumed that health services research in the coming years will increasingly rely on data linkage, bringing together several data sources on the basis of suitable common key variables. These data sources can encompass routine healthcare data (including administrative and claims data, see Glossary), clinical data and other, partly unstructured

every-day and contextual data on health status as well as socioeconomics and demography. To meet these demands, it is to be expected that research will increasingly take place in larger networks; however, the expectation that data informing decisions in health policy and practice are characterized by high validity, completeness, timeliness, representativeness and quality should remain unaffected. These realities are gaining recognition as reflected in recent policy initiatives (e.g., in the context of the European Union (EU), in the proposal for a European Health Data Space (EHDS)).

The main goal of this review is to provide insights into possible approaches to the linkage of different data sources in the context of health services research aiming to inform policy, based on existing practices in Europe and beyond. As such, the focus of this review is secondary use of data (i.e., the use of data collected for other purposes in analysis aiming to answer broader questions). The review was constructed in such a way as not to replicate existing work that compares the available evidence base for health services research in different countries (e.g., Oderkirk, 2021; OECD, 2015; Cole et al., 2015; Gothe, 2014), but to take a more in-depth look and to distill takeaways for policy-makers.

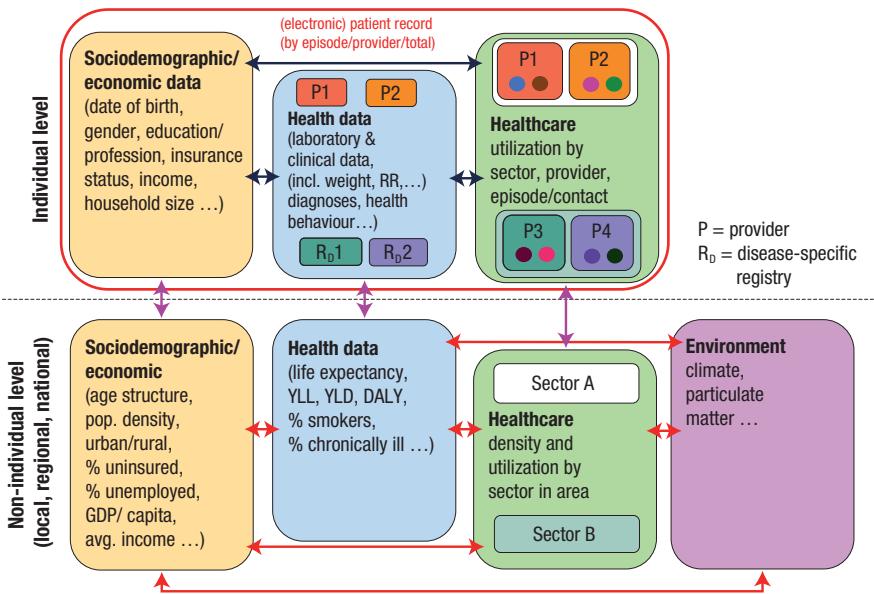
1.1 Methodological approach

Analytical framework

In every healthcare system the development – and possibility of linkage – of relevant databases is closely related to its organizational design, in particular the structure of service provision and the corresponding billing and reimbursement mechanisms. The accessibility of these databases is also influenced by data protection regulations and standards, including cultural considerations. The analysis in this review aimed to present findings on data structure and flow in a manner that is, to the extent possible, system-independent and transferable, while simultaneously also capturing country-specific regulatory frameworks.

To select and analyse the different approaches (or “case studies” from this point onward), a conceptual framework was developed that distinguishes between datasets with relevant data at two different levels (at the individual level: “individual level data”, i.e. microdata; and at the non-individual level), and from three or four large content blocks (sociodemographic or economic data, health data, healthcare data, and environmental data) (see also Figure 1.1).

FIGURE 1.1 Conceptual framework for the review



Source: Authors' own.

Using health data (marked in blue in Figure 1.1) as an example, individual-level health data concerns data for an individual patient, such as laboratory and clinical results, vital signs (body temperature, pulse rate and respiration rate), as well as diagnoses and health behaviour. Non-individual level data include aggregated data in areas such as life expectancy, years of life lost (YLL), years lost to disability (YLD) and disability-adjusted life years (DALY), as well as population characteristics such as prevalence of risk factors and chronic illness.

In Figure 1.1, the arrows are intended to symbolize the (possible) linkages that are available for individual data either by storing them in common databases (e.g., based on an electronic health record) or by linking them via unique personal characteristics (e.g., a [unique] patient identifier). Equally relevant, however, are the possibilities for linking individual data with non-individual data. For the purpose of this review, it was not expected that data(bases) are necessarily structured as in Figure 1.1; the categories are rather intended to categorize the content of the data contained in each case study. Examples for variables from the different data categories in Figure 1.1 are shown in Figure 1.2.

FIGURE 1.2 Possible content of different data categories

Socioeconomic Determinants	Health	Healthcare	Environment
Individual levels ✓ Age ✓ Sex ✓ Other demographic factors (e.g., immigration background, marital status) ✓ Employment ✓ Vocational training status ✓ Insurance status ✓ ...	✓ Exam results ✓ Lab results ✓ Diagnoses/symptoms ✓ Risk factors ✓ Procedures conducted ✓ Medication plans ✓ Medical devices ✓ Health outcome measures (e.g., falls, pain management, wrong medication) ✓ ...	✓ Exam results; lab results ✓ Lab results; Diagnoses/symptoms ✓ Physician appointments ✓ Transfers ✓ Hospitalisations ✓ Diagnostic procedures ✓ Therapeutic procedures ✓ Medical devices ✓ Medical aids prescribed ✓ ...	n/a
Aggregate levels ✓ Median household income ✓ GDP/capita ✓ Population in city/county ✓ % below the poverty line ✓ Education level ✓ Marital status ✓ % unemployed ✓ Insurance coverage ✓ % single parents ✓ % home owners ✓ % in need of support ✓ Age structure ✓ Div. Aggregate ✓ ...	✓ Life expectancy ✓ Disability-adjusted life years (DALYs) and components (YLL/YLD) ✓ % people with risk factors ✓ % chronically ill people ✓ % person with limitations ✓ ...	✓ Supply density/degree of utilisation ✓ ...	✓ Air purity ranking ✓ Climate indicators ✓ Toxins/waste ✓ Distance metrics ✓ Degree of development/ land use in rural areas ✓ ...

Source: Authors' own.

Selection of comparison countries and case studies

Three main criteria were used to select the countries for comparison in this review with the aim of ensuring both diversity and representation:

- healthcare system structure:** both statutory health insurance and national tax-funded health systems, with different degrees of decentralization of service provision and healthcare system management;
- data availability and use:** for this purpose, the periodic survey on Health Data Governance of the OECD was used (OECD, 2015), which describes the activities of several OECD countries with regard to the regulation, development and use of health-related databases.

While a newer iteration of the survey is now available (Oderkirk, 2021), this was published after the data collection for this review had been completed; and

- c. **interesting examples of relevant initiatives;** these were selected based on discussions with health services research experts within the networks of the European Observatory on Health Systems and Policies.

In line with the scope and purpose of the Observatory's HiT reviews, the selection included interesting cases within the EU, including Austria, Denmark, Finland, France, the Kingdom of the Netherlands, Slovenia, Sweden and the United Kingdom (UK). All these countries have implemented specific rules for regulating health-related data also digitally, as per the EU General Data Protection Regulation (GDPR). The analysis was extended to other OECD countries to capture additional case studies of particular interest. Though the regulatory structures and legal frameworks vary considerably by country, data privacy legislation largely reflects the OECD 2013 privacy framework and guidelines (OECD, 2013). The United States (US) and Canada were added, as they tend to have more industry self-regulation and more fragmented approaches. Australia and New Zealand were included owing to the high rates of electronic health record use in parts of the health sector and the ability to use integrated data for research, underpinned, however, by a federalized and nationwide approach, respectively. Lastly, the Republic of Korea was included to highlight the potential of a unified claims database with its large dataset available for research.

This review was not intended to be comprehensive or systematic with regard to the selection of included countries, nor does it contain an exhaustive presentation of their datasets. The objective was rather to identify varied and potentially useful case studies that could provide insights into the different approaches to data linkage that exist worldwide and to perhaps provide ideas for future more or less "disruptive" solutions for countries working on expanding their use of data. Table 1.1 shows the final selection of countries used in this review, the primary reason of interest per country and the scoring in terms of data availability, maturity and use according to the OECD (OECD, 2015).

TABLE 1.1 Selection of countries and reasons for inclusion

COUNTRY	REASON OF INTEREST	RANKING OECD (2015)*		
		Data availability	Use of data	Rank**
Australia	Decentralized distributed data sources are merged into study-specific metadata with the participation of the data holders via a central trust centre.			Not covered
Austria	Introduced electronic patient records with potential research accessibility; several health insurers, regional and central health reporting systems.			Not covered
Canada	Regional linking initiatives bring together several datasets based on different structural models.	High	High	12
Denmark	All routine data and, increasingly, clinical data will be combined; academic institutions will have comparatively easy access.	High	Medium	5
Finland	New digital networking options facilitate the consolidation and provision of health data.	High	High	8
France	A new data authority enables project coordinators to easily access pseudonymized, linked data securely for research purposes. The authority primarily provides access to the national health data system's databases, including on outpatient claims, hospital discharges, death and disability registries for about 99% of the population.			Not covered
Netherlands (Kingdom of the)	Data from several service providers in the outpatient sector are merged to enable integrated care and can be used for research purposes.	Medium	Medium	17
New Zealand	The national 'Integrated Data Infrastructure' combines a variety of individual-level centrally-collected data (up to and including the census) and allows their use on request for projects that are in the public interest.	Medium	High	10
Republic of Korea	A central health insurance company (NHIS) makes the raw administrative data of all >50 million Republic of Koreans (HIRA) available for research remotely on a central server on an application basis or at seven research centres and provides summary patient population and national statistics more openly.	High	High	2
Slovenia	New eHealth solutions attempt to streamline existing fragmented information systems through secure data to improve care coordination, facilitate communication among providers and increase the availability of medical, economic and administrative data at the individual, regional and national level for research purposes.			Not covered

COUNTRY	REASON OF INTEREST	RANKING OECD (2015)*		
		Data availability	Use of data	Rank**
Sweden	Well established, linkable patient registries; innovative use of linked data to support clinical guidelines.	High	High	4
United Kingdom (England)	Clinical Practice Research Datalink collects de-identified data from more than 10 million patients (source: family doctors).	Medium	High	14
United Kingdom (Scotland)	Data of all inhabitants from various sources (pharmacies, general practitioners, hospitals, public authorities such as the police) are held in both regional and national centres and can be linked with the Community Health Index.	Medium	High	7
USA	Research institutions enter into agreements with organizations and individuals on a private law basis, offer knowledge in return for data donation.	Medium	Medium	18

Note: * low / medium / high = 0–33% / 34–66% / 67–100% of all datasets available or regularly linked for research purposes; ** among the 22 countries surveyed in view of a number of variables relating to data availability, maturity and use (1st place corresponds to the highest and therefore best score; this was occupied by Iceland in 2015).

Case studies were prioritized that allow for the consolidation of intersectoral routine data, data from electronic patient or health records, and registers. They were identified following desk research and/or recommendations of the country experts.

Information collection and validation

In order to describe the included case studies and the relevant context in the comparison countries, a survey instrument (Table 1.2) was designed, piloted, adapted and finally applied in several steps.

TABLE 1.2 Survey instrument for describing case studies

GENERAL QUESTIONS PER COUNTRY	SPECIFIC QUESTIONS PER CASE STUDY
<ul style="list-style-type: none"> Structure of the healthcare system, organization of service provision. Legal framework for data use and security (general and specific to health) Strategies for health data Implementation and responsible actors If applicable, additional data innovations of the last years* Publications that can serve as case studies for data use 	<ul style="list-style-type: none"> Primary objective: research, governance/quality assurance, coordination of care, remuneration, etc. What data are collected and by whom? Are the collected data structured/unstructured? On which level are data aggregated (regional/national)? and who is responsible for it? Are plausibility checks performed? By whom? Are the data of high quality? Where are data located/stored and for how long? Who owns the data? ("ownership") Can the data be linked via UPI or are other methods used? Who is responsible for the linking? Data access: Who is authorized? What does the process look like? How long does it take until data can be made available? What is currently the last available data year? Previous use of the data for research purposes? How representative is the population covered by the data for the country/region? For non-individual data: what is the smallest unit for which linking is possible? How many of these units are there?

Note: * Data collection was carried out before the COVID-19 pandemic; for advances in data collection and use catalysed by the pandemic, see Chapter 9.

Information on the questions listed in Table 1.2 was initially identified for all included countries and case studies based on publicly available sources, mainly based on online searches for published and gray literature, as well as official information from the websites of the respective healthcare system actors. The reviews of the HiT series of the Observatory were used, supplemented in most cases by official documents on data regulations and strategies of the individual countries. Information on the specific questions per case study was first accessed via the respective websites, supplemented by current published descriptions of the corresponding databases as well as information from published healthcare research studies that were conducted based on the datasets in question. Subsequently, one profile per country was created and forwarded to suitable country experts for validation and completion. Suitable experts were identified from the networks of the Observatory. Research for this review began before the COVID-19 pandemic; not all changes in data availability and access stemming from pandemic management may be reflected throughout the review, but major trends have been captured in a dedicated chapter (see below).

Time of latest available information varies slightly from country to country owing to the evolving nature of policy development concerning the secondary use of health-related data in countries as well as the dynamic production processes for this report.

1.2 Overview of the review

The following chapters present a descriptive synthesis of findings across case studies. Chapter 2 includes an overview of included countries and their respective case studies. Chapter 3 focuses on the (extended) possibilities of using routine data, while Chapter 4 illustrates examples of different linkage approaches with and without a unique patient identifier. The different access paths for researchers identified across case studies are summarized in Chapter 5, while Chapter 6 showcases specific reflections on using data from electronic patient or health records. Chapter 7 then moves beyond the specifics of included case studies to provide foundational considerations on data safety and protection, and Chapter 8 describes recent developments in cross-border data sharing and in particular the European Health Data Space. Chapter 9 provides considerations for data linkage in the context of COVID-19. The review ends with overall conclusions on the necessary characteristics of data to inform research relevant for policy (Chapter 10). The appendices contain key data for each country (including population and the administrative structure of the country, health expenditure, structure of service provision, data protection regulations, degree of technological development) as well as the most important information on the individual case studies in profile format.

2

Overview of included countries and case studies

This chapter provides an overview of the data linkage case studies included in this review, organized by country. It gives a short introduction to each country's health system (in particular delivery models and payment mechanisms which influence data structures) and then focuses on the objectives of each case study, their general structure, and the amount and type of data(sets) they bring together. A description of each case study in profile format can be found in Appendix I.

2.1 Australia

Australia has a federal administrative structure with six states, two internal territories and seven external territories. The healthcare system is tax-financed with universal access to healthcare services, which are billed through the Medicare Benefits Schedule (MBS) and the Pharmaceutical Benefits Scheme (PBS). Medicare sees itself as a “public health insurance system” (Glover, 2019). Private supplementary insurance allows, among other things, access to ambulance services, dental services and glasses, contact lenses and hearing aids (all not covered by Medicare), a wider choice of providers, faster access or lower co-payments; almost half of all Australians make use of it. Medicare is funded by the federal government, but the organization and delivery of healthcare is decentralized. General practitioners (GPs) have a gatekeeping function, but patients do not have to register with specific GPs. Service providers in primary care are mainly remunerated on a fee-for-service basis; publicly funded

hospitals are reimbursed through global budgets and flat rates per case, private hospitals on a fee-for-service basis.

The Medicare Benefits Schedule and Pharmaceutical Benefit Schedule

The Medicare Benefits Schedule (MBS) maintains a current and historical data repository of all MBS-subsidized, billed services (outpatient and inpatient); similarly, the Pharmaceutical Benefit Schedule (PBS) maintains data on all billed prescriptions financed by the PBS as well as prescriptions that cost less than the co-payment limit and are therefore the responsibility of the citizen. The MBS and PBS data are available on an insurance-related basis, linked via the “Medicare Number”, and are regularly analysed for the management of the insurance programme, for example with regard to utilization and costs, for budget planning, and for the development of risk-sharing agreements with drug manufacturers.

The Population Health Research Network

The MBS and PBS data are also accessible for health services research and can be linked to other datasets via the services of regional linking agencies and the Population Health Research Network (PHRN), which forms the national network of these agencies. In addition, the PHRN is a secure data laboratory and a coordination centre for electronic data processing services. It was founded in 2009 as part of a federal government initiative on collaborative research infrastructure. While regional datasets vary in some cases, the PHRN’s linked data collections contain a wide range of data: health-related data, including MBS/PBS billing data, births, deaths, neonatal and cancer registries, limited clinical data, and other registry data. In addition, there are data on housing conditions, justice and education. This Australian government investment in linking infrastructure has led to an increase in the use of linked data for research purposes (Young & Flack, 2018).

My Health Record

A more recent initiative in Australia is the implementation of the electronic patient file “My Health Record”, which was piloted in 2016 and rolled-out in 2019. Australian citizens had until the end of January 2019 to indicate if they did not want to participate (“opt out”) and 10% of the population made use of the “opt out” option. My Health Record includes routine data generated during care, laboratory findings and the results of imaging procedures, as well as documentation on care planning and discharge management. The primary motivation for implementing My Health Record was to improve care coordination and promote patient empowerment. Nevertheless, enabling research in the public interest and with the aim of improving care is also mentioned among the objectives of the initiative. To date, My Health Record data are not yet available for secondary use. A secondary use framework outlining how information in My Health Record can and cannot be used has been established, and a committee has been put in place to implement the framework and establish the technical infrastructure to do so. A Data Governance Board to consider applications for data and oversee future My Health Record data research projects is under development. There is a second “opt-out” option for citizens who want to participate for the primary use of My Health Record, but do not wish for their data to be available for secondary use. My Health Record users should also have the option of opting out of the use of their data for research, even if they wish to use the patient file.

2.2 Austria

The healthcare system in Austria is financed from several sources: the federal government, the nine provinces and the social insurance agencies all contribute to the overall budget. Income for the system comes from general tax revenues (40%) and contributions to the statutory social health insurance (60%), which covers more than 99% of the population. Hospital care (outpatient and inpatient) is largely financed by the *Länder*. Outpatient primary and secondary care is regulated by overall contracts between the statutory health insurance funds and healthcare providers. General practitioners have no formal “gatekeeping” function and much of the care is provided in hospitals. Patients

have a free choice of general practitioners, specialists and hospitals. Outpatient care financed by health insurance funds is characterized by a mixture of per-capita flat rates and individual service payments; privately practising doctors must first be paid directly by the patients, but the costs can be reimbursed subsequently and proportionately by the health insurance funds. Acute care hospitals are reimbursed on a case-by-case basis. The proposed Austrian eHealth strategy of 2007 aimed to improve the integration and compatibility of existing information systems of different types and to enable a stable exchange of information through a common eHealth infrastructure.

Electronic Health Records (Elektronische Gesundheitsakte)

As a basic component of the proposed eHealth Strategy of 2007, the use of electronic health records (Elektronische Gesundheitsakte, ELGA) was introduced with the Health Telematics Act 2021. ELGA allows structured health information from patients (e.g., medical examinations, medication prescriptions, laboratory and radiological findings, discharge letters from the hospital) to be made available to the patients themselves and to authorized service providers. The legal basis for the introduction of ELGA entered into force in January 2013 and the ELGA portal was activated in 2014. Via the portal, patients and service providers can access the described data. Patients can restrict access to selected personal health data and see who has accessed their individual records. Patients are automatically registered for ELGA (no “opt in” is required, see Chapter 6), but can opt out. Currently, the ELGA infrastructure offers e-reports (discharge letters, lab and radiology reports) and e-medication as core services. Additional eHealth services include the electronic vaccination registry. Different from the core ELGA services, the vaccination registry does not allow an opt out. Additional eHealth services and ELGA core components are under development. The development of the ELGA infrastructure is also linked to the opportunities and requirements at European level, e.g., regarding cross-border exchange of e-prescription data.

Austrian Health Information System (Österreichisches Gesundheitsinformationssystem)

The Austrian Health Information System (Österreichisches Gesundheitsinformationssystem, ÖGIS) is used for health reporting and planning and is designed as a geographic information system (GIS). The data available in the system can be presented in aggregated form as maps and time series and cover areas such as mortality, cancer incidence, procedures, hospital density, and subjective health status. Specifically, microdata from the DRG system (including sociodemographic information, diagnoses, billed services) and from outpatient care (physician visits, billed services) are included. Macro data from the causes of death statistics, cancer statistics, hospital statistics, statistics on outpatient service providers, surveys on health behaviour, subjective health assessment and use of the healthcare system (aggregated by age, gender and region of the respondents), demographic and socioeconomic data and data on accessibility are also included. The Regional Health Information System (Regionales Gesundheitsinformationssystem, REGIS) is part of ÖGIS and enables regional analysis of ÖGIS data in the form of maps.

Additionally, the revisions of the Federal Statistics Act 2000 in 2021 and 2022, together with the provisions and update of the Research Organization Act, established the Austrian Micro Data Centre (AMDC). The AMDC is hosted by the Federal Statistics Agency (Statistik Austria) and makes available both the data collected at Statistik Austria and additional registry data that can be included in the AMDC on the basis of directives agreed upon by two responsible Ministers. Experience with these options is still being generated.

2.3 Canada

Canada's ten provinces and three territories have primary responsibility for the organization and delivery of healthcare services and for the supervision of healthcare providers. They operate their own universal health coverage systems, which must adhere to central government guidelines (the Canada Health Act of 1984) in order to receive federal cash transfers. Provincial/territorial health coverage programmes are financed by tax revenues. Approximately two-thirds of all Canadians have private supplementary insurance that covers

non-reimbursed services, such as medications, dental treatment and vision care. General practitioners serve a gatekeeping function and can be freely chosen by patients. “Gatekeeping” here is based primarily on financial incentives: specialists receive lower remuneration for patients without referral and so will usually require patients to have a referral from their general practitioner prior to the initial visit. Primary and specialist care by physicians are generally remunerated through a fee-for-service model (in Ontario about half of primary care physician earnings are from capitation-based payments), while hospitals (which provide secondary and tertiary care) are reimbursed through global budgets. This review covers one national data linkage initiative as well as several linking initiatives at the provincial level.

The Social Data Linkage Environment

In order to enable the use of already available microdata from different sectors in combination, the Social Data Linkage Environment (SDLE) was founded at the Federal Statistical Office, Statistics Canada. A centrally maintained “key register” containing new identifiers based on a number of identifying characteristics (see Chapter 4) is used to link a wide range of datasets, including the Labour Force Survey with sociodemographic and national health surveys (with health, care, sociodemographic and economic microdata), the hospital discharge database (with microdata on services provided and diagnoses), census data, environmental macro data and a range of provincial data on crime, education, mental health, etc. Thus, the SDLE is not an integrated database, but is a secure environment that allows for the linking of other databases, which have also been widely used for scientific research.

ICES, Ontario

The former Institute for Clinical Evaluative Sciences is now known only by its acronym ICES and holds a repository of datasets that can be linked for research and analysis to improve the health of Ontario’s residents, over 90% of which are publicly visible in the data repository. These include data on the interactions of Ontario residents with the healthcare system (healthcare microdata), microdata from surveys on health, utilization of services, sociodemographics

and economics, migration data (sociodemographic and economic microdata), and registry data (including cancer, stroke and cystic fibrosis registries). Macro data from the census provide sociodemographic information on an aggregated level, which can be linked to micro data via geographical features. In addition, ICES enables the linking of data from new research projects, which were also developed outside the institute, to existing datasets.

Population Data British Columbia, British Columbia

Unlike ICES, Population Data British Columbia (PopData BC) does not have its own research agenda but allows external researchers to link their own primary data with the data held in its secure research environment. These datasets include microdata on healthcare utilization (including medication prescriptions and hospital discharges), vital statistics, migration data, occupational health billing data and educational data. Income data are aggregated by postal code and divided into strata. PopData BC has supplied data access for many researchers over the course of its history.

Manitoba Population Data Research Repository, Manitoba

The Manitoba Population Data Research Repository at the University of Manitoba enables the University itself to conduct research based on its data and can also provide access to external parties. The data come from several local, provincial and national sources and include health data at the individual level (from several registries), healthcare utilization data, data on education and justice, census data and a number of more specific sets. The data are de-identified before they reach the repository (for a comparison of different approaches to de-identifying microdata, see Chapter 4). Only aggregated data may be used in publications and presentations. Advantages of repository data include the wide coverage of the Manitoba population (representativeness) and the long-time span of certain datasets.

2.4 Denmark

In Denmark healthcare is planned, coordinated and delivered by the five regions, which represent the middle administrative level in the country (state, regions, municipalities). Securing healthcare for the population via public hospitals and financing private service providers in the outpatient sector also corresponds to the main responsibility of the regions as a whole. By contrast, health promotion and prevention measures and public primary care are the responsibility of the municipalities. The health system is tax-financed with universal access; almost 40% of the population opt for private supplementary insurance to cover co-payments or non-reimbursed services, while about 25% take out packages for a wider range of private service providers. For the vast majority of the population, general practitioners have a “gatekeeping” function for other outpatient services and for inpatient care. Primary care is remunerated on the basis of individual service remuneration or capitation flat rates, while hospitals are remunerated via global budgets and flat rates per case.

The Danish healthcare system has a high degree of digitization – as early as 1994 the development of an eHealth strategy was commissioned to enable the electronic exchange of relevant data between service providers. All hospitals and general practitioners now work with electronic health records: findings are transmitted electronically, as are almost all referrals and medication prescriptions. Data in the respective health file are stored locally on the basis of regional eHealth applications; data from the regional platforms are merged in the Danish Health Data Network. All residents in Denmark have a unique patient identifier from the Central Person Register (Centrale Personregister, CPR), which is also used in all registers.

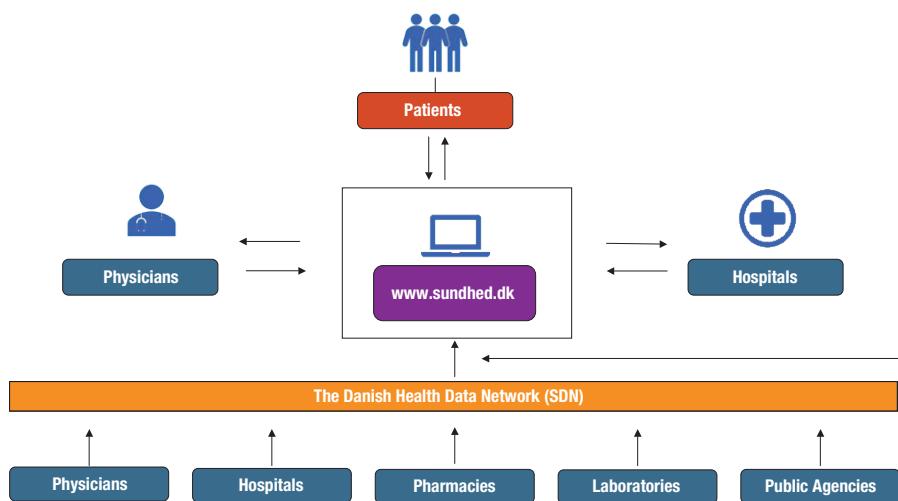
Statistics Denmark is the main provider of register data for research purposes and maintains regularly updated information at an aggregated level, which is accessible free of charge. Access to microdata requires that the requesting research groups are appropriately qualified.

Sundhed.dk

In 2001 the five Danish regions, together with the national government, founded the national eHealth portal Sundhed.dk, which is the access point to the data of the underlying network for patients and service providers (outpatient

care providers, hospitals and pharmacists) (see Figure 2.1) (Sundhed.dk, 2016; Sundhed.dk, 2022). Among other things, patients can obtain information about their diagnoses, treatments and findings, book appointments, renew prescriptions and monitor their own adherence; they can also determine which service providers can view which of their personal data. Service providers can also use the platform to communicate with each other, bill for services, and access relevant evidence (textbooks, scientific articles, clinical guidelines).

FIGURE 2.1 Sundhed.dk: Data flow and actors



Source: Sundhed.dk, 2022.

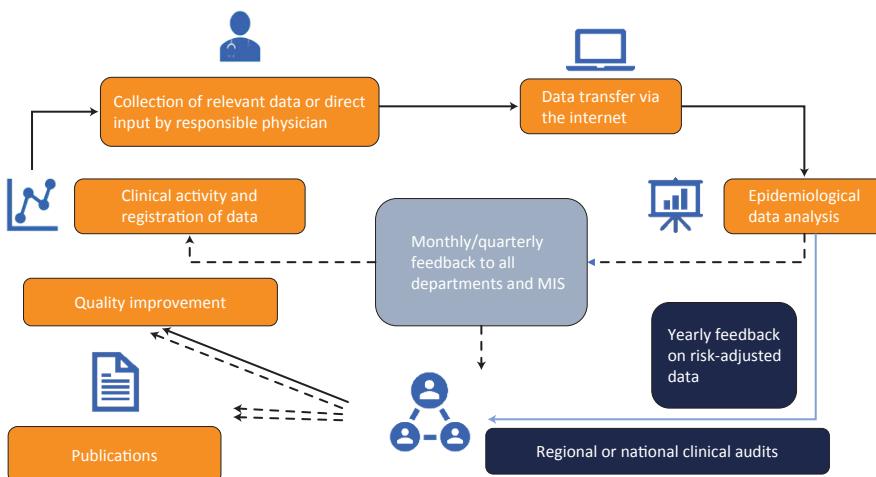
Clinical Quality Programme (Regionernes kliniske kvalitetsudviklingsprogram)

Denmark, like the other Scandinavian countries, has a long tradition of maintaining and using clinical and epidemiological registers to monitor and improve the quality of care.

The Danish Clinical Quality Programme (Regionernes kliniske kvalitetsudviklingsprogram, RKKP) manages 80–85 clinical registries with different targets and populations (e.g., stroke, lung cancer, intensive care, dementia), including clinical data, diagnoses, treatments, prescriptions, laboratory findings, patient-reported endpoints and sociodemographic

information. The registries report regularly to clinicians and managers; additionally, results on an aggregated level are published on an annual basis (see Figure 2.2). Clinicians have access to microdata for patients treated in their practice or clinic for data validation and quality improvement; other clinicians, administrative staff and the public receive only aggregated evaluations of the data contained in the registries. Participation is mandatory for all healthcare providers treating patients with the appropriate clinical indications; the programme does not require patient consent for data processing.

FIGURE 2.2 Core processes of the Danish Clinical Registries



Source: RKKP, 2022.

2.5 Finland

The Finnish healthcare system is highly decentralized in its administrative structure, financed by several mechanisms (tax, insurance contributions, private expenditure) and combines three parallel systems for primary care: the municipal system, the national health insurance system and occupational healthcare. The main system is organized by the municipalities, which are responsible for primary care, and secondary and tertiary care is co-responsible through membership of 20 hospital regions. General practitioners also have a control function (“gatekeeping”) for specialized outpatient and inpatient

care, but patients can choose where they receive the relevant services. Due to the high degree of decentralization, interoperability between the widely used electronic health records has so far been low (Keskimäki et al., 2019); the national electronic health data repository (Kanta) was implemented in stages between 2010 and 2016 and comprises electronic data on prescriptions, treatments, findings and utilization of health services. Service providers have access to nationwide data, patients to their own health and social care data. More than 40% of the population (2 million Finns) now make use of it. Kanta is maintained by the Social Insurance Institution of Finland (Kansaneläkelaitos, Kela); the Finnish Institute for Health and Welfare (Tervyden ja hyvinvoinnin laitos, THL) is responsible for structuring the data after they arrive from individual controllers.

Findata

The new law on the secondary use of health and social data (2019) established the Findata authority, which grants permission for the use of data in research and collates and links samples of existing datasets as requested for research purposes. Findata was launched following the Finnish Presidency of the European Council in 2020. It is envisaged that in the future Kanta Services will be a one-stop shop for all available health-related data and that Findata will be able to sample all stored data directly from Kanta, rather than going to each data controller and adding to their administrative burden.

2.6 France

France's health system is largely centralized, based mainly on a social health insurance (SHI) system that covers all residents. Around 92% of the population are covered by the main fund (National Health Insurance Fund for Salaried Workers, Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, CNAMTS), approximately 7% of the population are covered by the agricultural fund and the remaining population are covered by other small funds. Since 2009 regional health agencies have played a larger role in managing healthcare provision at the local level, but SHI and central government continue to oversee the health system. Over the past two decades the state has also become more involved in controlling health expenditure funded by the SHI system by

setting a national health spending target. Complementary health insurance is very common, with around 95% of the population having VHI, mainly to cover co-payments and to attain better coverage for medical goods and services insufficiently covered by the SHI, like dental and vision care. France has a voluntary gatekeeping system, with incentives for patients to visit their primary care physician before going to specialist care. With regards to payment, hospital care is largely paid through DRGs, whereas outpatient care is mostly fee-for-service (FFS).

Since 2016 the French National Health Data System (Système national des données de santé français, SNDS) has been a warehouse of continuous pseudonymized administrative data on about 99% of the French population. It contains all the care presented for reimbursement. The SNDS makes it possible to link several databases: outpatient claims (National Inter-Plan Health Insurance Information System (Système national d'information inter-régimes de l'Assurance maladie, SNIIRAM), public/private hospital discharge summaries (Le Programme de médicalisation des systèmes d'information, PMSI), death and medical causes of death (Centre d'épidémiologie des causes de décès, CépiDC) and disability (Caisse nationale de solidarité pour l'autonomie, CNSA). These databases contain data on demographics, medical information, outpatient reimbursements, hospitalizations, medicines/devices, survey/interview data, registry data, biobank/sample/specimen data and customer record data.

Health Data Hub

The Health Data Hub (HDH) was set up in 2019 as a public interest group of 56 stakeholders to provide access to health data to improve quality of care and patient support. Together with Cnamts, the HDH works to make data available from the expanded SNDS for research, which means that the database contains information from nearly the entire French population of 66 million people. The HDH also has data from additional databases, including data on COVID-19, and plans to add other databases to the catalogue, including cohorts and other studies. The copies of the catalogue databases, pseudonymized so that all identifying personal information is removed, are stored and updated regularly on the HDH platform. The non-catalogue databases are accessible only during the project time frame. The HDH began in 2019 with 27 pilot projects; by December 2021 more than

4,600 projects were displayed on the HDH website (<https://www.health-data-hub.fr/projets>). Research results from completed studies are also presented on the website.

2.7 Netherlands (Kingdom of the)

The Dutch healthcare system is financed by insurance contributions and tax revenues. All residents must take out statutory health insurance with a private health insurance company. The contributions are collected centrally and distributed to the insurance companies on the basis of a risk adjustment formula. About 85% of the population also have private supplementary insurance that covers co-payments or non-reimbursed services (such as dentistry). A referral from the family doctor is necessary for outpatient specialist and inpatient care (“gatekeeping”). Although the insured do not have to register, most people register voluntarily with their family doctors. These are remunerated through a combination of per capita flat rates and performance-based payments. Specialist care is hospital-based and, depending on the specific contractual situation, is financed either on an individual service basis or through flat rates per case. All patients have a unique identification number (“burgerservicenumber” – citizen service number). Work is under way on a central information management system (Wammes et al., 2019); almost all general practitioners and hospitals use electronic health records, but interoperability needs to be improved.

NIVEL Primary Care Database

The NIVEL research institution operates a primary care database, called the Nivel Primary Care Database (Nivel Zorgregistraties eerste lijn), which brings together routine data on health problems and the use of primary care services for treatment of health problems in order to monitor the health and healthcare of the population (Schweikardt et al., 2016). The patient collective of the panel of participating practices is considered representative and comprises 1.7 million subjects (10% of the Dutch population); in addition to general practitioners, other primary care providers also participate, including GP out of hours services, physiotherapists, dieticians, exercise therapists and speech therapists. Participating healthcare providers receive regular reports with aggregated data and NIVEL uses the database for its own research

purposes; external researchers can also request the use of the data, which can be obtained in anonymized format. It is possible to link the data with other databases and sociodemographic information using pseudonyms based on the citizen service number.

PHARMO

PHARMO is an independent research organization that has built up a broad network of databases since its foundation in 1999. Here, depending on the research question or application, adapted, indication-specific “cohorts” can be created and microdata from several sources can be linked, including data from family doctors’ practices, outpatient and inpatient pharmacies, laboratories, cancer, mortality and neonatal registers, pathology findings, hospital discharges and patient-reported endpoints. PHARMO conducts its own research on specific issues, mainly with regard to pharmacoepidemiology and the effect of drugs after market access; external researchers can apply for the data for a fee and receive a customized dataset for processing.

2.8 New Zealand

New Zealand has a largely tax-funded health system that provides preventive, primary care, and inpatient and outpatient services, regionally administered through four districts. Although around one third of the population has private supplementary insurance, this is minimally used and primarily funds services not provided by the public system, or that are subject to lengthy waiting lists (such as elective surgery). Secondary and tertiary care is free of charge to all New Zealand residents. Individuals are encouraged to be registered with a GP /Primary Health Organization at birth, although this is not mandatory. Primary care is funded on a subsidized fee-for-service system, whereby general practitioners receive both out-of-pocket copayments from the individual and capitated government subsidies.

BOX 2.1 Health data in the IDI

- Results from the health check at pre-school age (4 years, “B4 School Checks”) – from 2011
- Cancer registrations – since 1995
- Chronic diseases – from 2007
- Accounting data of the family doctors – starting from 2002
- “Health tracker” (a kind of health census of the Ministry of Health) – 2006–2014
- Vaccinations – from 2006
- Laboratory settlements – from 2003
- Services for puerperium/newborn babies – from 2003
- Cause of death statistics – since 1988
- National Booking Reporting System – from 2003
- Need for care (SOCRATES)
- Billing data for outpatients – from 2007
- Regulations – from 2005
- Registrations with family doctors – from 2003
- Demographic information from the National Health Index – from 2004
- Mental health (only referrals, billed services and demographic information; no diagnoses or findings) – from 2008
- Private hospital discharges – from 2001
- Publicly financed hospital discharges – from 1988

Source: Stats NZ, 2020.

Integrated Data Infrastructure

A range of data from the healthcare system is held centrally, including individual-level data from hospital settings and from disease registers. This is accessible to public health organizations for service planning or delivery and is also held together with data from other areas (social services, education, income and employment, housing, justice, social conditions, population statistics), in the Integrated Data Infrastructure (IDI).

This research database is maintained by Statistics New Zealand. It was established to support the best possible use of existing data to answer complex questions.

Within data domains in the IDI, unique identifiers can be used to link data at the micro level (e.g., the National Health Index number for the health sector); to link data across domains, a probabilistic procedure is used (see Chapter 4). It is also possible to link other data that are not stored in the IDI, such as environmental data, by means of location. The IDI uses the so-called “five safes” framework and the Ngā Tikanga Paihere framework. Researchers who wish to have access to IDI data must adhere to the following: “safe people, safe projects, safe settings, safe data, safe output”; and the 10 tikanga (Te Ao Māori – Māori world concepts) which help establish goals, boundaries and principles to guide and inform data practice for researchers. The IDI data are already being used by several research groups (Milne et al., 2019). Among the case studies considered in the review, the IDI presents the highest degree of data integration.

Unique to New Zealand is the Treaty of Waitangi, the agreement between the New Zealand Crown and Māori representatives in the mid-nineteenth century, which made the protection and promotion of the rights and wellbeing of the Māori, among other things, a foundational governing principle of the country. The Treaty has implications and considerations with respect to Maori data sovereignty, including around public health and health data. This became particularly debated during the COVID-19 pandemic.

2.9 Republic of Korea

The Republic of Korean healthcare system is basically a national health insurance system; only the expenditures for 3% of the population are financed by taxes. National Health Insurance (NHI) receives income-based contributions from employers and employees or self-employed persons and finances the health expenditure of its insured persons accordingly. Although the range of benefits is relatively broad, co-payments for certain types of services can be high. More than 60% of the population have private supplementary insurance for co-payments and non-reimbursed services. Primary care is provided in healthcare centres, which are mainly private and are financed by individual payment for services. Physicians working in hospitals can also provide services for primary care of patients. In addition to individual services, acute-care

hospitals are remunerated via flat rates per case for seven diagnosis groups; nursing care services are financed via flat rates per day. There is no control function (“gatekeeping”) in primary care. The National Health Insurance Service (NHIS) collects the contributions and reimburses service providers, while the Health Insurance Review and Assessment Service (HIRA) reviews the billed services used, evaluates the appropriateness of the care provided and is responsible for designing the benefits catalogue and co-payments.

The National Health Insurance Service and Health Insurance Review and Assessment Service Database

As the central agencies for processing the billed services for the insured population of the Republic of Korea, NHIS/HIRA have extensive data on examinations, treatments, prescribed drugs, procedures and diagnoses of approximately 50 million people, in addition to their sociodemographic characteristics and information on the service providers involved. The HIRA data can be linked to other databases, so that missing information (findings, causes of death, disease severity, etc.) can be added. The HIRA data were only released for research purposes in 2009; since then, their use has increased, so that Republic of Korea is among the countries with the highest OECD index (see Table 1.1). Possible areas for research based on HIRA data include adherence, prescription patterns, benefit utilization, burden of disease, adverse events and policy evaluation. Although HIRA data include an “anamnesis” of outpatient and inpatient care of the insured from birth (see Chapter 3), only data from the last five years are released for research purposes.

2.10 Slovenia

The Slovenian healthcare system is financed through a combination of public and private sources, including employment-based SHI contributions, taxation and other government revenue streams, VHI premiums and out-of-pocket spending. SHI is compulsory for all residents, has near-universal coverage and is administered by the Health Insurance Institute of Slovenia (Zavod za zdravstveno zavarovanje Slovenije, ZZZS). More than two-thirds of the population have complementary insurance to protect against high co-insurance rates. Health policy and the management of the healthcare system are centrally

located at the national level, as is the administration of SHI. However, the organization and delivery of healthcare are decentralized. Primary care is primarily provided by a network of community-based primary healthcare centres, owned and managed by the municipalities, offering public health, primary and sometimes secondary care services by many different providers. Patients choose a personal physician in primary care from among four specialities – family medicine, gynaecology, dentistry and paediatrics – who serves as a gatekeeper. Hospitals are mostly located in regional centres. Overall, there is a capped annual national healthcare budget. Within this context, several payment mechanisms are used to remunerate services rendered: services in primary care are mainly paid by a mix of capitation and FFS payments, flat-rate and/or exclusively FFS. Secondary-level specialist care is paid on an FFS-basis (hospital-based) or by a combination of FFS and capitation. Acute hospital care is remunerated via a case payment model based on DRGs, while non-acute hospital care may be paid based on prospectively determined number of bed days, FFS or via case payments. Like the rest of the health system, governance of digital health solutions is centrally managed; the National Institute of Public Health (Nacionalni inštitut za javno zdravje, NIJZ) took over governance from the Ministry of Health (MoH) in 2015. There are more than 20 eHealth project solutions at various stages of implementation (NIJZ, 2018).

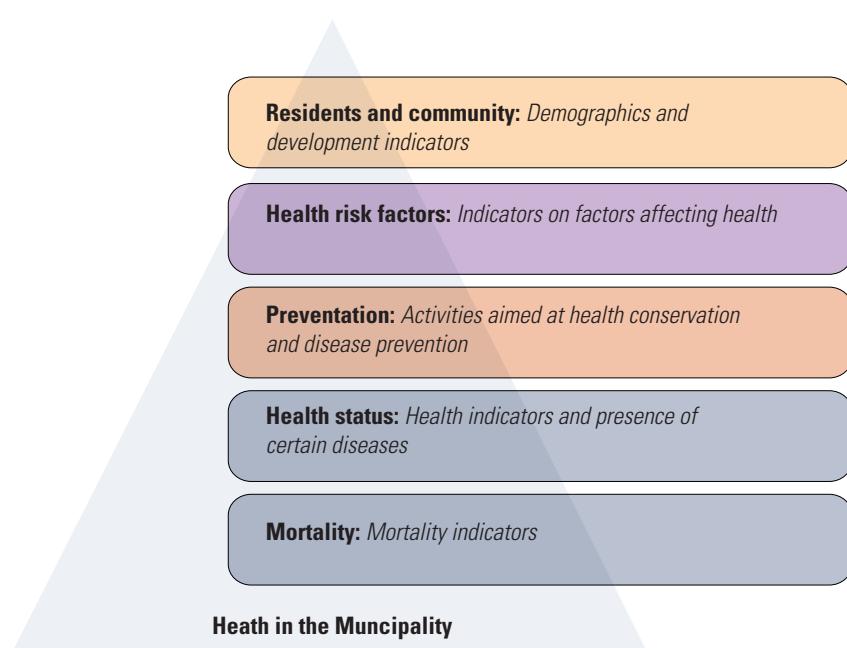
Health “All in One” Portal (Zdravje Vse na Enem Mestu Portal)

Rolled out in 2017, the patient data portal (Zdravje Vse na Enem Mestu, zVEM) portal serves as a connecting service for all essential eHealth solutions. It is one website for all health information and a real-time repository for the latest healthcare data for users of the Slovenian healthcare system as data are made available immediately after being entered. It does not collect data, but provides access to accumulate data from other sources, services and databases. NIJZ, in its capacity as eHealth authority, may save information and manages the portal. Databases include the Central Registry of Patient Data (Centralni register podatkov o pacientu, CRPP), which contains data on demographics, health and healthcare utilization, e.g., exam reports, diagnostics, discharge letters, lab reports, patient summaries and COVID-19 test results, and

operates like an EHR in Slovenia; e-Prescriptions (including dispensations), e-Appointments (including e-Referrals), data from the e-registry of vaccinated persons, telestroke and teleradiology are also included (NIJZ, 2018). All data are physically located in the two national eHealth data centres in Slovenia – in Ljubljana and Maribor – and records are active for five years before being archived (they remain accessible). Linkage occurs at NIJZ but can also be conducted by institutional healthcare providers. Data can be linked via several identifiers, including the unique health insurance number/identifier, which every resident in Slovenia has, or a personal ID number. Patients and providers authenticated by digital certificate have full access to original data records in any of the eHealth databases. Others have access to specific data according to their legal standing and purpose of request (e.g., private citizens versus representatives of MoH or a public health authority). Researchers may request data and access is provided after review from the ethics board or data commissioner and upon obtaining an appropriate digital identity. Patients, however, have the right to forbid access to patient summary records.

Health in the Municipality (Zdravje v občini)

For each Slovenian municipality the NIJZ provides insight into the health status of the population. The project, Health in the Municipality (HiM), was launched in 2016 and presents various health and social indicators categorized into different thematic strands (see Figure 2.3), including socioeconomic, demographic, health status (e.g., risk factors, mortality) and selected environmental and other data impacting on the living conditions of populations, aggregated at the municipal and European NUTS3 region level. Data available through this project are collected and contained within multiple databases managed by NIJZ and linked to the repositories of seven other organizations (see Appendix I). To achieve insights, data from different sources are linked by unique identifiers, including personal ID numbers and health insurance numbers. The portal is refreshed annually, and the latest available data are from 2021, though the majority of data covers 2015–2020 (as of December 2021). Indicators are stored permanently. Presented as structured data and often in ready-made graphics and tables, they are used for analytical decision-making and research to improve healthcare, design new services and understand differences among regions.

FIGURE 2.3 Indicators included in HiM

Source: Authors' own, based on NIJZ, 2016.

Health Data Portal (Podatkovni portal)

NIJZ's Health Data Portal (Podatkovni portal) is the main platform for health reporting and health system planning in Slovenia. It is a single point of access to all health-related databases – an ever-expanding data repository – showing information on health, the health system and impacts on population health. Most data are collected by NIJZ, but links to data from nine other institutions are provided (see Appendix I). Altogether, there are data covering health status, socioeconomics, health service utilization, resources and consumption, including prescribed medicines, and expenditures as well as links to HiM, the environmental agency and the cancer registry. Data linkage across these sources is performed by NIJZ. The main advantage of the portal is the preparation and storage of data in the desired scope and format of display for anyone to access. The data are ranked by individual thematic strands. Content is prepared at national level. The full database is updated six times a year, open to anyone and stored permanently. The latest available data are from 2020. Presented

as structured data, they are used for analytical decision-making and research purposes to understand the dynamics of health and healthcare at different territorial levels and other aggregated levels.

2.11 Sweden

In Sweden it is the 21 counties (12 county councils and 9 regional authorities) that are responsible for providing healthcare. Local authorities, on the other hand, are responsible for healthcare and social services. Income for the system comes from local and regional taxes; a system for the regional redistribution of income allows the regions to meet the needs of their inhabitants. Only about 10% of the population have private supplementary insurance packages for faster access to specialized medical care and non-reimbursed services (optional services). Primary patient care is provided by private (40%) or public (60%) service providers (general practitioners, nurses, midwives, psychologists, physiotherapists and gynaecologists, often organized in group practices). Registration is required in most regions, but family doctors do not have a strict gatekeeping function; patients usually register with a (shared) practice rather than with a doctor. Service providers in primary care are remunerated through a mixture of mainly capitation, individual service remuneration and performance-based payments. Outpatient and inpatient specialist care takes place in public hospitals or private clinics; it is remunerated via flat-rate payments per case (DRGs), which are supplemented by performance-related components and cost or volume covers. Similar to Denmark and other Nordic countries, Sweden has well-established epidemiological and clinical registers.

Registerforskning

In order to enable research based on clinical registries using multiple datasets, the Swedish Research Council has established the Registerforskning unit to help researchers find and process the right data. An electronic tool (Register Utiliser Tool, RUT) has been developed to further simplify this process. All registers in the tool are structured according to the Generic Statistical Information Model, but the tool will not be able to access microdata directly. The registers that can be accessed via Registerforskning include data on health and healthcare, tax, education, justice, sociodemographics, etc. The personal

identification number (Personnummer, PN) can act as an identifying feature for linking data across multiple registers.

2.12 United Kingdom

The National Health Service (NHS) in the United Kingdom was established with the aim of ensuring that healthcare meets the needs of all residents, is responsive and independent of patient ability to pay, and is free of charge at the time of use. Since its foundation in 1948, the system has been developed along these basic principles. It consists of four separate organizations for the different parts of the UK (NHS England, NHS Wales, NHS Scotland, Health and Social Care Northern Ireland), which operate independently but follow the NHS Common Basic Principles (UK Government, 1948). The system is funded by taxpayers' money and is known for the central role of GPs in primary care, acting as an entry point for patients and "gatekeepers" for further care in the system. In England 42 integrated care systems (ICSs) were established in July 2022, replacing clinical commissioning groups (CCGs), with the responsibility for planning and providing care. In Scotland there are 21 NHS Boards responsible for delivering this function. In England GPs are funded through a mixture of capitation, individual performance payment and performance-related payments, while hospitals are remunerated through per-case flat rates, with separate budgets for mental health, education and research. In Scotland the so-called "purchaser-provider-split" has not been implemented, so that the remuneration of services is more integrated.

Clinical Practice Research Datalink

The important role of general practitioners is also reflected in the internationally renowned database Clinical Practice Research Datalink (CPRD), which brings together data from GPs' practices in the UK on the basis of the respective electronic practice information systems. CPRD is supported by the Medicines and Healthcare products Regulatory Agency (MHRA) and the National Institute for Health Research (NIHR). The CPRD collects de-identified microdata from a network of GP practices that have decided to participate and that may opt out at any time (see Chapter 4). The data collected

include demographic information, information on service providers, clinical presentation (diagnosis, symptoms, clinical signs), referrals, vaccinations, laboratory findings and medication prescriptions. Data from English practices can be linked to other datasets so that the process of care beyond primary care can be researched; practices must have agreed in advance to participate in the linking programme. The datasets that can currently be used for linking are listed in Box 2.2. NHS Digital, which is not part of NHS England's transformation directorate, is responsible for data linkage; the personal NHS

BOX 2.2 Databases that can be linked to CPRD data

- Hospital Episode Statistics (HES) data for/from:
 - Inpatients
 - Outpatient services
 - Emergency departments
 - The Diagnostic Imaging Dataset
 - Patient reported outcomes measures (PROMs)
- Death registrations from the statistical office
- Cancer data from NHS Digital National Disease Registration Service (NDRS) (formerly Public Health England (PHE)):
 - Cancer registrations
 - 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)
- Mental Health Services Data Set (MHSDS)
- Small-scale socioeconomic: Index of Multiple Deprivation (IMD), Townsend Deprivation Index, Carstairs Index, Patient postcode-linked deprivation measures, Practice postcode-linked deprivation measures, Rural-Urban Classification
- COVID-19 data: Second Generation Surveillance System (SGSS) positive virology test data, COVID-19 Hospitalisation in England Surveillance System (CHESS), Intensive Care National Audit and Research Centre (ICNARC) data on COVID-19 intensive care admissions.

number can be used as a unique identifying feature, but other features are also used depending on the data to be linked (see Chapter 4). As of December 2021, the CPRD data cover 60 million patients and have been used for more than 2700 peer reviewed scientific publications that have researched issues such as drug safety, drug use, the effectiveness of health policy measures, the organization of service delivery and the development of risk factors (MHRA 2022b).

Electronic Data Research and Innovation Service

In recent years Scotland has taken a number of strategic steps to develop and use routinely collected data. The Electronic Data Research and Innovation Service (eDRIS) provides a “one-stop shop” for the use of existing administrative and health data. It is designed to help researchers and policy analysts identify the right datasets for their projects, operationalize research questions and gain access to linked data. All publicly owned administrative data can be requested for research, planning or evaluation processes in the healthcare system. The linking is done by a trusted third party based on a so-called “Population Spine” (see Chapter 4), which contains the master data of all persons who have come into contact with NHS Scotland.

Scottish Primary Care Information Resource

Like the CPRD (see above), the Scottish Primary Care Information Resource (SPIRE) was established in Scotland to bring together data from Scottish GP practices for research and analysis purposes. These can be linked via eDRIS with other datasets of the Scottish NHS. Among other things, SPIRE data will be used to determine the quality of care and inform the GMS Contract (see the completed Quality and Outcomes Framework: ISD Scotland, 2016). SPIRE data will be extracted on request from the practice information system of participating practices (the software is provided and operated by MSDi) and will include sociodemographic data of patients, findings, procedures, drug prescriptions, diagnoses, utilization of services. Only structured information can be extracted for research purposes, never the entire electronic health record; thus, for example, notes from doctor-patient discussions are not passed on.

Potentially identifying data must be pseudonymized before they leave the practice. All practices in Scotland participate in SPIRE, unless they decide to opt out.

2.13 United States

The United States' healthcare system is known for its fragmentation in terms of financing and organization of service delivery and the lack of universal access to healthcare services. Two-thirds of the population have private health insurance with varying catalogues of services and co-payment requirements. The majority of privately insured persons are covered by employer-financed plans (56% of the population in 2017; see Commonwealth Fund, 2019). Approximately one-third of the population are covered by public social security systems; the most important are Medicare for those over 65 (17% of the population) and Medicaid for low-income earners (19% of the population). The 2010 Affordable Care Act introduced the individual purchase of health insurance packages for all otherwise uninsured persons, and this has contributed to a decrease in the uninsured rate (8.8% in 2017). The mandate has since been abolished. In primary care physicians are mainly active in small individual or group practices; they usually have no “gatekeeping” function and patients have free choice if the service providers belong to their insurance company's network. Remuneration varies depending on the insurance system (Rice et al., 2020). In primary care it mainly comprises individual payments (with regulated or negotiated fees, for publicly or privately financed services accordingly), capitation and performance-based incentives (in some insurance systems). Specialist care takes place in doctors' offices and hospitals. Large companies offer integrated care systems that bring together health insurance, outpatient and inpatient healthcare in the “managed care” model. The largest of these is currently Kaiser Permanente, with branches in eight countries and more than 12 million insured persons.

Similar to the structure of the system, the availability of data for research is also fragmented; nevertheless, there is no lack of initiatives aimed at transparency and facilitating research and analysis activities. However, while the US has many data sources for healthcare use and at least one good one for costs, there is very limited access to information from electronic health records for researchers or the public (with the exception of the Veterans Affairs system).

Centers for Medicare and Medicaid Services Data

The Centers for Medicare and Medicaid Services (CMS) are part of the Department of Health and Human Services (DHHS) and maintain more than 400 records (CMS Data) on public healthcare in the American system. These include socioeconomic and demographic data, and health and healthcare data at the individual and macro levels. Many of these datasets are freely accessible on an aggregated level as “public use files (PUFs)” via data.cms.gov. De-identified microdata are available on request either as “limited datasets (LDS)” or as “research identifiable files (RIFs)”; the latter can be customized according to the research question before they leave the institution and their use is much more cost-intensive compared to LDS (see Figure 3.1). The CMS databases contain several identifying characteristics that can be used for linking, but CMS does not publish a data dictionary (see Chapter 4).

Kaiser Health Connect System

Based on the electronic health records of the patients entered at Kaiser Permanente, the Kaiser Health Connect System brings together information from different levels: clinical findings, examinations, diagnoses and treatments with appointment bookings, registrations and billing data. Kaiser Permanente uses Epic, the largest EHR systems in US, which contains records on approximately half of the US population (Shull, 2019). The Kaiser Health Connect System is the largest private clinical information system in the world. It offers two interfaces: MyChart, where patients can view their information, and EpicCare, which is used by healthcare professionals to coordinate patient care. Kaiser Permanente has its own research team and uses the linked data available in Kaiser Health Connect to find the best ways to deliver healthcare to its members.

National Patient-Centered Clinical Research Network

The 2010 Affordable Care Act established, among other things, the Patient-Centered Outcomes Research Institute (PCORI). PCORI leads the PCORnet (National Patient-Centered Clinical Research Network) initiative, which

aims to simplify and optimize the conduct of clinical research based on large datasets and strong participatory networks by addressing patients' needs more explicitly. PCORnet represents an umbrella network of so-called "Research Networks"; in 2019 nine of these networks were located at clinical care systems and two at insurance organizations; the patient-driven networks, which were originally equally integrated, have been used in an advisory capacity from April 2019. PCORnet operates via a Common Data Model (CDM), which is the key to the preparation of data from participating partners. The PCORnet Coordinating Center manages the so-called "Front Door", which allows potential data users to be advised about the possibility of PCORnet data and to submit applications for use. The PCORnet data includes demographic variables, duration of observation in the network, contacts with the healthcare system, diagnosis, procedures and diagnostic tests, vital signs, prescribed drugs, laboratory findings, health status (diagnosed and self-reported), patient-reported endpoints, deaths and causes of death, information on the service providers involved, and, if applicable, data from clinical studies on the patients included.

Surveillance, Epidemiology, and End Results-Medicare initiative

The Surveillance, Epidemiology, and End Results-Medicare (SEER)-Medicare linking initiative links the data on cancer incidence, staging and survival rates collected by the SEER programme's cancer registries to the Medicare programme's accounting data (see above). The SEER registries are funded by the National Cancer Institute and represent the most comprehensive population-based oncology dataset in the world. Data include demographic information, tumour-related data including identification and staging, and care-related billing data. Possible research questions that can be answered using the linked data include the care given to patients before and after a cancer diagnosis (immediate and long-term perspective), the use of diagnostic tools and procedures, and the costs of cancer care.

CancerLinQ

In the field of oncology CancerLinQ sees itself as a “big data analytics platform”, which aims to improve the quality of care based on data from the daily care of cancer patients. CancerLinQ is a non-profit subsidiary of the American Society of Clinical Oncology (ASCO) and is thus managed by oncologists; clinicians who subscribe to the database can access patient-related information. Only de-identified data are available for research, which is provided according to the project proposal. CancerLinQ collects the following data from more than 100 participating practices: patient and practice information, diagnoses, pathology findings, staging, vital signs, clinical evaluations, laboratory findings, radiology findings and imaging, treatment plans, prescribed and locally administered drugs, radiotherapy treatments, surgeries, other outcomes, and participation in clinical trials. There are also unstructured data that require human curation assisted by natural language processing. The CancerLinQ database itself contains linked data and CancerLinQ subscribers can view identifiable patient information in the clinical database, but data may not be linked to other datasets.

2.14 Summary of case studies presented in this chapter

The case studies presented above can be roughly divided into five categories:

- i. **large administrative datasets**, which, due to a central administration point, make it possible to map questions about the care process of individual patients across sectors and offer possibilities for linking to other data (e.g., MBS/PBS, NHIS/HIRA, ÖGIS, CMS)
- ii. **centralized locations that merge existing datasets** of different data holders and with variable content; here, two possibilities stand out: the data remain with the data holders and are linked on request (e.g., HDH, PHRN, SDLE, eDRIS) or are stored in a centralized location (e.g., IDI);

- iii. **databases which combine routinely recorded data** and make it available to clinicians and researchers (e.g., CPRD, NIVEL PC Database, SPIRE, Kaiser Health Connect);
- iv. **patient-centric electronic platforms** that allow (inter alia) patients to gain access to their own information and which could potentially be considered as a possible basis for clinical data for research (e.g., ELGA, My Health Record or Sundhed.dk)
- v. **tools that facilitate research** across multiple datasets (e.g., RKKP, Registerforskning).

TABLE 2.1 Overview of case studies and their contents (data categories correspond to the blocks in Figure 1.1)

Country	Case Study	FURTHER									
		MACRO DATA					MICRO DATA				
		SD	Health data	Health care	SD	Health data	SD	Health data	Health care	E	-
France	HDH	Access to National Healthcare Data System databases and other selected databases and studies	✓	✓	✓	✓	-	-	-	✓	✓
New Zealand	IDI	Research on multidimensional and complex issues	✓	✓	✓	✓	-	-	-	✓	✓
Netherlands	PHARMO	Research: Use of "Real World Evidence", promotion of innovation in healthcare	-	✓	✓	✓	-	-	-	✓	✓
Slovenia	zVem Portal	Nivel PC Database	✓	✓	✓	✓	-	-	-	✓	✓
Republic of Korea	HIM	NHIS/HIRA	✓	✓	✓	✓	-	-	-	✓	✓
Zealand	PP	Provides access to individual health data for various research purposes	✓	✓	✓	✓	-	-	-	✓	✓
France		Provides selected health and social indicators at the regional level: (1) population and community, (2) risk factors, (3) prevention measures, (4) health status, (5) mortality	✓	✓	✓	✓	-	-	-	✓	✓
Slovenia		Various research purposes (analytical and decision-making) at different territorial levels	✓	✓	✓	✓	-	-	-	✓	✓

Country	Case Study	Primary goal	MICRO DATA			MACRO DATA			FURTHER		
			SD	Health data	Health care	SD	Health data	Health care	E	-	-
Sweden	Registerforskning	Supporting research by providing information on register data (health, tax, justice, education, etc.)	✓	✓	✓	-	-	-	-	-	-
	OPRD	Research on primary care and possibility of linking to other data	✓	✓	✓	-	-	-	-	-	-
	UK -	eDRIS SPIRE	Research and policy analysis: Information on and provision of (linked) data Research and care optimization through the consolidation of data from primary care	✓	✓	✓	✓	✓	-	-	-
UK - Scotland	CancerLinQ	Research – consolidation of observational data ("real world cancer data") from several sources	-	✓	✓	✓	-	-	-	-	-
	CMS Data	Transparency for supply management and research	✓	✓	✓	✓	✓	✓	✓	-	-
	Kaiser Health Connect	Improving clinical care for individual patients, research to improve future care overall	✓	✓	✓	✓	✓	✓	-	-	-
	PCORnet	Research – central point of contact for further decentralized data from service providers, health insurers and so-called patient-powered research networks	✓	✓	✓	✓	✓	✓	-	-	-
US	SEER Medicare	Research on the care of cancer patients	-	✓	✓	✓	✓	✓	-	-	-

Note: SD = Socio Demographic, E = Environment. * Some of the regional linking units also use environmental data.

3

Expanding the possibilities of routinely collected data

This chapter brings together insights from those case studies presented in Chapter 2 that allow for the secondary use of comprehensive, intersectoral data collected during the course of healthcare provision. This includes large administrative datasets that link data from different sectors to fulfill their primary purpose; the routine collection of health and healthcare data from different providers into one repository for secondary use with the potential to link to other (routinely collected) data; and the linking of existing datasets, which are available in a decentralized format.

3.1 Datasets that bring together data from different sectors to fulfill their primary purpose

When all information on the interaction of individual patients with the healthcare system is collected centrally, there is a higher likelihood that comprehensive data will be available that allow for processes of care provision that span service providers and sectors to be mapped (over time). For healthcare data this is perhaps easiest in single-payer systems. Three of the case studies (Australia, the Republic of Korea and the US) emphasize the possibility of using existing, integrated routine data for research questions. This refers to datasets that are already created for the primary purpose in an integrated (linked) form across several service areas. These are on the level of individuals covered by the insurance/social security system in question, are available over several years and cover various service areas. The generation of these datasets

is done by the respective government or health insurance fund and the data are primarily used for accounting purposes or internal analyses. Datasets that are only generated or linked for research purposes (secondary use) are to be distinguished from these datasets (and are presented in the next section). As a rule, these are datasets that consist mainly of accounting data (claims data) (see Table 3.1).

In Australia the two large datasets of the MBS and the PBS contain information on all benefits provided that fall under MBS or PBS. Since almost the entire Australian population has access to healthcare through MBS/PBS, and a large proportion of benefits are fully or partially covered, the datasets can be considered “complete”. A comparison with primary data has shown that PBS billing data are a good alternative for research into the provision of care (Harris et al., 2017), although they tend to underestimate usage (Gisev et al., 2018). In addition, both datasets contain the Medicare ID, which allows for links between the two datasets as well as to other data (see PHRN, below). Access to the data are clearly regulated for research purposes and the data are used accordingly for scientific analysis.

In the Republic of Korea all accounting data for over 50 million insured persons flow into the database of the national insurance agency (see Chapter 2). These data are generated in the course of reimbursing service providers: invoices are transmitted electronically from service providers to HIRA, where they are checked. The settlement data of reimbursed claims are stored in the “Data Warehouse” for five years. Access for research purposes was established in 2009. As of December 2019, research projects can request data from the previous five years, although an extension to ten is planned. Identification numbers of insured persons and service providers are stored in encrypted form to protect personal information. Since HIRA data are limited to billing data (i.e., healthcare data), individual research projects so far have had to use and link other datasets (e.g., the national health surveys with health data) to answer broader questions. For example, linkage has been done using regional variables or based on several variables that were present in both datasets (e.g., date of birth, gender, etc.) (see examples in Table 4.1). Microdata are available to researchers in universities and government agencies, and to a limited extent to interested parties from industry and the private sector, including private insurance companies (see Chapter 5). This data accessibility to for-profit insurers and other commercial purposes is not without controversy (see e.g., Song, 2017).

Similar access possibilities are available for the routine data of public insurance programmes in the US (see Figure 3.1). The CMS data are described here as an example (see Chapter 2 on the context in the healthcare system). CMS allows researchers to access Medicare billing data, which includes contacts with the healthcare system and therapeutic treatments including drugs, procedures and other services. An overview of studies conducted with Medicare data from 1979 to 2016 shows a significant increase in use, especially in the years after 2010 (in 2016 alone there were 205 publications; Mues et al., 2017). The issues under investigation include the study of patterns of morbidity, mortality and disease burden, the comparative assessment of the effect of drug therapies, the mapping of the costs of different care pathways, the analysis of the effects of the behaviour of healthcare providers on care and the influence of health policy measures on care providers and patient outcomes. Thus, these data provide a comprehensive insight into the care of older patients. In order to provide access to the (correct) data and to support the optimal use of the existing datasets, CMS cooperates with the Research Data Assistance Center (ResDAC), which serves as a gateway and signpost for interested parties (ResDAC, 2022). ResDAC is also responsible for reviewing applications for scientific feasibility (see Chapter 5). Despite the broad coverage of the population over 65 in the Medicare programme, the Medicare data do not reflect the entire collective of associated patients; there are specific insurance options that are not included. Furthermore, similar to other billing data described here, these data are limited by the fact that they reflect only billed and thus reimbursed services.

TABLE 3.1 Scope of datasets bringing together information from different sectors and examples of studies using their data

COUNTRY	DATASET	WHAT DO THE DATA COVER?	ACCESS FOR RESEARCH/ LINKABILITY/ AVAILABILITY	EXAMPLES OF PUBLISHED ANALYSES
Australia	MBS	Information on all services provided that are fully or partially covered by the Medicare Benefit Schedule, i.e., a large part of outpatient and inpatient care Dental services, for example, are excluded from Medicare. Includes both current and archived data	Yes Linking with other databases possible	i) Description of the use and costs of paediatric care and regional comparison (Freed & Allen, 2018) ii) Time course of autism diagnoses by gender (May & Williams, 2018) iii) Incidence, regional variation and trends in keratinocytic neoplasia (Adelson et al., 2018)
	PBS	Information on all services provided that are fully or partially covered by the Pharmaceutical Benefit Schedule The PBS data can also be linked to the MBS data via the individual Medicare ID	Yes Linking with other datasets possible	i) Increase in the use of anticoagulants and costs after new drugs enter the market (Morgan et al., 2018) (ii) MBS and PBS data: Prediction of initiation of second-line treatment for diabetes in the Australian population (Fiorini et al., 2019)
Republic of Korea	NHIS and NHIDB	Treatments, prescribed drugs, procedures, diagnoses, examinations, sociodemographic characteristics of patients	Yes Linking with other datasets possible (based on individual projects)	i) Monitoring the use of antibiotic usage patterns from 2007 to 2014 (Park et al., 2017) ii) Influence of insurance coverage on use of anticoagulants (Ko et al., 2018) iii) Environmental pollution and pregnancy hypertension (linked data, Choe, Jun & Kim, 2018; for the structure of the NHIS cohort used, see Lee et al., 2017) iv) Linking clinical registry data (stroke) and HIRA data (Kim et al., 2018)
US	CMS data	Contacts with the healthcare system, therapeutic treatments including drugs, procedures and other services	Yes Linking with other databases possible (e.g. SEER-Medicare, individual projects)	i) Differences in costs between university and other hospitals (Burke et al., 2019) ii) Effect of annual screening on screening rates, referrals, utilization and costs (Ganguli et al., 2019)

Source: Authors' own.

FIGURE 3.1 Overview of CMS data formats according to ResDAC

	Public Use File (PUF)	Limited Data Sets (LDS)	Research Identifiable File (RIF)
Level of detail	aggregated data	De-identified individual level data	Individual level data
Budget	free		
Level of customization	none	none	customizable
Time to obtain	0–2 weeks	2–3 weeks	3–5 months
Storage needed			

Source: Adapted from ResDAC.

3.2 Merging of routinely generated data into a new database for secondary use

The previous section highlighted possibilities of using healthcare data. When it comes to health data, the same principle applies: the likelihood of comprehensive information being readily available rises if there is a central point of data collection within the system. This is perhaps easiest in healthcare systems that include a gatekeeping function in primary care, such as in the UK. Chapter 2 and the country profiles in Appendix I provide country-specific information on care purchasing and gatekeeping.

Among the case studies included, the most typical example of a database that brings together data collected by GPs is the CPRD (for the history of its origin and embedding in the British system, see also Gothe, 2014). CPRD brings together data from practices that have enrolled in the programme (“opt-in”). Patients in these practices can be excluded from the data transfer (“opt-out”), in accordance with the national opt-out regulation of the NHS. The data held by CPRD can be requested for observational studies or can be used to support experimental studies. They can also be linked to other NHS Digital datasets to provide richer, cross-sectoral insights (see also the next

section and Chapter 5). As a result, more than 2700 peer reviewed scientific studies have already used CPRD data.

In Scotland SPIRE was established in 2016 to bring together data from Scottish practices and enable Scottish GPs to compare and plan care better, and to help the Scottish NHS monitor and improve the quality of care. Again, patients can opt out of having their data used by SPIRE. The initiative emphasizes that SPIRE is not a comprehensive central database of all information on Scottish patients but aims at facilitating the target-specific aggregation of parameters from GP data for specific, approved purposes. Data will not be stored for longer than necessary to process applications and information will be destroyed afterwards. The available information on SPIRE is still new and not complete (see SPIRE Rollout, SPIRE, 2019).

The database on primary care at the Dutch research institute NIVEL also uses data from primary care, but here the data from the electronic practice management systems also come together from service providers other than GPs (see Chapter 2). In the Kingdom of the Netherlands this model is also used by other institutions or initiatives (see, for example, Smeets et al., 2018). What all initiatives have in common is that, in addition to the research opportunities, the data are also explicitly used for feedback to the participating practices, including the possibility of comparing one practice's results with those of other practices.

3.3 Linking of independently created datasets

Several of the identified case studies link different data sources with different content using one or more linking variables. In contrast to the datasets described in the previous sections of this chapter, this means that the data are not already collected in an integrated manner for the primary purpose of use or are merged to form a new dataset but must first be linked for the secondary purpose of use (in this case for research). This may be necessary either because the data for the primary purpose only cover a very limited area or because an additional extension to an already extensive dataset is necessary in order to answer research questions accordingly (e.g., linking healthcare data with occupational/unemployment data or clinical data with accounting data).

TABLE 3.2 Merging of data in primary care

COUNTRY	INITIATIVE	COVER	OPT-IN/OPT-OUT	ACCESS FOR RESEARCH	EXAMPLES
United Kingdom	CPRD	Approx. 738 practices in 2018 (Wolf et al., 2019)	Patients can opt out of data sharing; GPs voluntarily participate in the transmission to CPRD. Likewise, the consent of the practice is required for linking the data	Access is possible for scientists, legislators/NHS employees and others (including industry)	>2300 Articles (CPRD bibliography) i) Changes in the causes of death of COPD patients, 2005–2015 (Gayle et al., 2019) ii) Comparative benefit assessment in COPD drug therapy (Suissa, Dell'Aniello and Ernst, 2019)
United Kingdom (Scotland)	SPIRE	Rollout in implementation	Opt-in for practices; patients can exclude the use of their data	Yes, on application basis (planned)	None yet
Netherlands (Kingdom of the)	NIVEL primary care database	Representative sample of about 10% of the population	Opt-in for practices; opt-out for patients	Yes, on application basis; NIVEL own research	i) Use of non-scheduled primary care during the nursing reform according to income (Jansen et al., 2019) ii) Use of email counselling in primary care (Huygens et al., 2018)

Source: Authors' own.

The overarching goal of the initiatives covered here is to provide better data access for research. This will involve linking data from different sources, especially from different public institutions. For example, data on healthcare can be linked with data from the education sector and crime statistics. In almost all initiatives, the possibility of data access is regulated through application procedures (see Table 3.3). An overview of the research conducted on the basis of the data or of all positively evaluated applications is also provided by

almost all initiatives on the corresponding homepage, so that a high degree of transparency is achieved, and it is also clear to the general public how the data are used.

Both the corresponding information in Chapter 2 (Table 2.1) and the examples listed in Table 3.3 clearly highlight the potential of centralized linkage units for combining independently created datasets. The extent of consolidation of the different datasets varies among the case studies, which affects the linking process. For example, the HDH, PHRN and the SDLE link the data upon request (project-specifically), while the source datasets remain basically distributed. In contrast, the “integrated” IDI merges incoming data by means of a “spine” (see Chapter 4). An interesting aspect of the PHRN’s approach, is that the regional authorities at the state level have their own linking agencies, which cooperate in the PHRN’s network, but also perform their own linking. Additionally, some data within PHRN is routinely linked (e.g., birth registers, perinatal data (pregnancy, birth, etc.) and death registers), while other data are linked on a project basis. In general, a single entity is responsible for carrying out the data linkage, which is most often the population statistics agency within a country or region. Case studies such as eDRIS in Scotland and PHARMO in the Kingdom of the Netherlands, and the new Findata Authority in Finland are not displayed in Table 3.3, but in principle support similar considerations.

Slovenia’s NIJZ is the sole data controller and main performer of data linkage for two case studies that merge and link independently collected datasets for research purposes. The HiM database, for example, merges data from separate databases across different sectors to provide access information on population health, healthcare utilization and environmental factors aggregated at different levels: national, statistical region, administrative unit and municipality. These data are meant for researchers, policy-makers and other individuals to encourage health promotion and prevention activities (NIJZ, 2016). Similarly, the Health Data Portal links independently gathered data related to health and healthcare via unique identifiers and presents data aggregated at different levels for researchers and interested parties.

TABLE 3.3 Models with a central location linking distributed datasets (different data holders)

COUNTRY	CASE STUDY	WHAT IS LINKED?	WHERE ARE THE DATA LOCATED?	WHO CARRIES OUT THE LINKAGE?	OPPORTUNITIES FOR RESEARCH	EXAMPLE OF PUBLISHED ANALYSES
Australia	PHRN	A distinction is made between data that are routinely linked and data that can be additionally linked. There are differences in the additional data by responsible region (jurisdiction).	The data are located at the respective institutions.	Depending on the requested datasets and the responsible region (jurisdiction), there is a responsible Data-Linkage Unit (DLU). The DLUs are all part of the PHRN.	Access for research is clearly regulated and runs through an application procedure. In addition, there are fees for linking or data use.	Overview of individual studies: PHRN research publications (Australia, PHRN, https://www.phrn.org.au/publications/research-publications/).
	SDLE	Routinely linked data from: birth registers, perinatal data, death registers, cancer registers, mental health, public emergency departments, private and public hospital cases.	The data are stored in a secure environment of Statistics Canada; there is a separation of the DRD, the source data files and the key registry (see Chapter 2).	Statistics Canada.	Yes, access to some linked data or data linkage can be requested.	Interesting examples: i) Suicide risk after imprisonment in Queensland (Spittal et al., 2014). ii) Improving the estimation of mortality ratios in hospitals (Spilsbury et al., 2017).
Canada		Data are available from the Labour Force Survey, national health surveys, hospital discharges, census data, environment-related macro data and a range of provincial data on crime, education, mental health, etc.				Overview of positively evaluated applications: Statistics Canada, Record (https://www.statcan.gc.ca/en/record/2020).

COUNTRY	CASE STUDY	WHAT IS LINKED?	WHERE ARE THE DATA LOCATED?	WHO CARRIES OUT THE LINKAGE?	OPPORTUNITIES FOR RESEARCH	EXAMPLE OF PUBLISHED ANALYSES
Canada - Ontario	ICES linked data	Data from various administrative sources, survey data and register data are linked. Only the population of Ontario is represented in the data.	The data are stored on ICES servers in a closed computing system.	ICES	Access for researchers outside ICES has been possible since 2009 and is subject to a successful application.	<p>Overview of publications: ICES published journal articles (ICES, Publications, Journal Articles, https://www.ices.on.ca/Publications/Journal-Articles).</p> <p>Interesting examples:</p> <ul style="list-style-type: none"> i) Access to mental health services for care patients (Perlman et al., 2019). ii) Air pollution and risk of atrial fibrillation and stroke (Shin et al., 2019).
Canada - British Columbia	PopData BC	Individual-level, de-identified and longitudinal data on the population of British Columbia. Data come from various "partners" (including cancer registries, Department of Health, Statistics Canada). Own datasets of the applying researchers can also be linked.	Data are stored on the central server ("Secure Research Environment").	PopData BC	Yes, access is regulated for research. There is an application system.	<p>List of research projects: PopData BC projects (https://www.popdata.bc.ca/ra/projects).</p> <p>Interesting example:</p> <ul style="list-style-type: none"> i) Newborn outcomes in fathers on medication for multiple sclerosis (Lu et al., 2014).

COUNTRY	CASE STUDY	WHAT IS LINKED?	WHERE ARE THE DATA LOCATED?	WHO CARRIES OUT THE LINKAGE?	OPPORTUNITIES FOR RESEARCH	EXAMPLE OF PUBLISHED ANALYSES
Canada - Manitoba	Manitoba population research data repository	Administrative, survey, registry and other data will be linked for the inhabitants of Manitoba. Data come from various organizations.	Data are stored on the Manitoba Centre for Health Policy servers.	Manitoba population data research data repository.	Yes, access is regulated for research. There is an application system.	<p>Interesting examples:</p> <ul style="list-style-type: none"> i) Child health and equity (Nickel et al., 2014). ii) Risk of autism with prenatal exposure to antibiotics (Hamad et al., 2019). iii) Relationship between the social determinants of health and quality of care (Katz et al., 2018).
France	HDH	The HDH is authorized to make data available from the expanded SNDs (National Health Data System). The HDH also has data from a catalogue of additional databases.	Data are stored on a Microsoft Azure Cloud, physically based in the Paris region.	A third party (e.g., National Health Insurance Fund (CNAM)).	Yes, access is regulated for research. There is an application system.	Overview of projects (HDH, Projects, https://www.health-data-hub.fr/projects).

COUNTRY	CASE STUDY	WHAT IS LINKED?	WHERE ARE THE DATA LOCATED?	WHO CARRIES OUT THE LINKAGE?	OPPORTUNITIES FOR RESEARCH	EXAMPLE OF PUBLISHED ANALYSES
New Zealand	ID	Contains microdata of New Zealand residents and households. Some data are longitudinal and de-identified. The data come from various ministries, surveys and non-governmental organizations.	Data are stored on the servers of the IDI, with >166 billion pieces of information as of September 2018 (Mline et al., 2019)	Statistics NZ	Yes, access is regulated for research. There is an application system.	<p>Overview of all research projects: Stats NZ, Storehouse (https://cdm2045.contentdm.oclc.org/digital/collection/p20045coll1).</p> <p>Interesting examples:</p> <ul style="list-style-type: none"> i) Relationship between rurality, field exposure and attention deficit hyperactivity disorder (Donovan et al., 2019). ii) Correlation of certain diseases and the existence of private (supplementary) insurance (Warren & Zhang, 2018). iii) Financial burden of injuries for older workers (Davie & Lilley, 2018). <p>Examples: Inequalities in health future challenges for intersectoral cooperation (Gabrijeljic et al., 2021).</p> <p>Alcohol consumption and impact on health of alcohol use in Slovenia in the period 2013–2018, trends (Rados et al., 2022).</p>

COUNTRY	CASE STUDY	WHAT IS LINKED?	WHERE ARE THE DATA LOCATED?	WHO CARRIES OUT THE LINKAGE?	OPPORTUNITIES FOR RESEARCH	EXAMPLE OF PUBLISHED ANALYSES
Slovenia	HiM	Microdata describing socioeconomic and demographic data, health and some healthcare, e.g., on prevention measures, and some environmental data based on modelling. The data cover five broad categories: residents and community, health risk factors, prevention, health status, and mortality. This database presents information on population health at the level of Slovenia, statistical regions, administrative units and municipalities.	Data are stored at NJZ.	Initial data are linked by each participating institution as necessary	Open access.	Examples: Inequalities in health future challenges for intersectoral cooperation (Gabrijelcic et al., 2021). Alcohol consumption and impact on health of alcohol use in Slovenia in the period 2013–2018, trends (Rados et al., 2022).
Slovenia	Health Data Portal	—	NJZ	Open access.	Examples: Tobacco-attributable mortality in Slovenia 1997–2019 (Koprivnikar et al., 2021). Inequalities in health future challenges for intersectoral cooperation (Gabrijelcic et al., 2021). Suicide in Slovenia and around the world (Rosker et al., 2021).	Source: Authors' own.

4

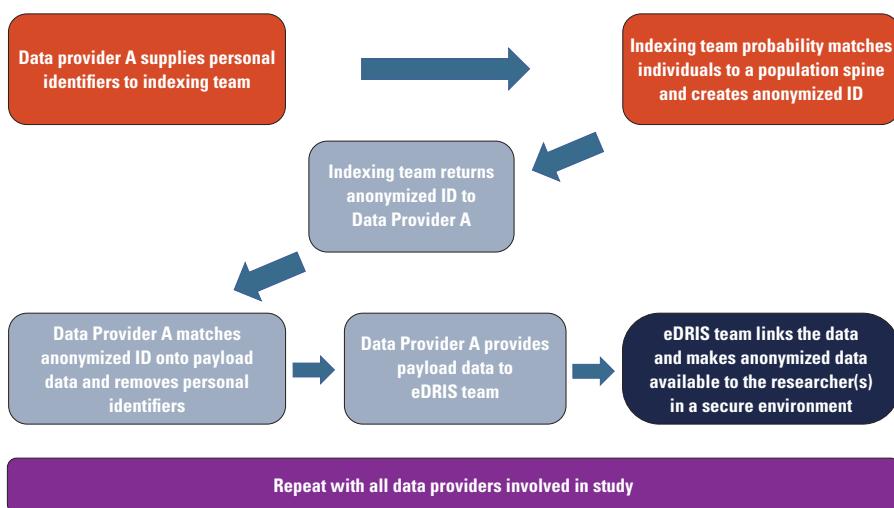
Differences in the linking process

Chapter 3 highlighted the potential of linking databases for the purposes of healthcare research. This can involve linking both existing datasets and new results from primary research with existing routine data. Successful linking presupposes the possibility of unambiguously assigning information from separate datasets to one individual. This requires suitable key variables that are present in all datasets to be linked. For reasons of data protection law (see Chapter 7), it is often necessary to use pseudonyms instead of unique personal identifying features from the original data, such as an insurance or ID number, for a personal link; the same pseudonymization procedure must be used for all data sources to be linked. Scheibner et al. (2020) noticed key differences concerning the definitions of anonymization, pseudonymization and de-identification in countries across the EU, Australia, Canada and the US. The responsibilities for the pseudonymization of the data as well as the process of de-identification can also contribute significantly to the preservation of anonymity and the protection of privacy. Anonymized data (see differences between types of de-identification of data in the List of Key Terms) are more difficult, if not impossible, to link.

A detailed technical description of possible de-identification and linking options goes beyond the scope of this review. Nevertheless, some interesting approaches can be highlighted. For example, in the sample of the case studies collected here, there is a cluster that uses a procedure based on an underlying “anchoring” list of persons to identify these persons in the datasets to be linked by means of identifying characteristics and thus to link data. In the New Zealand IDI this is the “IDI Spine”. This is a list of individuals that combines information from the tax authorities with birth and visa data. The aim of the spine is to identify all possible individuals who have ever lived in

New Zealand (target population) and to record them once (Black & Statistics New Zealand, 2016; Stats NZ, 2014). The microdata from the sectors mapped in the IDI are then linked to the spine and can thus be linked to each other, even without a unique identifying feature across the different sectors. The “Population Spine” in Scotland, which is used for linking via eDRIS, works similarly. It maps all persons in Scotland who have been in contact with NHS Scotland, including their identifiers (ISD Scotland, 2019). Figure 4.1 shows the process in Scotland that starts by matching only the identifiers in the records of included persons with the spine.

FIGURE 4.1 Linking approach in Scotland (eDRIS)

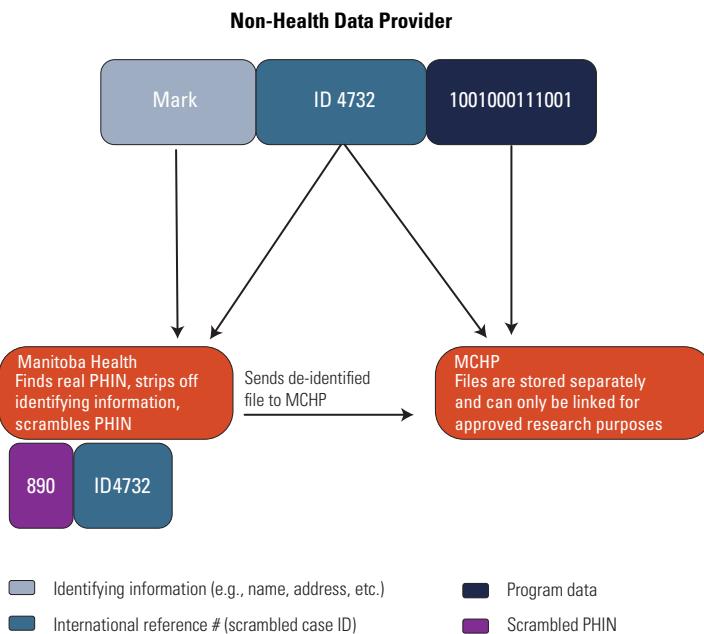


Source: Authors' own, adapted from NHS Scotland, 2020.

The concept of a list of target population with identifiers is also the basis of the Derived Record Depository (DRD) in Canada, which is used for linking via the SDLE. The DRD can be updated via its source data, as well as other datasets that update the personal information of individuals in the DRD. In addition to the DRD, the SDLE also has a “key registry”, where the linked IDs of the different datasets are stored separately so that the key registry can later be used to identify the corresponding cases in the original datasets (Trudeau, 2017). Records that contain the same unique person identifier can

be linked to each other, for example, healthcare records that contain the Health Insurance Number. Otherwise, deterministic and probabilistic approaches are used to match the datasets using non-unique identifiers. The principle of separation between the data containing the information for linking and the original data, as well as a “population repository”, are also applied in PopData BC. The result is a dataset with the linkage information, which can be used to identify cases from other datasets.

The approach used at the Manitoba Population Research Data Repository is highlighted here as an example of how distributing data components to different actors can aid with retaining de-identification during the linkage process. Datasets arrive without identifying characteristics such as names, addresses and telephone numbers. Data holders send this information separately to Manitoba Health (the Manitoba Department of Health), while they send only programme data and reference numbers to Manitoba Centre for Health Policy (MCHP) (the repository host). Thus, neither Manitoba Health nor MCHP have access to the complete dataset. Using deterministic and probabilistic methods, Manitoba Health links the identifying characteristics to personal health information numbers (PHIN), which are then sent to MCHP in encrypted form together with the reference numbers as crosswalk files. Every person who is insured in the statutory health system in Manitoba (see Chapter 2), as well as every person for whom a service has been billed in Manitoba, receives a PHIN. Once MCHP has this information, it can use the reference numbers to link the programme data to the encrypted PHINs. The underlying principle is that none of the actors involved has all the information: the data holders have no access to the encrypted PHINs, the repository has no access to the identifying characteristics, and Manitoba Health does not have the programme data (see Figure 4.2).

FIGURE 4.2 Linking approach in Manitoba

Source: Manitoba Centre for Health Policy, 2020.

Another interesting example is the linking of the CPRD data with other datasets (see Chapters 3 and 6). Before data from the practices flows into the CPRD, identifying characteristics (names, birth dates, postal codes and National Health Service Numbers) are removed via the practice software providers and replaced with pseudonymized patient and service provider identifiers. For practices that have agreed to the data linking, the software providers send the identifiers to NHS Digital, which checks the validity of the data and merges it with existing data from the CPRD to form a “cohort”. Data holders of the datasets to be linked (see Box 2.2) also send identifying characteristics to NHS Digital together with a pseudonymous patient identifier, the “Link ID”. NHS Digital matches the identifiers of the external data holders to the CPRD identifiers and generates a “linker” file with pairs of pseudonymized identifiers (practice and patient ID, Link ID) for each patient (Padmanabhan et al., 2019).

France's HDH has two other approaches to linking data from the SNDS (National Health Data System) with external databases. In one, the external database producer generates a temporary de-identified NIR (social security number, Numéro d'inscription au répertoire des personnes physiques), which is shared with SNDS along with common variables via the SAFE https procedure. SNDS pseudonymizes this NIR and prepares the dataset. The HDH links in direct deterministic matching the extracts from both the external and mainstay SNDS databases (Health Data Hub and L'Assurance Maladie Caisse Nationale, 2019). In the second, HDH uses statistical data linkage from a set of discrete variables and variables common to both the SNDS and external datasets (Health Data Hub and L'Assurance Maladie Caisse Nationale, 2019).

The case studies described above show that linking on the basis of other identifying characteristics is possible even without the presence of a unique identifier and has been successfully used for several years. Examples can be seen in Table 4.1. The distribution of responsibilities for data linkage among actors within the health system varies – nevertheless one can observe that the use of trusted third parties of different affiliations is common.

TABLE 4.1 Examples of linking methods used

COUNTRY	INITIATIVE	TYPE AND PROCESS OF THE LINK
Australia	PHRN	Probabilistic linking without a unique identifier based on other identifying characteristics (first name, last name, date of birth, address). PHRN supports the corresponding local (<i>state, territory</i>) institutions in data linkage.
	SDLE	StatCan links on a secure server; deterministically or probabilistically based on the presence of identifying information from the records to be linked.
Canada	Manitoba Population Research Data Repository	Personal Health Number; de-identification and linking of data are carried out by the same unit of the Ministry.
	HDH	HDH links in direct deterministic matching the extracts from the cohort/external databases and those of SNDS databases based on pseudonymized NIR and common variables or via statistical data linkage from a set of discrete variables and variables common to both the SNDS and external datasets.
New Zealand	IDI	Unique identifiers within the sectors if possible. Alternatively probabilistic linking, based on existing personal data in the spine. The linking is done by Statistics NZ.
Slovenia	zVEM Portal	Unique identifiers such as health insurance number and patient number. Linking is done by NIJZ and institutional healthcare providers.
United Kingdom	CPRD	NHS Number and other personal and practical data; deterministic approach.
United Kingdom (Scotland)	eDRIS	Community Health Index and other identifying characteristics, probabilistic matching via spine.

Source: Authors' own.

5

Access to health-related data for researchers

For the majority of case studies included in the review, an application for data use must be submitted to the responsible institution before data can be made available for research purposes. In many cases this is done digitally and is supported by a central office, even if data from several data holders have to be integrated (e.g., PHRN, eDRIS, HDH). For example, the application process at PopData BC is among the most accessible and transparent in Canada, following a three-step process: (1) the application must be submitted including ethics and review processes and consent documents; (2) the PopData BC staff in charge help applicants to obtain the approval of the necessary data holders; and (3) after successful approvals, contracts of use (including costs) and non-disclosure agreements can be signed. To access the data, successful applicants must then complete data protection training and set up their working environment within the PopData BC secure research environment.

For all case studies that provide access for researchers, the application rules were either explicitly described on the website or users were referred to the appropriate place. The timing for approval and processing varies by case study as well as by the complexity of the data linkage required. For instance, the IDI, NIVEL Primary Care Database and HIRA aim to make the data available for researchers in six weeks or less.

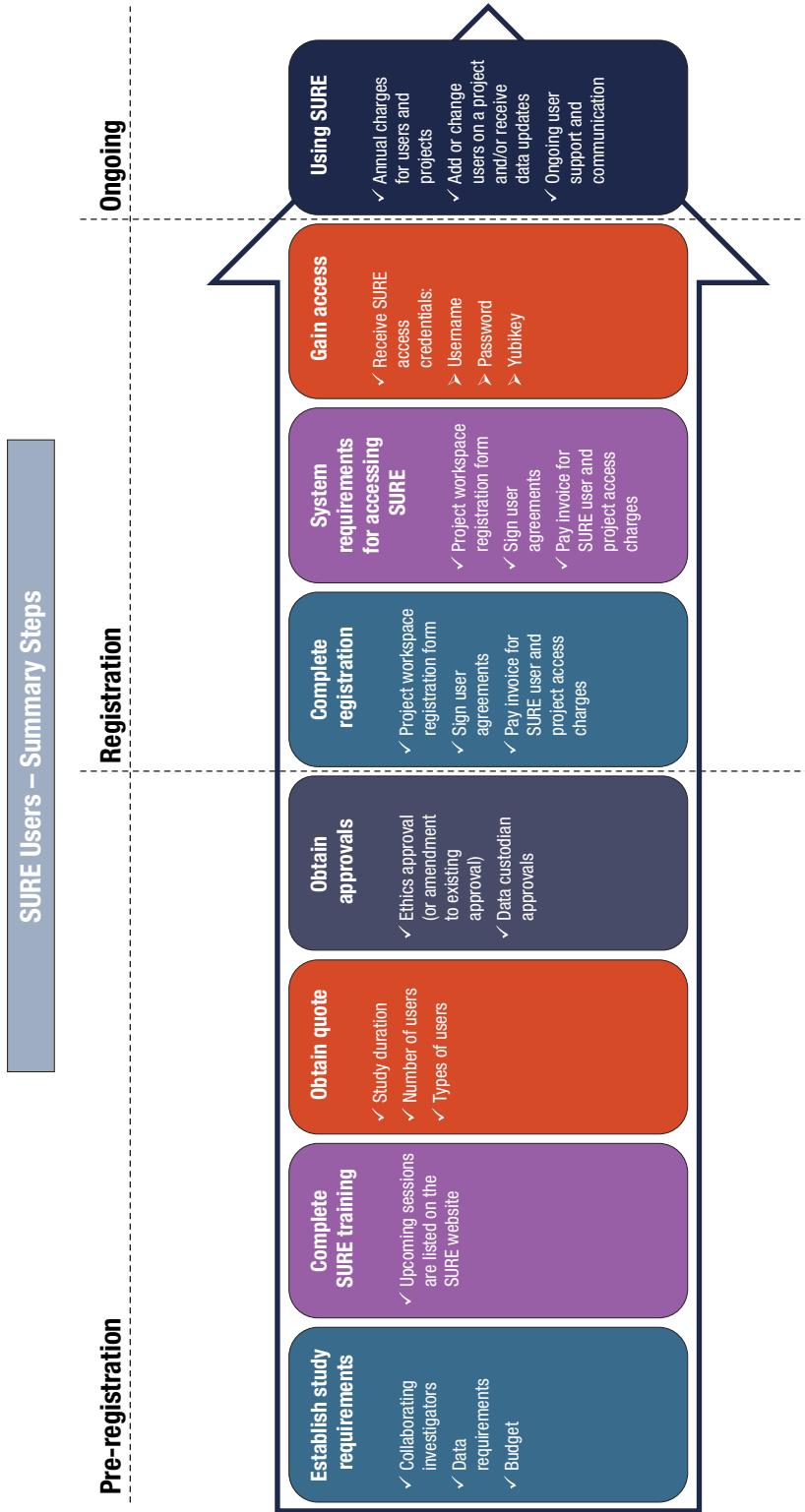
Since data in the healthcare system are fundamentally sensitive, three access paths are generally possible for researchers: direct data transfer, which requires an appropriate data protection and data security concept; in-house analysis, which is possible at suitable guest workstations of the institution holding the data; and remote inquiry, which requires separate servers with

“remote access” via VPN connection (see for example, All European Academies et al., 2021). As shown in Table 5.1, all three access methods are represented in the case studies included in this review. The use of guest workstations and remote access via VPN seem to be increasingly applied.

An interesting example to illustrate the application and usage processes is the Secure Unified Research Environment (SURE), which facilitates access to PHRN-linked data in Australia (Figure 5.1). It is managed by the Sax Institute in Sydney and represents a data processing environment. Here, data managers and researchers can upload data and researchers can analyse data. It is typically used to process de-identified microdata and linked data. During processing, the data remain on Sax servers in Sydney, but after successful completion of the analysis, aggregated data can be downloaded to local computers and belong to the researchers. Authorized users of SURE can work anywhere in the world and enter the environment via secure login.

Secure processing environments are used elsewhere as well; in France the HDH prepares a unique secure project environment on its platform for each project, to which it gives remote access to researchers. In the Republic of Korea access to the HIRA microdata are possible by appointment and on the basis of an application and review process via seven research centres or by remote access. In addition, HIRA also provides certain samples of hospital care (13% of the insured), the elderly population (20%) and the paediatric population (10%) that are accessible without a review process (Kim, Kim & Kim, 2014). Aggregated data are publicly available as summary statistics.

FIGURE 5.1 Requesting and using data via SURE



Source: Sax Institute, 2019.

TABLE 5.1 Data access processes for selected case studies

CASE STUDY	WHO CAN REQUEST THE DATA?	WHERE/BY WHOM MUST THE DATA BE REQUESTED?	WHO DECIDES ON APPLICATIONS?	HOW LONG DOES THE EVALUATION PROCESS TAKE?	HOW AND WHERE CAN RESEARCHERS WORK WITH THE DATA?	WHO HOLDS THE DATA DURING THE ANALYSIS?
PHRN (Australia)	Researchers whose projects meet content and ethical regulations	PHRN Online Application System for cross and multi-jurisdictional applications or data linkage unit for single jurisdiction application	Data linkage unit(s)	Varies according to the datasets to be linked	SURE, remote-access data research laboratory	Data remains on the servers in Sydney
PopData BC (Canada)	Canadian researchers working at a research institution	PopData's Data Access Unit (DAU)	Data holder	Varies according to the feedback from data holders	Secure Research Environment (SRE), can be used over a firewall using encrypted VPN	Data holder
HDH (France)	Anyone working in public interest that obtains authorization, e.g., researchers, policy-makers and individuals	(CESREES and the National Commission for Data Protection and Liberties (CNIL)	CESREES and the National Commission for Data Protection and Liberties (CNIL)	CESREES approval takes one month; CNIL review takes two months	Remote access to secure project environment given to process (not retrieve) data on the Health Data Hub	HDH
NIVEL Primary Care Database (Netherlands (Kingdom of the))	Researchers who comply with certain regulations (including the obligation to publish the results)	NIVEL via a form (available on website)	NIVEL on the basis of the rules of procedure (central associations of service providers and data protection committee)	On average three weeks, but depending on complexity and data protection specifics	Depending on the application, either remote access facility or dataset directly available	NIVEL
IDI (New Zealand)	Analysts and researchers in the public sector, universities and NGOs	Statistics NZ (online)	Statistics NZ, + other authorities if applicable	Review process: six weeks	Statistics NZ secure data lab environment (at Statistics NZ offices; institutions can apply to Stats NZ for their own data labs)	Statistics NZ

CASE STUDY	WHO CAN REQUEST THE DATA?	WHERE/BY WHOM MUST THE DATA BE REQUESTED?	WHO DECIDES ON APPLICATIONS?	HOW LONG DOES THE EVALUATION PROCESS TAKE?	HOW AND WHERE CAN RESEARCHERS WORK WITH THE DATA?	WHO HOLDS THE DATA DURING THE ANALYSIS?
HIRA (Republic of Korea)	Scientists, public sector, private sector (the amount of data varies)	HIRA	HIRA	Two to four weeks after application until data are available	Seven research centres or remote access (subject to conditions)	HIRA
zVEM Portal (Slovenia)	Patients and healthcare providers have access to original records. Others have access to specific data in accordance with their legal basis (e.g., MoH, public health authority, researchers)	NIJZ; access enabled by obtaining appropriate digital identifier through SI-PASS system	Ethics board or data commissioner provides access	Data are immediately available after being entered for patients/providers.	One of two national eHealth data centres (in Ljubljana and Maribor)	NIJZ
HiM (Slovenia)	Open data	Open data portal for anyone to access	N/A	The portal is refreshed annually	Located at NIJZ	N/A
Health Data Portal (Slovenia)	Open data	Open data portal for anyone to access	N/A	Portal is updated approx. six times/year	Located at NIJZ	N/A

Source: Authors' own.

6

Considerations for electronic patient records

Data from electronic patient or health records offer a further source of analysis for health services research. This review has so far adopted the terminology of each case study when describing digital records containing patient information; however, the fine differences between different terms merit consideration. The US Department of Health IT distinguishes between electronic medical records, electronic health records and personal health records (see Box 6.1 and, on the content of electronic health records, Häyrinen, Saranto & Nykänen, 2008).

Electronic health records, or applications that allow patients and providers to access all health-related information in a coherent way, are a key element in the implementation of eHealth in many countries. Denmark is the best

BOX 6.1 Definitions according to the US Department of Health IT

Electronic medical record: simplifies a digital version of the traditional paper-based medical record; primary purpose: diagnosis and treatment for clinicians in the same organization (practice, hospital). Not intended for exchange with other service providers.

Electronic health record: wider use, contains information from all providers involved in the care of patients and all these providers are allowed to access the information contained in it. Electronic health records “follow the patient”.

Electronic personal health record: they contain the same information as electronic health records, but are structured in such a way that patients can access and manage them. The aim is to enable patients to store and manage their information in a private, secure and confidential environment.

Source: Office of the National Coordinator for Health Information Technology, 2019.

example among the countries in this review. Through the underlying health data network, which contains health and healthcare information from all service providers (see Figure 2.1), patients can access their data on Sundhed. dk. Sweden has a system that is similar in principle, but technically different with a comparatively more limited implementation (Davoody et al., 2019). The primary purposes of such systems are to improve the coordination of care and to increase transparency for patients, not least to promote their empowerment. Regular access to the data is available to the patients themselves and to the service providers. Slovenia already has envisioned the use of data for research purposes in its zVEM Portal, which operates as an expanded electronic health record, providing secure access to all eHealth services. While patients and healthcare providers have access to original data, others, including health authorities and researchers, may access specific data in accordance with their legal basis and project approval.

In recent years some countries included in this review have introduced the implementation of electronic patient records, which could potentially create new datasets for research. In Australia and Austria patients have the option to opt out of using electronic records altogether, or to exclude certain information from being stored in the file or from being visible/accessible to certain providers (see Table 6.1). In Australia patients are additionally informed when a service provider accesses their file. My Health Record data can be used for research purposes, provided patients have not objected to the use of their data for research (using an opt-out system). Although the Austrian model initially foresaw a similar clause for research use, the political will to implement it is said to have been lacking. It is obvious that depending on how many and which patients decide for or against the use of such applications, and in accordance with the precise design options for access rights, the completeness and representativeness of the data could be significantly limited.

TABLE 6.1 Access options and patient rights for electronic patient records

COUNTRY	INITIATIVE	ACCESS	OPT-OUT/ DETERMINATION OF DATA VISIBILITY	ADDITIONAL OPPORTUNITIES FOR RESEARCH
Australia	My Health Record.	Patients and service providers have regular access to the data.	Patients: Opt-out of My Health Record, restrict selected service providers, hide certain documents; notification when service providers access My Health Record; opt-out from use for research. Service providers: Opt-out	Access to the data will be available in the future for a fee, excluding data from citizens who have opted out of My Health Record or have opted out of secondary use of their My Health Record data for research purposes.
Austria	ELGA (Electronic Health Record).	Patients and service providers (implementation phase), the latter with time restrictions.	Patients: Whole, partial (for individual applications, i.e., e-reports or e-medication) or situative (for specific treatments/healthcare service providers).	Regulated but not implemented.
Denmark	Sundhed.dk	Patients and service providers have regular access to the data.	No	No
Slovenia	zVEM Portal	Patients and service providers authenticated by a digital certificate have access to original records; others have access to specific data based on legal status and project approval (e.g., Public Health Authority, MoH, researchers).	Patients: Opt-out Service providers: Opt-out	Researchers can submit requests for data and then the appropriate procedures are carried out in accordance with the law. Results must be shared before dissemination.

Source: Authors' own.

7

Considerations for privacy and governance of digital health data

Historically, the relationship between patients and their health system has been anchored in expectations of privacy, which are underpinned by laws related to privacy, confidentiality, patient rights and data protection. The proliferation and availability of digital personally identifiable health- and healthcare-related data from diverse sources challenge traditional approaches to treating data. The established technical, ethical, legal and scientific procedures for thinking about and dealing with health-related data, including collection, storage, analysis and exchange, need to be updated in order to be better fit for purpose within digital infrastructures. For example, expectations of privacy and confidentiality (and corresponding regulatory regimes) must extend to the guardians, collectors, and processors of digital health-related data as well as to providers as the data are vulnerable at each phase. Especially when linked to other sources, personal health-related data can offer expansive opportunities for secondary research (OECD, 2015; Boyd et al., 2012; Scheibner et al., 2020) as described in previous chapters. However, while the potential value of secondary health data for research is great, leveraging and realizing this has implications for privacy and data security in a digital context.

This chapter first explores the overall rights to privacy and data protection and the associated obligations of nations. Second, it presents several principles and regulatory considerations for privacy and data security in the context of digital personal health-related data, with a view to research. How these are accounted for structures national governance approaches and determines the balance between data protection, availability and accessibility for secondary uses domestically. In addition, differences in national regulatory environments based on divergent handling of these considerations can help to either

facilitate or hamper how data are used and shared beyond/across jurisdictional boundaries for secondary research purposes. Third, it explores potential risks to privacy and data security when using linked health data. Fourth, as it is nearly impossible to gather – let alone disentangle – the many policies involved in privacy and data protection in one country, including as they pertain to health, this chapter will provide examples of existing data governance approaches from the review countries to better understand current regulatory practices and possible incongruencies. Finally, it will discuss the European Union's legally binding General Data Protection Regulation (GDPR), and its implications for cross-border sharing of data and barriers. Importantly, the chapter reviews the status of data privacy and protection before 2020 and the COVID-19 pandemic. Additional information on how COVID-19 has changed the usage of data and the consideration for health data protection is briefly covered in Chapter 9 on data linkage and the COVID-19 pandemic.

7.1 Rights to privacy and data protection regarding secondary use of digital health data

Privacy and data protection are two separate but related rights, both considered vital to a sustainable and well-fortified democracy. They are instrumental in preserving and promoting other fundamental values and rights, like freedom of speech and most countries have legally recognized the right to privacy, either constitutionally or through legislative action (European Data Protection Supervisor, 2020). Privacy is also a recognized human right, enshrined in the Universal Declaration of Human Rights (UDHR; Article 12), the European Convention of Human Rights (ECHR; Article 8) (European Court of Human Rights (2021), and the European Charter of Fundamental Rights (ECFR; Article 7) (Box 7.1). All countries in this review are legally bound to the UDHR, while those in the EU are also beholden to the ECHR and ECFR.

The concept of data protection derives from the right to privacy, operationalizing and implementing that right in a regulatory context. As such, data protection deals with safeguarding information relating to an identifiable person as well as its accessibility for said person; that is, information that could be leveraged to infringe on a person's right to privacy. Over time, it has also evolved into a right itself; for example, the ECFR includes an explicit right to the protection of personal data (Article 8). For the purposes of this

BOX 7.1 The rights to privacy in international consensus documents**Article 12 of the UDHR states:**

No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

Source: United Nations, 1948.

Article 8 of the ECFR states:

Everyone has the right to the protection of personal data concerning him or her. 2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified. 3. Compliance with these rules shall be subject to control by an independent authority.

Source: Official Journal of the European Communities, 2000.

Article 7 of the ECFR states:

Everyone has the right to respect for his or her private and family life, home and communications.

Source: Official Journal of the European Communities, 2000.

review, personally identifiable data falls under “individual level data” in the conceptual framework (Figure 1.1) and can be sociodemographic/economic data, health, or healthcare data.

7.2 Risks to privacy and data security when using linked health data

In terms of using linked health data for research purposes within modern, digitalized health information infrastructures, there are historically several ways in which the processing of data risks privacy and security. These include but are not limited to:

1. Possibility of lost or stolen data implicated by merely sharing data between people and places (Understanding Patient Data, 2018).

2. Obligation of recipients of shared data to ensure sufficient protection to keep data confidential once received (IT Governance, 2020).
3. Potential discovery of a person if all data points in a set of linked data are not thoroughly de-identified (Understanding Patient Data, 2018).
4. Possible financial and psychosocial harms incurred by the potential misuse of personal health information (Understanding Patient Data, 2018).
5. The economic and reputational consequences suffered by institutions experiencing data breaches (IT Governance, 2020).
6. On a societal level, the loss of public confidence in institutions responsible for privacy and data protection, such as the legislative and socio-health functions of a country, or even providers, should violations occur (Robbin, 2001).

In addition, cyberattacks, another challenge to data security and protection, have been increasing against health providers and other guardians of digital personal health data especially in the context of the COVID-19 pandemic (Open Access Government, 2019). This is occurring as increasingly more personal health data are being generated, used and shared in innovative ways, challenging – sometimes temporarily bypassing – the existing approaches to data security and privacy protection in the quest for expediency in the face of a rapidly spreading and evolving virus (see Chapter 9).

To pre-empt and mitigate risks, different organizations have set out principles for data protection, security, and privacy in the use of digital health data.

7.3 Principles and regulatory considerations for digital health data protection and security

The importance of both scientific inquiry using digital (personal) health data and the dissemination of related research has been highlighted in international data protection recommendations and consensus-based guidelines

(Scheibner et al., 2020; UNESCO, 1997; UNESCO, 2005; WMA, 2002; OECD, 2013; OECD, 2015; OECD, 2017; Oderkirk, 2021; CIOMS & WHO, 2017). International research guidelines generally recognize four common and important principles for data protection governance, which reflect the balancing act policy-makers must carry out when facilitating secondary use of individual level health data while fulfilling the responsibility of governments to obtain individual permission/consent and patient involvement, preserve privacy and promote security to protect against possible violations. These principles are (1) free and informed consent; (2) accountability for confidentiality and risk mitigation; (3) patient/individual compensation for research involvement; and (4) efforts to ensure accessibility and maximum reuse of data collected for research and provide an overarching framework required of data governance regimes, according to international consensus.

For the full realization of the benefits of health data for research, as described previously, health-related datasets should be digitally linked together. This can take different forms: data can expansive and centralized or multisite (linked across jurisdictions, i.e., institutions, countries, regions), multilevel (taken at individual, subnational and national level) and multisectoral (representing the care continuum). To achieve this, new technical solutions are needed, while also complying with privacy and data security objectives. Technical solutions may include innovative and reliable de-identification procedures (to remove identifying information from data), trustworthy linkage approaches, and secure sharing and data storage. However, with the advent of new technological solutions also arise legal, ethical and digital skills challenges that national data governance regimes need to take into account (Blasimme & Vayena, 2016).

Data protection considerations across the legal, ethical, and digital dimensions can be separated broadly into two categories: data content and data processing. How governments define specific details of these categories determines in part to what extent data governance is privacy-protective and how much limit there is to access for research purposes. The OECD, which contributes substantially to thinking about health data governance, has put forth eight data protection principles to be met to safeguard privacy, while allowing for meaningful research. These also help to elucidate the two data protection categories (Figure 7.1):

(1) collection limitation principle (“There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.”);

(2) data quality principle (“Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up to date.”);

(3) purpose specification principle (“The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose”);

(4) use limitation principle (“Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9 except a) with the consent of the data subject; or b) by the authority of law.”);

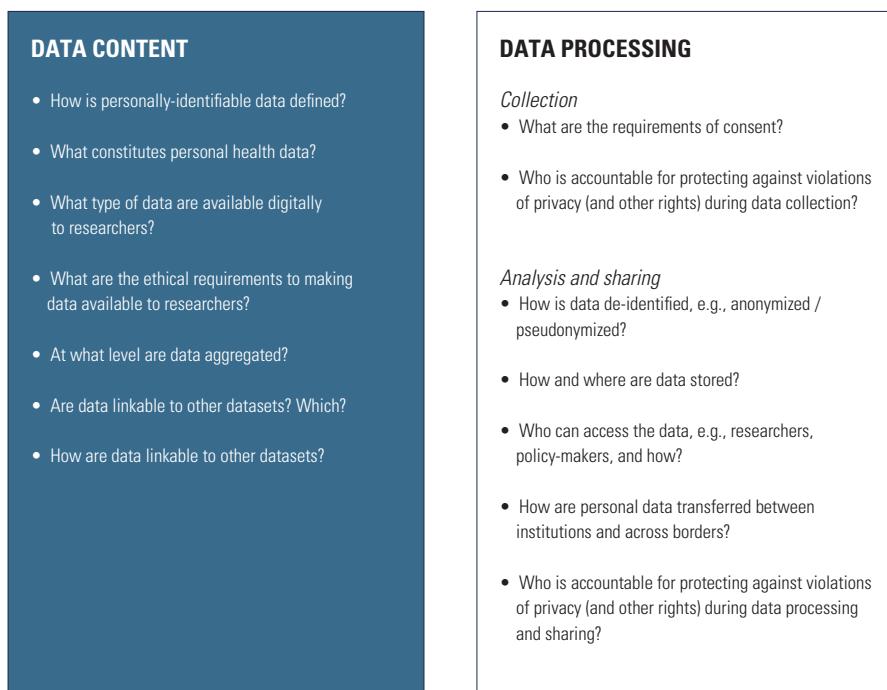
(5) security safeguards principle (“Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification or disclosure of data.”);

(6) openness principle (“There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller.”);

(7) individual participation principle (“Individuals should have the right: a) to obtain from a data controller, or otherwise, confirmation of whether the data controller has data relating to them; b) to have communicated to them, data relating to them i. within a reasonable time; ii. at a charge, if any, that is not excessive; iii. in a reasonable manner; and iv. in a form that is readily intelligible to them; c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and d) to challenge data relating to them and, if the challenge is successful, to have the data erased, rectified, completed or amended.”); and the

(8) accountability principle (“Data controller should be accountable for complying with measures which give effect to the principles stated above.”) (OECD, 1980; amended 2013).

Figure 7.1 provides examples of questions defining these categories and whose answers inform digital health data governance in a jurisdiction.

FIGURE 7.1 Data protection considerations for data content and processing

Source: Authors' own, derived from Blasimme & Vayena, 2016; Scheibner et al., 2020; OECD, 1980, amended 2013.

Data content refers to the nature and type of information being protected, including its completeness, accuracy and currency (up-to-dateness), as well as what it is used for. Not included in the OECD guidelines are specific technical attributes of the data, including at what level they are aggregated. Data content also pertains to technical aspects of data related to data linkage, e.g., do the data include personal national ID number or a healthcare card number, and whether linkage to other protected datasets is legal and possible (Scheibner et al., 2020). Answers to these questions help determine the level of protection needed and at which points and moments in the digital data infrastructure and process.

Once the content to be regulated is specified and definitions are established, procedures for processing the data must be set. Data processing can be separated into three different procedures: (i) collection, (ii) analysis and (iii) sharing. The goal of data regulation is to protect against potential violations

of privacy (see above) during collection, analysis and sharing of data, while also ensuring access and appropriate use of the data for secondary purposes. As such, questions around sufficient patient consent and outlining when and how data can be obtained for research purposes (and clarifying explicit exemptions) shape policies. Also important is to clarify who is accountable for data processing (and liable for infringements on data protection) and provide guidance for the processing and sharing of digital health data for secondary use. Absence of such guidance may undermine data accessibility in practice. The above list of considerations and corresponding questions is not exhaustive; however, it does describe how, depending on answers, differences in health data protection frameworks or approaches can arise. Incongruencies in how data content and data processing are defined and regulated in one location may create challenges for the processing and transfer of data across locations with different regulatory regimes, e.g., countries or regions, as standards of data protection may not be fully aligned, or, in some cases, reconcilable.

Turning to health-related personal data, specifically, building on its 2013 guidance the OECD has made several recommendations for the design of national health data governance frameworks, which operationalize aspects of data content and processing and the considerations of data privacy and protection mentioned above (OECD, 2015, 2017; Oderkirk, 2021). In 2015 the OECD recommended the establishment of (1) a health information system, (2) a legal framework, (3) public communication, (4) a system of certification/accreditation of processors, (5) a project approval process, including ethics review, (6) data de-identification, (7) data security and management, and (8) a data governance review cycle (OECD, 2015). In 2017 and 2021 it re-emphasized transparency in data access; monitoring and evaluation; coordination among public and private actors processing health data; data linkage; informed consent (or appropriate alternatives); maximizing the potential of technology; population representation; ensuring proper training and skills for those processing data; and cross-border sharing. A national health data governance framework aligned with these principles encourages the availability and use of personal health data for secondary purposes to benefit health-related public welfare while also promoting privacy, personal health data and data security (WHO, 2017; Oderkirk, 2021). Moreover, between countries or jurisdictions with difference governance approaches, taking these principles into account entirely or even partly can help to reduce the differences in jurisdictional health data governance frameworks and facilitate cross-jurisdiction data sharing.

7.4 Overview of data governance approaches

Among the countries in this review there is no single common policy approach for regulating personal digital health data. While one country might have designated comprehensive legislation on health data privacy, another might have a national data governance framework, where health is incorporated – or both. In some countries additional legislative initiatives may have clauses that influence the practice of data protection, especially as it pertains to vulnerable populations or specific health areas. Federalized countries often manage a complex dynamic of state/regional and national laws when it comes to privacy and data protection as well as (health) data use and transfer (Oderkirk, Ronchi & Klazinga, 2013). There can also be fragmentation within public healthcare as countries struggle to update the regulatory dimension at the speed that new digital solutions are being introduced, which influences data protection and access. A study by the European Commission in 2019 found that one third of study countries have integrated neither eHealth policy into general healthcare policy nor aspects of patient safety and quality into eHealth policy, including France and the UK. Meanwhile, Austria, Denmark, Finland, the Kingdom of the Netherlands, Sweden and Slovenia have integrated their policies (European Commission, 2019).

Despite these possible differences, there are similar data principles underpinning most regulatory approaches, informed by international guidance and regulations like the OECD and EU's GDPR. Such principles, which align with those outlined above, include accessibility, accountability, accuracy, confidentiality, data minimization, integrity, purpose limitation, storage limitation and transparency (Blasimme & Vayena, 2016; European Commission, 2018; Caseguard, 2021). Variation on which among these is prioritized or supersedes others is partly due to the administrative structure of a country, to cultures of risk perception management and to public attitudes on privacy and data protection (Oderkirk, Ronchi & Klazinga, 2013; Rumbold & Pierscionek, 2017a).

In addition to the above shared principles, the 14 countries surveyed in this review all have some formal health data governance framework that is either established or being established (see Appendix II; Oderkirk, 2021; Australian Government, 2021). Five of these – Austria, Denmark, Finland, France and New Zealand – are codified in law. Canada was just beginning in 2021 to establish a pan-Canadian health data governance framework and Slovenia started to develop a national health data governance framework in 2019.

Though it is nearly impossible to provide a detailed overview here of the complex health data governance regulatory framework in one country, let alone a set of 13 countries, Appendix II outlines some of the seminal policies and elements that shape health data governance and regulation in each jurisdiction. When reviewing this, several patterns do emerge.

Depending on the administrative structure, data governance occurs at the national, regional or local levels of government or a combination

Data protection and privacy laws related to health information can be applied at the national (federal) level, as in Austria, Finland, New Zealand, the Republic of Korea, Slovenia, or the UK. Other federalized countries such as Australia, Canada and the US have a complicated ecosystem of national and regional/state laws. How these interact (and supplement or complement each other) varies, depending on the nature of the separation of responsibilities across these administrative levels. In Australia there is a range of entities responsible for datasets at each level of government, with no overarching health data governance framework. However, the 2020–2025 National Health Reform Agreement includes a vision to introduce a national approach to data governance, structures, and processes.

National laws or regulations for different sectors and obligations to special interests may generate conflicting or overlapping regulations

Sector-specific regulations, e.g., public versus private, and those related to industry, e.g., healthcare, entertainment or technology, in one country may overlap with or challenge each other and an overall framework. All study countries have separate regulations regarding obligations to privacy of public versus private sector actors. Countries may have explicit legislation pertaining to their legal, ethical and social obligations towards specific groups or populations, generally and specifically related to health, which shapes how health data processing and sharing are interpreted and implemented. New

Zealand's Treaty of Waitangi (New Zealand Constitutional Law Resources, 2017), for example, which outlines the special protections for and rights of Māori, including data sovereignty and non-discrimination, runs up against the mission of the health sector to generate data-driven policies and insights for system and population health improvement (see Chapter 2). In fact, countries may have overlapping pieces of legislation that confront each other (Durie, 1989; Waitangi Tribunal, 2019).

Entities have been or will be established as central authorities for the approval of requests to process personal health data

Many countries have newly established separate agencies or entities to take on key roles in data governance. Almost all countries in this review have already established or will establish central authorities for approvals of requests to use personal health data for secondary purposes. Some of these also play a greater stewardship role in health data governance overall. For example, along with introducing health data governance in 2018, the Republic of Korea established a "Healthcare Big Data Policy Deliberation Committee" that is responsible for data development, use and dataset linkages. In France the HDH established in 2019 defines the elements of shared data governance with stakeholders. Meanwhile, within the information governance framework for personal data of the Scottish government, there is a patient panel called the Public Benefit and Privacy Panel for health and social care data that evaluates applications for access to NHS Scotland health data for secondary purposes regarding the public benefit and privacy implications of proposed projects. In Slovenia and Sweden the Information Commissioner and Swedish Ethical Review Authority, respectively, approve requests for data processing for research projects, though in the latter data request approval has other requirements depending on the dataset and size of project as well. In the US healthcare providers follow the HIPAA Privacy Rule (Centers for Disease Control and Prevention). In Denmark subnational authorities approve requests, while in Canada provinces and territories have separate processes for personal health data request review and approval (European Commission, DG Health & Food Safety, 2021).

Health strategies to operationalize and implement data governance frameworks

A number of countries, including Australia, Austria, Finland and Denmark, have an explicit, dedicated health data strategy road-mapping a vision of future data processing and sharing. Slovenia is in the process of preparing to develop a new eHealth strategy. Many others have issued guidelines on operationalizing the principles, standards and criteria of privacy legislation relating to health information.

7.5 The General Data Protection Regulation of the European Union

Entered into force in 2016 and replacing the 1995 Data Protection Directive, the EU's GDPR is the core of Europe's digital privacy legislation and sets guidelines for the protection of personal information from residents of the EU. It is an enormous political and policy effort to streamline the fragmented regulatory reality of EU countries' digital data cultures and infrastructures to promote a standard level of data security and privacy protection, while allowing for mutual use of data and sharing across national boundaries (Hansen et al., 2021; Yuan & Li, 2019).

As background, GDPR is concerned with "the protection of natural persons with regard to the processing of personal data and on the free movement of such data" (European Commission, 2016). Understanding the value of making data accessible for secondary purposes, while ensuring privacy provisions, it "[aims to make] Europe fit for the digital age". The GDPR's new framework for the protection of personal data ensures the same data protection rights across the EU and obliges organizations (public and private) that process data to respect these rights, wherever these entities are located. Under the regulation, reuse and transfer of data across borders are limited to countries assessed as providing adequate protection.

The general principles of data protection introduced by the GDPR for Europe in the digital environment (for example, privacy by design or the prohibition of discriminatory profiling) are both relevant and applicable to personal health data. Moreover, further safeguards for personal health data have also been introduced; the regulation adopts a new definition of personal data to be protected under its scope; in it, it lists "genetic data", "biometric

data”, and “data concerning health” as special categories. The GDPR is binding for EU Member States and the (health) data governance requirements of the GDPR are also translated into national and state laws (when applicable). At least eight countries in this review (Austria, Denmark, Finland, France, the Kingdom of the Netherlands, Slovenia, Sweden, and the UK) have adopted specific rules for the processing of biological and health-related data according to the GDPR. The UK implemented the GDPR into national law via the Data Protection Act 2018 before the finalization of Brexit.

The GDPR, Recital 35, explicitly defines “health-related data” as

“Personal data concerning health should include all data pertaining to the health status of a data subject which reveal information relating to the past, current or future physical or mental health status of the data subject. This includes information about the natural person collected in the course of the registration for, or the provision of, healthcare services as referred to in Directive 2011/24/EU of the European Parliament and of the Council to that natural person; a number, symbol or particular assigned to a natural person to uniquely identify the natural person for health purposes; information derived from the testing or examination of a body part or bodily substance, including from genetic data and biological samples; and any information on, for example, a disease, disability, disease risk, medical history, clinical treatment or the physiological or biomedical state of the data subject independent of its source, for example from a physician or other health professional, a hospital, a medical device or an in vitro diagnostic test.”

Under the GDPR data protection based on this definition does not apply to data or research involving anonymized data (information not relating to an identified or identifiable natural person) but rather to pseudonymized data (data that can be restored to its original state with the addition of information) and non-de-identified data. If health and health system researchers strive to work with sensitive non-pseudonymized personal data, the GDPR also requires explicit consent from individuals to process sensitive personal data (Article 9), with some exceptions, including for certain individual and public health reasons like preventative medicine, medical diagnosis or providing healthcare services; this is subject to professional standards of secrecy or for reasons of public interest in the area of public health, such as protecting against public health threats or to ensure medical device quality. In terms of accountability for GDPR compliance, the regulation holds all processors of personal data – recipients and providers alike – responsible, regardless of public or private sector affiliation.

Despite the aspirations of GDPR, there remain different rules and regulations determining access to health data both within and between EU Member States, which impact researchers in the context of both in-country and cross-border research. This also makes it hard for researchers themselves to understand and navigate the rules governing their intended research with health data. Generally, the GDPR errs on placing significant limitations on the use of health-related data for secondary purposes; however, EU Member States required to implement GDPR into national law are permitted to introduce further conditions, including limitations or allowances on processing biometric, genetic or health-related data (European Medicines Agency, 2020). Paradoxically, whether these modifications always comply with GDPR's requirements is unclear, which has implications for sharing data across countries (European Commission, DG Health & Food Safety, 2021). Another barrier to cross-border data processing is that different countries have different expectations of consent, which particularly affects obtaining patient data for clinical research and for health systems research (European Commission, DG Health & Food Safety, 2021). Beyond consent, other differences in Member State ethical requirements for health-related research could hamper EU collaboration. The research ethics entities of countries have different mandates and scopes, as do the project approval processes, with some requiring multiple approvals or vetting by a national data protection authority (see Chapter 5). Further, countries define and interpret terms related to data content and processing differently, with consequences for secondary research (see Chapter 4). Within the EU, for example, one study found that the distinction made between public and private researcher varies across countries, as do the specified regulatory approaches to data that has been de-identified in a hybrid manner (anonymized and pseudonymized) (European Commission, DG Health & Food Safety, 2021).

All these differences impact on the accessibility of health data and implicate additional factors that affect the availability and accessibility of health data such as respect for data subjects' rights and available procedures in case a violation has occurred, among other things. Within the EU more common approaches to structure and allow for the reuse of data for research are necessary (European Commission, DG Health & Food Safety, 2021; Rumbold & Pierscionek, 2017b).

International data sharing and the EU health data space

In addition to the intranational data linkage initiatives described in this review, several international organizations hold non-individual level data that enable researchers to compare health, healthcare, sociodemographic or economic and environmental data across countries. These include but are not limited to the World Health Organization (WHO), OECD Health Statistics, Eurostat, the World Bank and the International Monetary Fund (IMF). Further, databases for specific initiatives, such as the Global Burden of Disease study, collect information at a global level as well (Vos et al., 2020). While these databases are valuable in cross-country comparisons at a national level and are often used for health systems research, they less frequently contain more granular data at a subnational or individual level. Within the EU some initiatives have already taken place for cross-country data sharing of individual level data. This chapter introduces several key initiatives that have implications for health services research, including the cross-border healthcare directive, European Reference Networks (ERNs) to the European Health Data Space (EHDS).

Directive 2011/24/EU on patients' rights in cross-border healthcare enables EU citizens to receive medical care and reimbursement for healthcare costs in another EU country (European Parliament, 2011). The Directive represented the first legislative initiative of the European Commission specifically related to health services, and this caused some controversy during the time of its implementation (Azzopardi-Muscat et al., 2018). The uptake of the number of patients seeking care abroad has increased over time, but only represents a small percentage of all patients (Peralta-Santos & Perelman, 2018). The European Commission indicates that most of these are driven by geographic or cultural proximity, which is supported by evidence from

Germany (Panteli et al., 2015) and the Kingdom of the Netherlands (Verra, Kroeze & Ruggeri, 2016). The ability to exchange data across borders for both patient care and research requires some level of standardization across countries (see Chapter 7), and the eHealth Digital Services Infrastructure promotes common standards for e-Prescriptions and patient summaries (European Commission, 2022a), described further below.

The 2011 Directive on cross-border care included a regulation on mutual recognition of prescriptions. The cross-border prescriptions Directive of 2012 further clarified this, yet initially saw relatively low uptake, with only six countries (Croatia, Czechia, Estonia, Finland, Luxembourg, and Portugal) able to share prescriptions outside of their borders as of 2020 (European Commission, 2021). This may be related to differences in legal requirements for prescribing medicines, international non-proprietary name (INN) policies, and which healthcare professionals are able to prescribe medicines.

The cross-border exchange of patient summaries, or health information important for patient treatment and care, has also begun in some countries. The eHealth Network, a platform of all EU Member States that organizes cooperation on digital health at the EU level, released its first guidelines on patient summaries for non-scheduled care in 2013 (eHealth Network, 2013), aiming at a standardized data collection process to enable the goals set out in cross-border Directive 2011/24/EU. In 2019 the European Parliament published Recommendation 2019/243/EU on a European Electronic Health Record exchange format, referencing the eHealth Network guidelines as the basis for developing standardized patient summaries and e-Prescriptions (European Parliament, 2019). The patient summary guidelines have been updated twice more, with the third released in June 2021 (eHealth Network, 2021). A report from the same year showed that 68% of countries have incorporated references to eHealth Network guidelines into national policy documents and 54% have structured the patient summary according to the 2016 guideline (European Commission, 2019). A list of the two highest levels of variables that the eHealth Network recommends including in the patient summary is presented in Table 8.1. An expanded example of the full guidance, also including another level of variables, definition and comments, and preferred code system, is provided in Table 8.2 for “Vaccination/prophylaxis information”. The work of the eHealth Network and the need for cross-border exchange of patient data to enable cross-border care paved the way for the adoption of the European Health Data Space, described further below.

TABLE 8.1 Patient summary administrative and clinical data

NESTING LEVEL 1 VARIABLE	NESTING LEVEL 2 VARIABLE
Administrative Data	
Identification	National healthcare patient ID
Personal information	Full name Date of birth Gender
Contact information	Address Telephone number Email Preferred health provider to contact Contact person/legal guardian
Insurance information	Insurance number
Clinical Data	
Alerts	Allergy Medical alert information (other alerts not included in allergies)
Medical history	Vaccination/prophylaxis information Resolved, closed or inactive problems Medical history
Medical problems	Current problems Medical devices and implants Procedures Functional status
Medication summary	Current and relevant past medicines
Social history	Social history observations
Pregnancy history	Current pregnancy status History of previous pregnancies
Patient provided data	Travel history Advance directive
Results	Result observations
Plan of care	Vital signs observations
Diagnostic tests	Therapeutic recommendations that do not include pharmacologic treatments, such as diet, physical exercise, planned surgeries

Source: Authors' compilation based on European Commission, 2021.

TABLE 8.2 Expanded example of patient summary data

NESTING LEVEL 2 VARIABLE	NESTING LEVEL 3 VARIABLE	DEFINITION AND COMMENTS	PREFERRED CODE SYSTEM
Vaccination/prophylaxis information	Disease or agent targeted	Disease or agent that the vaccination provides protection against	ICD-10* SNOMED CT GPS
	Vaccine/prophylaxis	Generic description of the vaccine/prophylaxis or its component(s)	SNOMED CT GPS ATC* (IDMP, when available)
	Vaccine medicinal product	Medicinal product name	For the time being, this should be the name of the medicinal product as registered in the country. In the future the information on the medicinal product can incorporate the identifiers from the implementation of the ISO IDMP Standards and the medicinal package's unique identifier
	Marketing Authorization Holder	Marketing Authorization Holder	EMA's Organizations System data (SPOR)
	Number in a series of vaccinations/doses	Order in the vaccination course	–
	Batch/lot number	A distinctive combination of numbers and/or letters which specifically identifies a batch	–
	Date of vaccination	The date when the vaccination was administered	ISO 8601
	Administering centre	Name/code of administering – centre or health authority responsible for the vaccination event	–
	Health professional identification	Name or health professional – code responsible for administering the vaccine or prophylaxis	–
	Country of vaccination	The country in which the individual has been vaccinated	ISO 3166
	Next vaccination date	The date when the vaccination is planned to be given/repeated (e.g., next dose)	ISO 8601

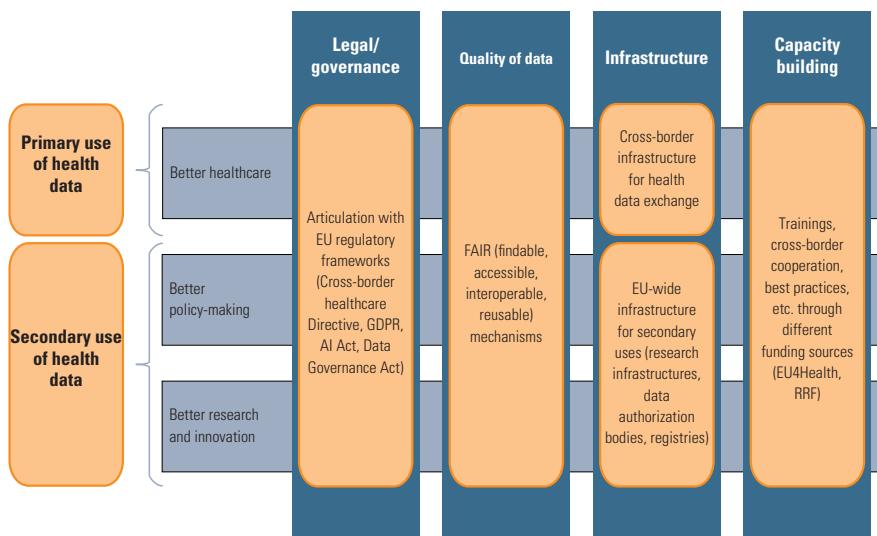
Note: * In a foreseeable future, the suggested preferred vocabularies might be superseded or complemented, as mentioned in Guidelines Article 11(2).

Source: Authors' compilation based on European Commission, 2021.

The COVID-19 pandemic also drew more attention to cross-border healthcare and the exchange of health data, particularly as it relates to the interoperability of contact tracing apps and digital vaccination certificates. Millions of EU residents have downloaded contact tracing apps, but they do not necessarily talk to each other, limiting the sharing of health data across borders. The EU Digital COVID certificate includes information about vaccination status, test results and recovery status that is standardized according to common specifications. Both efforts are described further in Chapter 9.

Sharing individual-level data internationally also supports research and knowledge sharing about rare diseases. The European Reference Networks (ERNs), which were launched in 2017, are a primary example of cross-border data collaboration in this area, and the largest platform for clinical research and translation worldwide (Tumiene et al., 2021). Over 1200 ERN centres exist with 26 countries participating in 24 virtual networks on complex and rare diseases (European Commission, 2017). Although one function of ERNs is patient care, research is another crucial function. Yet, decisions taken by national and European authorities have a high influence on the organization and funding of research in ERNs, and each research project requires its own funding application, representing a lack of funding sustainability (European Court of Auditors, 2019).

Health data exchange is increasingly becoming a larger policy objective at the EU level. In February 2020 the Commission published ‘A European strategy for data’, which includes ‘A Common European health data space’. In November 2020 the Commission and its German presidency announced a European Health Data Space (EHDS), and at the end of 2021 the European Commission announced the related legislative proposal, with concurrent goals of better healthcare, better policy-making, and better research and innovations. The full proposal on the EHDS was released on 3 May 2022 and distinguishes between activities supporting better healthcare delivery (primary use) and better research and policy-making (secondary use) (European Commission, 2022b) (Figure 8.1).

FIGURE 8.1 European Health Data Space

Source: Authors' compilation based on European Commission, 2019.

In order to enable secure exchange of health data, the proposal highlights the need for a European electronic health record exchange format. The proposal identifies 15 categories of electronic data that data holders should make available for secondary use, including the data contained in EHR systems; data impacting on health, including social, environmental and behavioural determinants of health; health-related administrative data, including claims and reimbursement data; electronic health data from clinical trials; and human genetic, genomic and proteomic data. While GDPR will prove a challenging prerequisite (see Chapter 7), other health system obstacles, including governance, financing, infrastructure and interoperability, should not be underestimated. For example, the proposal states that “electronic health data entailing protected intellectual property and trade secrets from private enterprises shall be made available for secondary use” (European Commission, 2022b). The proposal also outlines accepted and prohibited purposes for which data can be processed for secondary use, and identifies the need for a data permit, the possibility for data holders to levy data access fees, and the need for a secure processing environment. The Joint Action Towards the European Health Data Space (TEHDAS) aims to develop criteria and guidelines for the secondary use of health data (including their use for research purposes).

Data linkage in the COVID-19 pandemic

The global COVID-19 pandemic has challenged health systems worldwide to respond quickly to control, track and treat the acute virus and its long-term consequences, while also maintaining essential care functions and respecting equity and privacy considerations. In this way it has been a test of global health information systems. COVID-19 has laid bare weaknesses in national health information systems, including out-of-date information technologies, inadequate interoperability, lack of legal teams with competencies in topics such as health data privacy, and more (Negro-Calduch et al., 2021; Pooransingh et al., 2021). Countries with stronger health information systems, with elements such as centralized data collection mechanisms/points and strong coordination across sectors via unique patient identifiers, for example, were better positioned for a timely crisis response, as they have the tools to gather and link data from different sectors to capture insights into the personal and system level impact of the pandemic (Schmidt, Abboud & Bogaert, 2021).

At the same time the COVID-19 pandemic has also spurred considerable innovation in health and healthcare data generation, collection, sharing and linkage (de Bienassis et al., 2022; Negro-Calduch et al., 2021). Because of the novel, evolving, nature of the virus as well as its rapid spread, health systems and researchers have required unprecedented data integration and access to accurate, real-time data from myriad sources and disciplines. Data linkage within and between systems is essential for generating the health information needed to address the pandemic since data inputs for managing – tracking and treating – COVID-19 relate to different settings and levels of care. Data access and sharing of previous health status or healthcare usage of COVID-positive patients, for example, contribute to better understanding of the effect of comorbidities and the optimization of treatment – or design of prevention

measures (Salerno et al., 2021). Linking this information to sociodemographic data enables a more granular understanding of the spread and outcomes of COVID-19 among different socioeconomic groups (de Bienassis et al., 2022). Meanwhile, COVID-19-related data linked with mobility data supports more accurate forecasting of the virus's spread (Canino et al., 2022). Even in countries with little previous data linkage activity and decentralized health and healthcare data, the pandemic has pushed the health system to engage in data sharing and research partnerships (Lee and Haupt, 2021; Negro-Calduch et al., 2021; de Bienassis et al., 2022).

This chapter focuses on dimensions of data linkage during the COVID-19 pandemic around the themes described in previous chapters mainly in the study countries. However, examples from outside the study countries are given when particularly relevant. The following section elucidates some of the innovations around data collection and data sources used in the study countries during the COVID-19 pandemic, picking up on topics described in Chapter 3. This is followed by discussion about linkage between COVID-19-related data and other databases (linking to content in Chapter 4), enhanced engagement with researchers (linking to content in Chapter 5), integration with electronic health records (linking to content in Chapter 6) and international initiatives (linking to content in Chapter 8). In addition, the novelty of new sources and the speed of development of new solutions, catalysed by COVID-19, have implications for data protection (linking to content in Chapter 7) and may require extra consideration for the future.

9.1 New sources of data developed for the pandemic response

The emergence of COVID-19 presented a first challenge to health systems: to capture the considerable new COVID-related data, including diagnosis, associated care decisions and health outcomes. As such, new mechanisms for reporting and analysing this novel, real-time individual level health data were introduced. Some countries expanded the scope of existing data collection procedures to include COVID-19-related data. For example, established disease surveillance systems and other health registries provided a foundation for efforts in some countries but, they needed to be adjusted, expanded and

reorganized to keep up with the dynamics of the pandemic. For example, hospital discharge registries were modified to include COVID-19 variables as several countries updated the national version of ICD-10/11 according to COVID-19 coding advice and WHO/ECDC case definitions and recommendations (Negro-Calduch et al., 2021). Other adaptations included moving to electronic and more timely data collection systems such as death registries.

In other countries new information systems and novel datasets supplemented established health information systems, which can be plagued with irregular updating and lack of remote access for researchers, to fulfill the data needs during the pandemic. For example, specific registries for COVID-19 hospitalizations (e.g., the Kingdom of the Netherlands, France and Manitoba and Ontario, Canada), COVID-19 cohorts (e.g., Denmark and Sweden) and COVID-19 specific vaccination databases (e.g., France and the Kingdom of the Netherlands) were all introduced. Additionally, ICU information systems (e.g., in Australia) captured further information about the spread of the virus and health system capacity (Negro-Calduch et al., 2021; Pilcher et al., 2021). All these were linked to additional datasets for research purposes (see next section).

Voluntary digital contact tracing systems have been introduced by all study countries, primarily as part of a strategy to contain the virus spread (Milioris & Papageorgiou, 2021; Blassime, Ferretti & Vayena, 2021). But they may also be linkable to other datasets, depending on the maturity of a country's data linkage infrastructure. These applications are used primarily through proximity tracing (bluetooth) and collect aggregated statistics on the usage and functionality of the application and metrics data for public health surveillance (e.g., day, time and duration (usually 10–15 minutes) of a contact); whether the infected user is asymptomatic; first day of symptoms; and the date of testing. By country, an application may also include the capability to track symptoms (e.g., France, England, the Republic of Korea), acquire local epidemiological information (e.g., France, England), order COVID-19 tests and access results (e.g., England) and store a range of personal information (Blassime, Ferretti & Vayena, 2021). Countries may also retain anonymous data for epidemiological surveillance and research purposes, as with the ZOE COVID app in Scotland (Box 9.1); however, retention periods vary across countries. In England, for example, data are kept for 20 years; in Scotland it is held indefinitely (Blasimme, Ferretti & Vayena, 2021).

BOX 9.1 The ZOE COVID app in Scotland

ZOE COVID is a mobile application that allows users to voluntarily track their own symptoms and collect data on location, demographic information and any COVID-19 test results. The collected data can be shared with health researchers through SAIL (Secure Anonymised Information Linkage consortium) in Wales and the Health Data Research UK (HDR UK) gateway.

Participants of the ZOE COVID Study give their explicit consent to this upon setting up the app. As of 2 September, 2021 the data of new users can be used to study a wider range of health conditions, including other infections, dementia, cancers and more.

Sources: Menni et al., 2022; Zoe Health Study.

BOX 9.2 The use of geographic data during COVID-19

The conceptual framework used in this review (Figure 1.1) identifies sociodemographic/economic data, health data, healthcare data and environmental data as areas of interest for health researchers. However, during the COVID-19 pandemic, the use of Geographic Information Science and Technology (GIS&T) data in public health has gained particular attention (Smith & Mennis, 2020).

Each country presented in this study has developed GIS&T datasets and projects, many of which were initiated and supported by governments, such as COVID-19 dashboards using maps. GIS&T data were also used to model the effectiveness of non-pharmaceutical interventions and support surveillance, containment and contact tracing efforts (Smith & Mennis, 2020). In the US GIS&T data were used to investigate geographic health disparities including the disproportionate impact of COVID-19 on predominantly Black communities (Siegal et al., 2021). In the Republic of Korea mobile phone data were used to verify patient claims during contact tracing activities (Park, Choi & Ko, 2020; COVID-19 National Emergency Response Centre, Epidemiology & Case Management Team, Korea Centres for Disease Control & Prevention, 2020).

Connected to these contact tracing applications is the enhanced availability and use of Geographic Information Science and Technology (GIS&T). Combined with COVID-19-related data, these systems can provide real-time information essential for health policy decision-making and epidemic monitoring (Ahsan et al., 2022; Reeves et al., 2021) (Box 9.2).

However, using mobile phone data requires attention to equity and protection considerations related to populations surveyed, as well as the design, collection, analysis and interpretation of data (Anom, 2022). Further, while cell phone data can provide highly detailed information at numerous levels of aggregation, they can be biased (Szocska et al., 2021). Yet the opportunities for linking GIS&T data with other collected data abound. They have already been employed in the UK where self-reported data from the COVID Symptom Study app were linked with the Secure Anonymised Information Linkage (SAIL) databank, which includes multiple health and social care datasets, to enable near real-time monitoring of COVID-19 in small geographic areas (Fry et al., 2021).

9.2 National examples of COVID-related data linkage activities and innovations

Beyond new mobile technologies and collection mechanisms for capturing pandemic-related data, countries worked quickly to facilitate linking of these new data with other sources to gain insights into the virus's nature, behaviour and health impact on individuals and populations. In some cases, countries linked the newly created databases to each other. In the Kingdom of the Netherlands and France all hospitalized persons with a positive COVID-19 test are registered in the nationwide registries of COVID-19 hospitalizations; these data are subsequently linked to vaccination status and stored in national COVID-19 vaccination databases. Such pairing permits analyses related to one or several of these themes together (preprint Gier et al., 2021; Republique Francaise, 2022; de Bienassis et al., 2022).

Several countries leveraged the considerable amounts of already routinely collected data by their health systems, including health and administrative data. Australia, Austria, Canada, Denmark, France, the Kingdom of the Netherlands, Slovenia, Sweden and the UK, for example, all linked hospital and mortality data to COVID-19-related data to understand the impact of the pandemic on patient and population groups (Milioris & Papageorgiou, 2021; Carr et al., 2021; MHRA, 2022a; Gutpa and Aitken, 2022). Additionally, Australia linked infection, hospitalization, vaccination, medical care, medication and mortality data (Australian Institute of Health and Welfare,

2020). In the US health data records were linked to COVID-19-related data, while in Scotland, test results, health records, hospital admissions and deaths were linked (US, CDC, 2021; de Bienassis et al., 2022). Similarly, data from the established health information systems of Public Health England, the NHS (England) and the recent National Immunization Management System, which all include COVID-19-related data, have provided insights into determinants, infections, emergence of new variants and the efficacy of different interventions, as well as vaccine effectiveness (Armstrong et al. 2020; Bhattacharya et al., 2021; Majeed et al., 2022). Moreover, the data authority in Victoria, Australia, the Centre for Victorian Data Linkage (CVDL), intends to develop linked datasets to look at the impact of COVID-19 on service usage in mental health, cancer and chronic health conditions (Safer Care Victoria, 2021).

Despite the unprecedented speed of new health data generation and research in the time of pandemic, scientists are calling for other, more efficient, solutions, for example, linking COVID-19 trials to the medical databases and routinely collected data, which is not yet a widespread practice (Paprica et al., 2020; Pottegård et al., 2020). Prospective linkage to the databases could reduce the costs of trials (by shortening a follow-up time in experimental conditions), increase speed (especially in the case of real-time data) and possible monitoring of multiple outcomes, and reduce problems with loss to follow-up for long-term outcomes in clinical trials. Some institutions, like Health Data Research (HDR) UK and Health Data Research Network Canada, have already begun to support this activity (Paprica et al., 2020). For example, HDR UK and NHS Digital provided services on feasibility assessment of the international clinical trial *RECOVERY* (Randomized Evaluation of COVID-19 Therapy). *RECOVERY* aims to identify treatments for hospitalized COVID-19 patients based on real-time hospital data (ISRCTN registry, 2020).

Several countries specifically linked data from other health, care, and social sectors to COVID-19-related data. In Denmark COVID-19 cohort data on (positive and negative) PCR tests are retrieved every hour from the Danish Microbiology Database (MiBa) – part of the Danish Health Data Network – to identify new cases of COVID-19, and additional information is acquired by data linkage from national registries by unique identifier. Registries include the Civil Registration System, the National Patient Registry, the National Immunization Registry, the Register of Causes of Deaths, a

registry of long-term care facilities, the National Database of Work and Productivity, and the Danish Covid-19 Genome Consortium. De-identified and pseudonymized data on all tested individuals are also transferred daily to research databases, available for researchers after application (Pottegård et al., 2020; Schønning et al., 2021). In Sweden all residents with a positive SARS-COV-2 (COVID-19) PCR test are similarly included in the Swedish Covid-19 Investigation for Future Insights – a Population Epidemiology Approach using Register Linkage (SCIFI-PEARL) cohort (Nyberg et al., 2021), for timely response to scientific questions regarding COVID-19. The database encompasses a regularly updated observational study and a representative comparison cohort linked to national and regional healthcare data for timely insights for research on COVID-19. These databases are linked using the national personal ID so that information from the cancer and prescribed medicines registry can be incorporated in the analysis along with the COVID-19-related data, national patient register and two regional health databases. Slovenia sometimes linked other national data such as health insurance data for occupation and employment activity to its health and COVID-19-related data, including tests, cause of death and inpatient data (de Bienassis et al., 2022). In accordance with the concern that COVID-19 has had a disproportionately negative impact on lower-income people and other marginalized populations, Austria, Canada and the Republic of Korea linked COVID-19-related data to individual level sociodemographic data, allowing deeper analyses. In Austria a research project called the Austrian Corona Panel conducted a survey and analysed socioeconomic effects on COVID-19 outcomes (Kittel et al., 2020). In the UK certain established health data infrastructures already link census, health, education, workforce and social care data, but the scope was expanded to include COVID-19 tests, sequence, tracking and vaccination status. One study looked at COVID-19 outcomes among the homeless (Song et al., 2021). As part of the recovery plan in the UK (England), the government asked systems to improve the collection and recording of ethnicity data across care sectors, e.g., primary care, outpatient, mental health, community services and emergencies. In other countries socioeconomic data were collected separately as they could not be linked to newly collected COVID-19 data (de Bienassis et al., 2022).

9.3 Building on existing databanks and data linkage platforms

Some countries sought to build on existing platforms that support data linkage to enhance COVID-19-related insights. These include EHR (see below) and regional and national health data linkage projects. Denmark's COVID-19 cohort, for example, is only possible due to the country's robust data linking system. Meanwhile, the Welsh Multimorbidity e-Cohort (WMC), originally designed to support research on prevalence, determinants and healthcare needs of multimorbidity, was repurposed to meet the region's data needs for its COVID-19 response (de Bienassis et al., 2022). WMC uses SAIL (Box 9.1), which provides secure storage of de-identified individual level data for research to improve health, well-being and services and has integrated 28 new anonymized datasets into the platform. These span census, school and workforce data, and several COVID-19 related data sources, including symptom tracking, test results, shielded people list, and sequencing and vaccination data.

Several countries modified their usual regulation of data integration to facilitate timelier access to (personal) information. The UK invoked an exemption built into its National Data-Opt Out, applied when there is an overriding public interest in the use of data, to ensure timely, relevant, individual level data collection in the pandemic. In Australia states/territories (e.g., Victoria) expanded the frequency of reporting and data scope of existing notifiable disease surveillance systems and made improvements to enable linkage to these systems with inpatient hospital admissions, immunization, emergency department attendances and deaths (Safer Care Victoria, 2021; de Bienassis et al., 2022). CVDL also piloted a new population linkage spine methodology, where the datasets are linked to a “spine” of high-quality de-identified datasets with high population coverage, such as electoral enrolments, driver’s licence and births datasets, to increase the speed and accuracy of the linkage for monitoring short-term COVID-19 outcomes. This dataset could have applications for understanding the long-term health impacts of COVID-19 as well. At the Commonwealth level efforts are in place to link notifiable diseases surveillance systems with the national death register and with medical and pharmacy data. The list may be extended to admitted patient data, non-admitted emergency department data, aged care data and vaccination.

9.4 Innovations in data access and sharing

Countries have also innovated how data related to COVID-19 are shared both with the public and with researchers. New mechanisms such as applications to allow for quick analysis and reporting of data in a more digestible way have been established. Dashboards communicate COVID-19-related data to the public visually, including daily reported tests, cases and deaths, though sometimes additional indicators are also reported. Canada developed dashboards and interactive tools on excess mortality (McGrail, 2022), and an international interactive data map of COVID-19 cases by country for international benchmarking (Interactive data visualizations of COVID-19, 2022 ; Sadeghieh et al., 2022). Meanwhile, a Health Inequalities Improvement Dashboard in England (UK) contains expanded datasets on currently underrepresented aspects, such as people experiencing long COVID (de Bienassis et al., 2022).

For further research purposes other countries have developed additional platforms. For example, the Republic of Korea introduced a COVID-19 Epidemiological Investigation Support System (also called the “COVID-19 Smart Management System”) in 2020 (Park et al., 2020). This platform became a centre for data collection and multi-agency coordination and was developed through the application of the Smart City Data Hub Platform, which allows collection, process and analysis of massive volumes of urban data. The transfer of the system to the Korea Centre for Disease Control and Prevention (KCDC) permitted an automated epidemiological investigation process. First, by bringing such stakeholders as KCDC, the National Police Agency, the Credit Financial Association, and telecommunication and credit card companies onboard, the acquisition of the personal information needed for contact tracing became simple and rapid. Second, the automatic analysis of movement routes and tracing of the investigated case enables detection of infection sources and the transmission network for timely government interventions.

Countries have also adapted their policies for data sharing for research purposes, including on health systems. Austria, Canada, the Kingdom of the Netherlands and Sweden, for example, have introduced open data policies. Several countries have also created one-stop data repositories to make it easier for researchers and the public to find and use some data, including Australia, Austria, Canada, the Kingdom of the Netherlands, the Republic of Korea and the US (de Bienassis et al., 2022).

Additionally, some countries have adapted their data request processes for individual-level data for research projects, including Sweden, Australia and the Republic of Korea. Sweden fast-tracked processing requests for data and statistics for COVID-19-related research, including the required ethical review. Australia's one-stop-shop makes it easier to obtain data, including linked datasets. Similarly, the Republic of Korea and Scotland (UK) have streamlined and centralized some or all COVID-19-related data to improve efficiency of resource use and data access. Some are also progressing on allowing researchers to access data remotely, such as the Republic of Korea and Canada (de Bienassis et al., 2022).

9.5 Integration of COVID-19 data with electronic health records

Electronic health records for pandemic research can be a promising and efficient “real-time” way to generate new knowledge in the time of dynamic crisis. In Finland data from EHR, stored and managed by Kanta Services, have been used to follow the longitudinal progress of COVID-19 patients. The electronic health records also serve as frontline defence and triage against COVID-19 as it provides the necessary tools for users to report their symptoms and, based on that, allocate people to appropriate healthcare services (Milioris & Papageorgiou, 2021). In Alberta, Canada, results of COVID-19 testing are linked with the Alberta Electronic Health Record Information System (Connect Care) using the Alberta Healthcare Number (Baumgart, 2020). In England (UK) a novel nation-wide electronic health record resource enabled whole population research on COVID-19 and cardiovascular disease (Wood et al., 2021).

9.6 Examples of international data linkage and sharing

Innovation in data linkage and sharing occurred not only within national borders but across them. The nature of COVID-19 and the rapid development of, or acceleration of, existing data sharing activities led exceptionally quickly

to expedient data linkage and pandemic research across borders. The strong political, economic and academic ties in Europe, for example, facilitated the establishment of a unique EU-wide gateway, to enable interoperability between national digital contact tracing apps. By these means, travellers to and from Belgium, Croatia, Denmark, Estonia, Finland, Germany, Ireland, Italy, Latvia, Lithuania, Malta, the Kingdom of the Netherlands, Norway, Poland, Slovenia and Spain received contact tracing alerts when outside of their home country (and tracking information was shared) as of May 2022 (European Commission, 2022c). The international Consortium for Clinical Characterization of COVID-19 by EHR (4CE) gathered data of patients diagnosed with COVID-19 between January–April 2020 and laboratory tests from 96 hospitals across five countries within three weeks, providing early data on the disease progression. Associated descriptive statistics utilizing large-volume EHRs were also instrumental to identify at-risk populations (Dron et al., 2022). Box 9.3 offers two other examples of international cooperation in COVID-19-related data linkage.

BOX 9.3 Two examples of international health data research cooperation during COVID-19

1. Observational Health Data Studies and Informatics (OHDSI) Community initiated **Characterizing Health Associated Risks, and Your Baseline Disease In SARS-COV-2 (CHARYBDIS)** framework, which mobilizes real-world data (RWD) from descriptive epidemiological studies for tracking the evolution and management of the COVID-19 crisis (Prieto-Alhambra et al., 2021).

2. ELIXIR is another noticeable example in the field of multi- and international COVID-19 data collection and sharing research. It is the European research infrastructure for life science data, which strongly relies on the cooperation from 23 country nodes spanning over 200 institutes. ELIXIR's country nodes could either establish access to the national cloud and facilities for the COVID-19 research projects, or provide instruments for data search, access and share in the national settings. The aim of the ELIXIR Hub is an alignment of national infrastructures, European research institutes and the European Open Science Cloud (EOSC). As an European intergovernmental organization, ELIXIR aligns its actions with the development of European Health Data Space for COVID-19 (see Chapter 8) (Blomberg & Lauer, 2020).

COVID-19 has highlighted the critical need for accurate, interdisciplinary, real-time data access and sharing in the short-term to respond to and address the impact of the COVID-19 pandemic. It also presents as a unique opportunity in which health systems quickly developed and adopted data-based decision-making approaches and new forms of data integrations, where large volumes of real-time data could interoperate and allow for complex analyses. However, it is unclear if the innovations that have occurred are temporary efforts introduced in the acute phase of the pandemic or if they signal new ways of utilizing and leveraging data.

9.7 Data protection and privacy considerations

The advancement and innovation in data generation, usage, sharing and linkage during the COVID-19 pandemic have challenged the established national and international protocols and processes for data protection and privacy (see Chapter 7). Countries have been under pressure to provide data, especially at the individual level, as quickly as possible to researchers, practitioners and policy-makers. This challenges the established procedures for requesting and accessing data, which, in balancing privacy and data security with data usage, may slow data access or constrain data sharing. Such barriers have been demonstrated to hinder an efficient public health crisis response to COVID-19 (Basit, Lehmann & Medford, 2021; Kucharski et al., 2021; Schmidt, Abboud & Bogaert, 2021). In anticipation of such challenges, some national and international regulations on privacy and data protection, e.g., the UK, the Republic of Korea, Australia and the EU's GDPR, include a clause that allows for the superseding of public health and disease prevention laws during times of public health crisis. These allow for the processing of personal information without consent if necessary to protect against serious cross-border threats to health and threats to public health. As mentioned above, several countries have adapted, fast-tracked and streamlined usual data access procedures to protected sensitive data in the quest to get the right information to the right people at the right time. This applies as well in some cases to data protection regulation. The UK, for example, has extended its "opt-out" procedure several times (NHS Digital, 2022). For more information see the above section on Innovations in data access and sharing.

The new data sources and mechanisms for data sharing and linkage that have emerged during the COVID-19 pandemic challenge current regulations

in other ways. Innovative datasets are often not collected or processed by authorized bodies as per existing policy. For example, mobile phone companies and online reporting applications (e.g., digital contact tracing), as well as employers, have all processed personal data for a variety of specific purposes (including managing and protecting their workforce, customers and the public) during the pandemic. Because these do not fall under the data protection regulations in place and because they likely do not technologically adhere to criteria required of authorized bodies, these new entities and practices present a risk to data protection and privacy. They have also required the existing regulation and ethical practices to adapt to benefit from them. Box 9.4 presents

Box 9.4 Three country examples of new health data protection guidance adopted during the COVID-19 pandemic

Canada: Office of the Information and Privacy Commissioner of Alberta, Canada, issued a statement which requires developers of the Alberta Contact Tracing app to clearly outline the types of personal data being processed, and the purposes and conditions on which the data will be disclosed and retained (DataGuidance, 2020).

Australia: Australia passed a separate law regulating its COVIDSafe app privacy protection. The Privacy Amendment (Public Health Contact Information) Act 2020 specifies the circumstances for data collection, emphasizes its voluntary use and provides requirements for data storage and disclosure. It also identifies legal penalties which can apply for a misuse (OAIC, 2022).

Republic of Korea: The Personal Information Protection Act (PIPA) of 2011 in principle prohibits the collection, use and disclosure of personal data without prior informed consent of the individual whose data are involved. However, after the 2015 MERS outbreak, the changes in the Contagious Disease Prevention and Control Act (CDPCA) provisions and of PIPA and other privacy laws were superseded. The current version of CDPCA allows public agencies to collect, manage and share individual data that associate with confirmed or suspected disease cases. Location data, closed circuit television (CCTV) footage, mobile phone data, personal identification information and medical records are circulated among agencies and need to be linked with national health insurance information systems. The collection of personal and location information by the COVID-19 Smart Management System is compliant with Article 76-2 of the Infectious Disease Control and Prevention Act. Access to the personal information is strictly managed and granted to the authorized KCDC and local government health officials responsible for contact tracing (Park, Choi & Ko, 2020; Lee & Lee, 2020; Asia Pacific Observatory on Health Systems and Policies, 2020).

three country examples of health data protection guidance adopted during the crisis to accommodate new actors in the health data landscape.

Concurrent with the surge in health data production and sharing during the pandemic has been increased attention on the value of the data being generated. This has increased the vulnerability of personal health data, especially collected by new, often digital applications, entities for which both security regulations and technical practices are still in development. This has been exacerbated by the dispersion of resources (financial and human) and attention (organizational and political) throughout the pandemic, as well as an increased sophistication of ransomware. Overall, the security of health data applications and institutional information systems – and thus personal health data and patients' exposure to possible violations of privacy – has been threatened potentially more than ever before and healthcare institutions and other actors have been infiltrated by those seeking to leverage these vulnerabilities and capitalize on the opportunities represented by the newly generated data (Muthuppalaniappan & Stevenson, 2021). In Finland hackers attempted to leverage accessed medical records of mental health patients in therapy for their own gain by asking for payment in bitcoin to prevent the release of this personal information (Heikkilä & Cerulus, 2020). Ultimately, incidents like this have, however, led to more investment in digital infrastructure and data protection in some places. For example, after cyberattacks effectively shut down two hospitals in France, the government announced a large investment in cybersecurity in response (archyde, 2021).

Though the COVID-19 pandemic is an extreme situation, it has highlighted the need to maintain strong data and privacy protections in the face of constrained resources while not limiting legitimate access to health data for secondary research and the benefit of public health. In the future established procedures and standards, including ethics, will have to be reconciled with the precedents of data generation and processing practices that have arisen during COVID-19. This is especially important as the acute crisis ends and a more endemic phase begins. Doing so will ensure that the innovations introduced during the pandemic can be sustained, which will enhance health systems and support health system preparedness in future crises (Dron, 2022; De', Pandey & Pal, 2020). It will also enable continued generation of COVID-19-related data for long-term needs.

Other challenges also need to be addressed to leverage the advances in data generation, sharing and linkage made during COVID-19 in the future. For example, many of the innovations have been temporary or stand-alone. With some noticeable exceptions, many have not (yet) been integrated into national health information systems. Additionally, barriers covered in previous chapters such as diverging access protocols and differences in terminology and in linkage approaches still exist that undermine international data exchange and linkage (Chapter 7), though the future EHDS may catalyse the resolution of some of these issues. If the linking of data can be organized in a standardized and sustainable way nationally, regionally and internationally, it will strengthen the world's ability to respond quickly and appropriately to other viral threats and health crises. A last concern is the increasing burden that this additional data (and new) collection and processing place on already limited resources, both financial and human. Without the resources, including sufficient levels of trained and dedicated workforces at providers and in data agencies, to formalize and utilize new health data practices, the benefits of these will never be realized, even if the regulatory and technical underpinning are in place.

Conclusions

This review aimed to highlight the possibilities of linking data from different sources to support research aiming to inform policy decisions at different levels. It also showcased the differences between countries' health information systems and the way data are used, hoping to provide a basis for discussion and ideas generation to help policy-makers consider the evolution of their own data strategies. The COVID-19 pandemic catalysed several breakthroughs in the way data are collected, processed and used, and supranational initiatives, such as the EHDS, are capitalizing on the increased interest in the area to provide innovative solutions for the future. The case studies presented in this review can serve as examples to inform further specifications regarding which data are to be made available to whom and under what conditions, as well as which technical alternatives can be considered. However, it is necessary to underline that a much more specialized and comprehensive examination would be required to address the feasibility of the solutions highlighted here in different settings.

The cross-country variability of data infrastructures, and consequently of the possibility for making representative data of good quality available for research, is well established (see, for example, Oderkirk, 2021). For countries with fragmented data landscapes, this review clearly demonstrates the added value of comparatively easily accessible, comprehensive routine healthcare data that are available promptly, allow cross-sectoral analyses and can be linked to health and regional data on an individual level (e.g., MBS/PBS in Australia, NHIS/HIRA in the Republic of Korea and CMS Data in the US).

A number of the case studies presented are designed to enable access to linked data for research. These include both integrated models (such as the IDI in New Zealand) and central access points for requesting and linking distributed data from different data holders (e.g., PHRN in Australia, SDLE in Canada, HDH in France, eDRIS in Scotland; in all countries regional

initiatives are also active in parallel). Service platforms such as the ResDAC and the CMS Data Navigator in the US can also facilitate the secondary use of data for research.

This review also highlighted the potential of databases that are based on the routine documentation (usually electronic health records) of healthcare providers (e.g., CPRD, NIVEL Primary Care Database). A systematic review of the factors influencing the development of databases based on electronic practice management systems in primary care in English-speaking countries identified further initiatives in England, Canada, and the US, and found that the most important prerequisites for successful implementation are a robust technical infrastructure and strong scientific and political support (Gentil et al., 2017). An investment in this direction, such as that made in the UK with CPRD, can not only provide insights for one's own healthcare system, but also promote international cooperation and contribute to the international visibility of scientific excellence (see Vezyridis & Timmons, 2016); participating practices can benefit from periodic reports as well as an opportunity for benchmarking. However, it is important to note that a number of sources of bias can hamper research based on data from electronic health records. This concerns, among other things, the documentation and coding in the electronic record, the extraction and merging of the data, and the preparation for scientific practice (Agniel, Kohane & Weber, 2018; Verheij et al., 2018). The representativeness of the sample in networks of participating practices (such as the CPRD and the NIVEL database) must be taken into account, as well as the completeness of the data; both are influenced by the options available to patients to exclude their data from research access (see below).

Long processing times for usage requests and procedural hurdles in carrying out analyses often hamper the use of data for healthcare research. Different approaches have been highlighted in this review to overcome these challenges. Canada and the US shorten processing times by requiring actors to send a defined extract of data to the data processing centre. The creation of guest researcher workplaces (as in the “safe havens” of the British NHS, the secure Datalabs of New Zealand’s IDI or also in the Republic of Korea) as well as direct data access by means of a secure connection (e.g., via a terminal server, such as the SURE system in Australia or VPN access to the PopData BC can be considered to improve access. At the interface between data collection and research, the concept of a “one-stop-shop” for data, i.e., a body

that coordinates or carries out all bureaucratic steps and thus simplifies the processes for applicants, serves not only to relieve the burden on researchers or policy analysts, but also to improve and make more efficient use of all data and thus to formulate and answer far-reaching questions. It is clear from the strategic efforts in countries included in this review (see Profiles in Appendix I) that strong political commitment is crucial for the successful implementation of new approaches that support the use of data for health services research (see also Tew et al., 2017).

The following sections aim to distill and summarize the insights gained from the case studies described in this review and highlight important considerations for future action towards making quality data available for research aiming to support better policies. These considerations can be relevant at the national and the European level; the regulatory proposal for the EHDS released in May 2022 provides first directions but foresees a number of delegated decisions that will determine the content and shape of data to be shared in the future. Relevant initiatives specifically looking at this question, such as TEHDAS, are ongoing (see Chapter 8). The following thoughts are presented against the backdrop of a fundamental requirement for the highest data protection standards and the preservation of the anonymity of persons (see also Chapter 7).

10.1 What criteria should be met to enable meaningful healthcare research based on linked data?

Based on the case studies identified in this study, several preconditions must be met for meaningful linked datasets to be created and made available for research. These are presented here as additional considerations to the FAIR Guiding Principles for scientific data management and stewardship, which stipulate that data should be Findable, Accessible, Interoperable and Reusable (see Wilkinson et al., 2016).

Ideally, healthcare research should be able to access data that:

- cover the entire population (this might require special considerations for particular population groups, see the example of New Zealand in this report);

- are available at individual level (“microdata”), i.e., not only at aggregated summary level, in a structured and standardized format;
- cover, to the extent possible, all relevant sectors of care, and within sectors all services and providers;
- are available over longer periods of time and can be linked over the years – longitudinal studies are indispensable for the appropriate evaluation of innovations in healthcare and require that the data can also be analysed as a panel;
- also contain information on non-medical, i.e., sociodemographic/ socioeconomic parameters as well as regional information (e.g., zip code or place of residence); and
- can also be linked at an aggregated (e.g., regional) level (e.g., with information on the localization or density of service providers, regional/PC data on average utilization or service provision, but also with other databases, e.g., with regard to regional environmental factors).

Furthermore, the regulatory framework around making data available for research should observe a minimum set of parameters. In particular, there must be:

- transparent regulations with regard to usage rights;
- equal conditions for all scientific institutions to access data;
- as little time difference as possible between data collection and possible data access;
- prompt access to requested data (meaning sufficient capacities and investment in data structures); and
- possibilities for direct access to individual level data (microdata) – under the relevant data protection regulations.

10.2 How can coordinating bodies facilitating data linkage support policy-informing research?

For researchers, answering research questions that require data from multiple sources poses two major challenges: i) different access regulations to the various data are complicated to navigate, and ii) data are often available in pseudonymized format and without a unique linking key. Access to data and

carrying out data linkage are often accompanied by extensive applications for data use, data protection concepts and, if necessary, ethics board approval applications – and thus a particularly high workload. This may discourage or significantly delay important research initiatives.

A coordination and data linking hub could alleviate these problems and reduce the effort for access and linkage considerably. Internationally, there are many positive examples of the use of coordination and data linking points (including the HDH in France, eDRIS in Scotland, the SDLE and PopDataBC in Canada, and the PHRN in Australia). While the scope of the available data and the processes for both data linkage and data access differ, the guiding principles of these various initiatives are aligned:

- objective: to improve data availability for the scientific community in order to promote evidence-based health policy and improve public health and enable more efficient use of resources;
- use of already existing datasets;
- a single point of access for scientists, providing information on the available data and analysis options, as well as performing the linking of different datasets; and
- the coordination and data linking institution ensures compliance with data protection regulations

To maximize the possible research applications, it would be necessary for such a coordinating body to provide the following data in a linkable way or carry out the linking process and coordinate access:

- *linkable on individual level.*
 - data of all health insurers/payers, statutory and private;
 - data of other social security institutions, such as pension insurance, accident insurance, etc.;
 - health data from electronic health or patient records (see below);
 - survey data (e.g., regarding subjective health, experiences with the healthcare system);
 - register data; and
 - cause of death data
- *linkable on aggregated level (e.g., regions).*
 - at the level of service providers, aggregated data from associations of physicians and hospitals;
 - cause of death statistics; and

- additional data, e.g., from cancer and implant registers, existing population survey data.

A further goal could be to enable linking data from individual research projects with existing datasets. For example, data from surveys or from Randomized Controlled Trials (with the consent of the study participants) could be linked to existing administrative data to increase the efficiency of data collection and enable further analyses (e.g., PopDataBC in Canada). One conceivable approach would be for study participants to be asked to provide their health insurance number or other information that would allow the subsequent data to be linked via the data linking site. The data could then be linked via the data linking site or a trust centre using the information provided, and a linked and pseudonymized dataset could then be provided for evaluation. Such an approach would only be feasible with the explicit consent of the study participants and in compliance with data protection and ethical regulations. This pseudonymized dataset could also be made available to other scientists (if necessary, after compliance with a protection period) for further analyses. As shown in this review, data linkage is possible even without unique patient identifiers by combining different information. However, the use of a uniform pseudonym, which is assigned in a uniform way at different points during data generation, would be desirable to enable a secure linkage. A context-specific technical analysis would be necessary to determine the feasibility of such an approach in each setting.

10.3 How can health information from electronic records be leveraged?

In many countries comprehensive health data are not available for secondary use. Some initiatives based on electronic health or patient records have started to address this issue but are not without limitations. Datasets that require “opting-in” entail the probability of distortion due to non-participation, and in particular an unequal representation of different population groups. It is conceivable that especially persons who participate little in public life, e.g., due to illness, and patients with diseases that are perceived as stigmatizing, will not participate.

In the international comparison opting out is more common than opting in (see Chapter 6). Australia began implementation of My Health Record

in 2019; all persons who meet the basic criteria were given three months to opt-out, and if they did not opt-out during this period a record was created for them automatically. A person's created My Health Record can be cancelled by them at any time. In the opt-out-period about 10% of the population had objected to participation. With My Health Record the data belong to the patient. This means that besides opting out completely, it is possible to object to the use of My Health Record data for research purposes only or object to the documentation or disclosure of specific information or to specific providers. Thus, there may be a distortion through "social desirability" in the resulting dataset. In Austria patients can opt out of ELGA similar to the one in Australia, either completely or for certain applications (e.g., e-findings). Although the availability of ELGA data for research purposes has been announced by the regulatory authorities, it could be difficult to implement. According to anecdotal evidence, ELGA departures have tripled since the announcement of research accessibility. In the NHS in the UK the so-called "National Data Opt Out" was introduced in 2018, a service that enables patients to regulate access to their confidential data.

Reference networks of selected service providers can help organize health data collection efforts and provide information at the provider level. Selected reference practices could also be used to test which further information would be useful additions to already possible analyses and should therefore be included in larger studies. The Nivel Primary Care Database in the Kingdom of the Netherlands has a similar objective. Based on data collected in the course of practice in outpatient care, the database examines, among other things, the extent to which service providers cooperate in the care of individual patients and how health policy measures affect care. Participating practices are given access to periodic feedback reports, which allows benchmarking and can contribute to quality assurance. A similar model is offered by the CPRD in the UK, which is frequently used for research. Here, too, it is important that both healthcare providers and the patients registered with them can decide for or against participation. Data delivery by selected service providers will presumably always have high limitations regarding representativeness. Therefore, such options should only be complementary to efforts for strengthening the secondary use of comprehensive, representative administrative data. For healthcare systems without a gatekeeping system, there is an additional layer of complexity: system-related limitations are to be expected, especially with regard to data completeness, since usually no single service provider is in possession of all relevant data.

10.4 Outlook and future work

The aim of this review was to identify interesting case studies and provide impetus for new initiatives towards enabling broader, better access to health and healthcare data within and across countries. The scope of the review was selective: it is certain that other countries, or even other case studies in the selected countries, can also provide useful insights for future action. Additionally, a more thorough examination of individual structural, technical, and political aspects is necessary to provide optimal support for a potential implementation of the considerations described here in different settings. None of the case studies described in this review captures all the data categories included in the review's conceptual framework; for instance, the use of environmental data in linkage efforts could not be examined in detail and requires further research. Finally, the use and analysis of unstructured data are uncommon and comes with additional challenges, but it is core to the discussions around Big Data and could unlock important additional options.

While the work underlying this review was undertaken before information on the regulatory proposal on the EHDS was available, its findings are in line with current thinking. As with the EHDS consultation, also at the national or regional level it is desirable that future data users (especially in the area of healthcare research) are included or at least consulted when new data governance initiatives are developed. Moreover, the advances in data generation, sharing, and linkage made during the COVID-19 pandemic need careful consideration before they can be meaningfully integrated into health information systems. An important concomitant consideration is the resources required to ensure that additional data collection and processing will yield meaningful benefits; these include financing and a trained and dedicated workforce at the provider level and in data agencies. Without them, emerging regulatory frameworks and technical investments will not be sufficient to foster the availability and use of better data for health and health-related research. Finally, as health data from sources that allow patients and providers to decide on their participation (such as electronic health records) are prone to systematic bias and lack of representativity, they should not be considered a replacement for comprehensive routine data of good quality when it comes to research purposes. Consequently, efforts aiming to expand research access to data from such modalities should not distract from the importance of strengthening the availability and accessibility of robust, comprehensive routine data.

References

- Adelson P et al. (2018). Keratinocyte cancers in South Australia: Incidence, geographical variability and service trends. *Australian and New Zealand Journal of Public Health*, 42(4):329–33. Available at: <https://doi.org/10.1111/1753-6405.12806> (accessed 30 December 2022)
- 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
- Ahsan R et al. (2022). Applications of GIS and geospatial analyses in COVID-19 research: A systematic review. *F1000Research*, 28;9:1379. Available at: <https://f1000research.com/articles/9-1379/v2> (accessed 11 December 2022)
- All European Academies, European Academies Science Advisory Council and Federation of European Academies of Medicine (2021). *International Sharing of Personal Health Data for Research*. Brussels, FEAM. Available at: International-Health-Data-Transfer_2021_web.pdf (allea.org) (accessed 30 December 2022)
- Anom BY (2022). The ethical dilemma of mobile phone data monitoring during COVID-19: The case for South Korea and the United States. *Journal of Public Health Research*, .11(3):22799036221102491. Available at: <https://pubmed.ncbi.nlm.nih.gov/35911427/> (accessed 30 December 2022)
- archyde (2021). *Cyber attack: Paris hospitals victims of massive theft of health data*. (<https://www.archyde.com/cyber-%E2%80%8B%E2%80%8Battack-paris-hospitals-victims-of-massive-theft-of-health-data/>, accessed 7 June 2022)
- Armstrong J et al. (2020). Dynamic linkage of COVID-19 test results between Public Health England's Second Generation Surveillance System and UK Biobank. *Microbial genomics*, 6(7):mgen000397. Available at: <https://doi.org/10.1099/mgen.0.000397> (accessed 30 December 2022)
- Australian Institute of Health and Welfare (AIHW) (2020). *Impacts of COVID-19 on Medicare Benefits Scheme and Pharmaceutical Benefits Scheme: quarterly data*. Sydney, AIHW, Australian Government (<https://www.aihw.gov.au/reports/health-care-quality-performance/impacts-of-covid19-mbs-pbs-quarterly-data/contents/about>, accessed 20 December 2022)
- Australian Institute of Health and Welfare (AIHW) (2021). Data Governance Framework 2021, Public Edition (italics). Sydney, AIHW, Australian Government (<https://www.aihw.gov.au/getmedia/c3e00f60-c40d-4989-ad22-de1be3ab5380/Data-Governance-Framework-2021.pdf.aspx>) (accessed 06 July 2023)
- Azzopardi-Muscat N et al. (2018). The role of the 2011 patients' rights in cross-border health care directive in shaping seven national health systems: looking beyond patient mobility. *Health Policy*, 122(3):279–83. Available at: <https://pubmed.ncbi.nlm.nih.gov/29317109/> (accessed 15 September 2022)

- Basit MA, Lehmann CU, Medford RJ (2021). Managing Pandemics with Health Informatics: Successes and Challenges. *Yearbook of Medical Informatics*, 01:1–348. Available at: <https://www.thieme-connect.com/products/ejournals/html/10.1055/s-0041-1726478> (accessed 30 December 2022)
- Baumgart DC (2020). Digital advantage in the COVID-19 response: perspective from Canada's largest integrated digitalized healthcare system. *Npj Digital Medicine*, 3 (114). Available at: <https://www.nature.com/articles/s41746-020-00326-y> (accessed 30 December 2022)
- Bhattacharya A et al. (2021). Healthcare-associated COVID-19 in England: A national data linkage study. *Journal of Infection*, 83(5):565–72. Available at: <https://doi.org/10.1016/j.jinf.2021.08.039> (accessed 30 December 2022)
- Black A (2016). *The IDI prototype spine's creation and coverage*. Wellington, Tatauraga Aotearoa, New Zealand Government (Statistics New Zealand Working Paper No 16–03). Available at: <https://www.stats.govt.nz/assets/Uploads/Integrated-data-infrastructure/Your-information-in-the-IDI/IDI-prototypes-spine-creation-and-coverage.pdf> (accessed 18 November 2019)
- Blasimme A, Ferretti A, Vayena E (2021). Digital Contact Tracing Against COVID-19 in Europe: Current Features and Ongoing Developments. *Frontiers in Digital Health*, 3:660823. Available at: <https://doi.org/10.3389/fdgh.2021.660823> (accessed 30 December 2022)
- Blasimme A, Vayena E (2016). Becoming Partners, Retaining Autonomy: Ethical Considerations on the Development of Precision Medicine. *BMC Medical Ethics*, 17:67. Available at: <https://doi.org/10.1186/s12910-016-0149-6> (accessed 30 December 2022)
- Blomberg N, Lauer KB (2020). Connecting data, tools and people across Europe: ELIXIR's response to the COVID-19 pandemic. *European Journal of Human Genetics*, 28:719–23. Available at: <https://doi.org/10.1038/s41431-020-0637-5> (accessed 30 December 2022)
- Boyd JH et al. (2012). Data linkage infrastructure for cross-jurisdictional health-related research in Australia. *BMC Health Services Research*, 12:480. Available at: <https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-12-480> (accessed 30 December 2022)
- Burke LG et al. (2019). Comparison of Costs of Care for Medicare Patients Hospitalized in Teaching and Nonteaching Hospitals. *JAMA Network Open*, 2(6):e195229. Available at: <https://doi.org/10.1001/jamanetworkopen.2019.5229> (accessed 30 December 2022)
- Canino MP et al. (2022). Epidemic forecasting based on mobility patterns: an approach and experimental evaluation on COVID-19 Data. *Social Network Analysis and Mining*, 12, 116. Available at: <https://doi.org/10.1007/s13278-022-00932-6> (accessed 20 December 2022)
- Carr MJ et al. (2021). Impact of COVID-19 on diagnoses, monitoring, and mortality in people with type 2 diabetes in the UK. *The Lancet. Diabetes & endocrinology*, 9(7):413–15. Available at: [https://doi.org/10.1016/S2213-8587\(21\)00116-9](https://doi.org/10.1016/S2213-8587(21)00116-9) (accessed 30 December 2022)
- Caseguard (2021). *The UK GDPR, New Data Privacy for British citizens*. Available at: (<https://caseguard.com/articles/the-uk-general-data-protection-regulation-data-privacy/>), accessed 7 June 2022)
- Choe SA, Jun YB, Kim SY (2018). Exposure to air pollution during preconceptual and prenatal periods and risk of hypertensive disorders of pregnancy: A retrospective cohort study in Seoul, Korea. *BMC Pregnancy and Childbirth*, 18(1). Available at: <https://doi.org/10.1186/s12884-018-1982-z> (accessed 30 December 2022)
- Cole A et al. (2015). *Data governance arrangements for real-world evidence*. Office of Health Economics, London, UK. Available at: <https://www.ohe.org/wp-content/uploads/2015/11/420-Data-Governance-for-RWE.pdf> (accessed 18 November 2019)
- Commonwealth Fund (2019). *The US Healthcare System*. New York, The Commonwealth Fund (International Health Care System Profiles). Available at: https://international.commonwealthfund.org/countries/united_states/ (accessed 18 November 2019)

- Council for International Organizations of Medical Sciences (CIOMS) and World Health Organization (WHO) (2017). *International Ethical Guidelines for Health-related Research Involving Humans*. Brussels, CIOMS and WHO. Available at: <https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf> (accessed 7 June 2022)
- COVID-19 National Emergency Response Centre, Epidemiology & Case Management Team, Korea Centres for Disease Control & Prevention (2020). Contact Transmission of COVID-19 in South Korea: Novel Investigation Techniques for Tracing Contacts. *Osong public health and research perspectives*, 11(1):60–3. Available at: <https://doi.org/10.24171/j.phrp.2020.11.1.09> (accessed 30 December 2022)
- DataGuidance (2020). *Alberta: OIPC issues statement on Government's contact tracing app*. DataGuidance (<https://www.dataguidance.com/news/alberta-oipc-issues-statement-governments-contact>), accessed 7 June 2022)
- 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. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5922494/>
- Davoody N et al. (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). Available at: <https://doi.org/10.1186/s12911-019-0816-x> (accessed 30 December 2022)
- de Biennassie K et al. (2022). *Health data and governance developments in relation to COVID-19. How OECD countries are adjusting health data systems for the new normal*. Paris, OECD Publishing (OECD Health Working Papers). Available at: https://www.oecd-ilibrary.org/social-issues-migration-health/health-data-and-governance-developments-in-relation-to-covid-19_aec7c409-en;jsessionid=mp695RAHrJ2EqD_kTAld0sGX.ip-10-240-5-103 (accessed 7 June 2022)
- De' R, Pandey N, Pal A (2020). Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. *International Journal of Information Management*, 55:102171. Available at: <https://doi.org/10.1016/j.ijinfomgt.2020.102171> (accessed 30 December 2022)
- Digital Health Europe (2022). *Glossary*. Digital Health Europe (<https://digitalhealtheurope.eu/glossary/>), accessed 30 December 2022)
- Donovan GH et al. (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):e226–34. Available at: [https://doi.org/10.1016/S2542-5196\(19\)30070-1](https://doi.org/10.1016/S2542-5196(19)30070-1) (accessed 30 December 2022)
- Dron L et al. (2022). Data capture and sharing in the COVID-19 pandemic: a cause of concern. *The Lancet Digital Health*, 4: e748–57. Available at: <https://www.thelancet.com/action/showPdf?pii=S2589-7500%2822%2900147-9> (accessed 20 December 2022)
- Durie MH (1989). The Treaty of Waitangi and health care. *New Zealand Medical Journal*, 102(869):283–5. Available at: <https://pubmed.ncbi.nlm.nih.gov/2733903/> (accessed 30 December 2023)
- eHealth Network, European Commission (2013). *Guidelines on minimum/nonexhaustive patient summary dataset for electronic exchange in accordance with the cross-border directive 2011/24/EU: Release 1*. Brussels, European Commission, eHealth Network. Available at: https://ec.europa.eu/health/system/files/2019-02/guidelines_patient_summary_en_0.pdf (accessed 30 December 2022)
- eHealth Network, European Commission (2021). *Guidelines on the electronic exchange of health data under Cross-Border Directive 2011/24/EU Patient Summary: Release 3*. Brussels, European Commission, eHealth Network. Available at: https://ec.europa.eu/health/publications/guidelines-patient-summary_en (accessed 7 June 2022)

- European Commission (2016). *Respect for private and family life*. Brussels, European Commission (https://ec.europa.eu/info/aid-development-cooperation-fundamental-rights/your-rights-eu/know-your-rights/freedoms/respect-private-and-family-life_en, accessed 7 June 2022)
- European Commission (2017). *European Reference Networks*. Brussels, European Commission. Available at: (https://ec.europa.eu/health/system/files/2019-03/2017_brochure_en_0.pdf, accessed 7 June 2022)
- European Commission (2018). General Data Protection Regulation (GDPR), Regulation (EU) 2016/679, *OJL 119, 04.05.2016; cor. OJL 127, 23.5.2018*. Available at: <https://gdpr-info.eu/> (accessed 11 May 2022).
- European Commission (2019). *eHealth, Interoperability of Health Data and Artificial Intelligence for Health and Care in the EU Lot 1 – Interoperability of Electronic Health Records in the EU*. (2020) SMART 2019/0056. Brussels, Publications Office of the European Union (Final report, <https://ec.europa.eu/newsroom/dae/redirection/document/79897>, accessed 7 June 2022)
- European Commission (2021). DG Sante, Unit B3. *European Health Data Space*. Brussels, European Commission (https://edps.europa.eu/sites/edp/files/publication/european-commission_dg_sante_25_january_2021_en.pdf, accessed 7 June 2022)
- European Commission (2022a). *Electronic cross-border health services*. Brussels, European Commission (https://ec.europa.eu/health/ehealth-digital-health-and-care/electronic-cross-border-health-services_en, accessed 7 June 2022)
- European Commission (2022b). *Proposal for a Regulation of the European Parliament and the Council on the European Health Data Space*. COM/2022/197 final. Strasbourg, European Commission. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0197> (accessed 7 June 2022)
- European Commission (2022c). *Mobile contact tracing apps in EU Member States*. European Commission. Brussels, European Commission (https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/travel-during-coronavirus-pandemic/mobile-contact-tracing-apps-eu-member-states_en, accessed 7 June 2022)
- European Commission, DG Health and Food Safety (2021). *Assessment of the EU Member States' rules on health data in the light of GDPR*. Brussels, Publications Office of the European Union (Specific Contract No SC 2019 70 02 in the context of the Single Framework Contract Chafea/2018/Health/03, https://ec.europa.eu/health/system/files/2021-02/ms_rules_health-data_en_0.pdf, accessed 7 June 2022)
- European Court of Auditors (2019). *Special Report No. 7 – EU actions for cross-border healthcare: significant ambitions but improved management required (pursuant to Article 287(4), second subparagraph, TFEU)*. Luxembourg, European Court of Auditors (https://www.eca.europa.eu/Lists/ECA/Documents/SR19_07/SR_HEALTH_CARE_EN.pdf, accessed 7 June 2022)
- European Court of Human Rights (2021). *Guide on Article 8 of the European Convention on Human Rights, Right to respect for private and family life, home and correspondence*. Strausbourg, Council of Europe (https://www.echr.coe.int/Documents/Guide_Art_8_ENG.pdf, accessed 7 June 2022)
- European Data Protection Supervisor (2020). *Data Protection*. Brussels, European Data Protection Supervisor (https://edps.europa.eu/data-protection_en, accessed 7 June 2022)
- European Medicines Agency (EMA) (2020). *GDPR and the secondary use of health data: Report from EMA workshop held with the EMA Patients' and Consumers' Working Party (PCWP) and Healthcare Professionals' Working Party (HCPWP) on 23 September 2020*. EMA/532436/2020. Amsterdam, EMA (https://www.ema.europa.eu/en/documents/report/report-workshop-application-general-data-protection-regulation-gdpr-area-health-secondary-use-data_en.pdf, accessed 7 June 2022)

- European Parliament (2011). *Summary of European Parliament Council Directive 2011/24/EU on the application of patients' rights in cross-border healthcare*. Document 32011L0024. Brussels, European Parliament (<https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=CELEX:32011L0024>, accessed 7 June 2022)
- European Parliament (2019). *Commission Recommendation (EU) 2019/243 of 6 February 2019 on a European Electronic Health Record exchange format*. Document 32019H0243. Brussels, European Parliament (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.039.01.0018.01.ENG, accessed 7 June 2022)
- Fiorini S et al. (2019). Predicting diabetes second-line therapy initiation in the Australian population via time span-guided neural attention network. *PLOS ONE*, 14(10):e0211844. Available at: <https://doi.org/10.1371/journal.pone.0211844> (accessed 30 December 2022)
- 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):582–7. Available at: <https://doi.org/10.1111/1753-6405.12819> (accessed 30 December 2022)
- Fry R, et al. (2021). Real-time spatial health surveillance: Mapping the UK COVID-19 epidemic. *International Journal of Medical Informatics*, 149:104400. Available at: <https://doi.org/10.1016/j.ijmedinf.2021.104400> (accessed 30 December 2022)
- Gabrijelcic M et al., eds. (2021). *Inequalities in health future challenges for intersectoral cooperation*. Ljubljana, NIJZ. Available at: <https://nijz.si/publikacije/inequalities-in-health-future-challenges-for-intersectoral-cooperation/> (accessed 12 March 2023)
- Ganguli I et al. (2019). Association Of Medicare's Annual Wellness Visit With Cancer Screening, Referrals, Utilization, And Spending. *Health Affairs*, 38(11):1927–35. Available at: <https://doi.org/10.1377/hlthaff.2019.00304> (accessed 30 December 2022)
- Gayle AV et al. (2019). Changing causes of death for patients with chronic respiratory disease in England, 2005–2015. *Thorax*, 74(5):483–91. Available at: <https://doi.org/10.1136/thoraxjnlg-2018-212514> (accessed 30 December 2022)
- Gentil ML et al. (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). Available at: <https://doi.org/10.1186/s12911-017-0538-x> (accessed 30 December 2022)
- de Gier B et al. (2021). COVID-19 vaccine effectiveness against hospitalizations and ICU admissions in the Netherlands, April–August 2021. Preprint. medRxiv. Available at: <https://doi.org/10.1101/2021.09.15.21263613> (accessed 30 December 2022)
- Gisev N et al. (2018). To what extent do data from pharmaceutical claims under-estimate opioid analgesic utilisation in Australia? *Pharmacoepidemiology and Drug Safety*, 27(5):550–5. Available at: <https://doi.org/10.1002/pds.4329> (accessed 30 December 2022)
- Glover L (2019). *The Australian Healthcare System*. New York, The Commonwealth Fund (International Health Care System Profiles). Available at: <https://international.commonwealthfund.org/countries/australia/> (accessed 18 November 2019)
- Gothe H (2014). Kapitel 19: Routinedaten im Ausland. In E Swart (ed.), *Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven*, 2nd, completely revised and extended edition. Bern, Huber: 304–13.
- Interactive data visualizations of COVID-19 (2022) [web site]. Ottawa, Government of Canada (<https://health-infobase.canada.ca/covid-19/visual-data-gallery/>, accessed 30 December 2022)
- Government of the UK (2021). Office for National Statistics. *Guidance, Linking with anonymized data – how not to make a hash of it*. London, Government of the UK, Office for National Statistics (<https://www.gov.uk/government/publications/joined-up-data-in-government-the-future-of-data-linking-methods/linking-with-anonymised-data-how-not-to-make-a-hash-of-it>, accessed 6 June 2022)

- Gupta S, Aitken N. (2022) 'COVID-19 mortality among racialized populations in Canada and its association with income', Statistics Canada, 30 August 2022. Available at: <https://www150.statcan.gc.ca/n1/pub/45-28-0001/2022001/article/00010-eng.htm?mpaign=statcan-all-content-22-23> (accessed 06 December 2022)
- Hamad AF et al. (2019). Prenatal antibiotics exposure and the risk of autism spectrum disorders: A population-based cohort study. *PLOS ONE*, 14(8):e0221921. Available at: <https://doi.org/10.1371/journal.pone.0221921> (accessed 30 December 2022)
- Hansen J, et al. (2021). Directorate-General for Health and Food Safety. *Assessment of the EU Member States' rules on health data in the light of GDPR*. Brussels, Publications Office of the European Union (https://health.ec.europa.eu/publications/assessment-eu-member-states-rules-health-data-light-gdpr_en, accessed 20 December 2022)
- 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). Available at: <https://doi.org/10.17061/phrp2751744> (accessed 30 December 2022)
- Häyrinen K, Saranto K, Nykänen P (2008). Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics*, 77:291–304. Available at: <https://pubmed.ncbi.nlm.nih.gov/17951106/> (accessed 30 December 2022)
- Health Data Hub (2022a). *Notre offre technologique*. Paris, Health Data Hub (<https://www.health-data-hub.fr/offre-technologique>, accessed 6 June 2022)
- Health Data Hub (2022b). *Projets*. Paris, Health Data Hub (<https://www.health-data-hub.fr/projets>, accessed 30 December 2022)
- Health Data Hub and L'Assurance Maladie Caisse Nationale (2019). *Schémas d'appariement et de circulation des données*. Paris, Health Data Hub (<https://www.health-data-hub.fr/sites/default/files/2021-05/Sch%C3%A9mas%20d%C3%A9partement%20et%20de%20circulation%20des%C3%A9tats%20du%20Cnam.pdf>, accessed 6 June 2022)
- Heikkilä M, Cerulus L (2020). *Hacker seeks to extort Finnish mental health patients after data breach*. Politico, 26 October 2020 (<https://www.politico.eu/article/cybercriminal-extorts-finnish-therapy-patients-in-shocking-attack-ransomware-blackmail-vastaamo/>, accessed 7 June 2022)
- Huygens MWJ et al. (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. Available at: <https://doi.org/10.1136/bmjopen-2017-019233> (accessed 30 December 2022)
- ICES (2022). *Publications, Journal Articles*. Ontario, ICES (<https://www.ices.on.ca/Publications/Journal-Articles>, accessed 30 December 2022)
- ISD Scotland (2016). *Quality and Outcomes Framework (QOF): Prevalence, achievement, payment and exceptions data for Scotland, 2015/2016*. Edinburgh, NHS Scotland (<https://www.isdscotland.org/Health-Topics/General-Practice/Quality-And-Outcomes-Framework/>, accessed 18 November 2019)
- ISD Scotland (2019). *eDRIS Frequently Asked Questions*. Edinburgh, NHS Scotland (<https://www.isdscotland.org/Products-and-Services/eDRIS/FAQ-eDRIS/>, accessed 18 November 2019)
- ISRCTN registry (2020). *ISRCTN50189673: A randomised trial of treatments to prevent death in patients hospitalised with COVID-19 (coronavirus)*. BMC. Available at: <https://doi.org/10.1186/ISRCTN50189673> (accessed 30 December 2022)
- IT Governance (2020). *What is data loss and how does it work?* IT Governance. 6 October 2020 (<https://www.itgovernance.co.uk/blog/what-is-data-loss-and-how-does-it-work>, accessed 6 June 2022)
- Jansen T et al. (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. Available at: <https://doi.org/10.1136/bmjopen-2018-026426> (accessed 30 December 2022)

- Kand H, Kwon S, Kim E (2020). *COVID-19 Health System Response Monitor: Republic of Korea*. New Delhi, Word Health Organization Regional Office for South-East Asia. Available at: <https://apo.who.int/publications/i/item/9789290228219> (accessed 30 December 2022)
- Katz A et al.(2018). Association of the Social Determinants of Health With Quality of Primary Care. *The Annals of Family Medicine*, 16(3):217–24. Available at: <https://doi.org/10.1370/afm.2236> (accessed 30 December 2022)
- Keskimäki I et al. (2019). Finland: Health system review. *Health systems in transition*, 21(2):1–166. Available at: <https://apps.who.int/iris/handle/10665/327538> (accessed 30 December 2022) 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*, 36:e2014008. Available at: <https://doi.org/10.4178/epih.e2014008> (accessed 30 December 2022)
- Kim TJ et al. (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):e343. Available at: <https://doi.org/10.3346/jkms.2018.33.e343> (accessed 30 December 2022)
- Kittel B et al. (2020). The Austrian Corona Panel Project: monitoring individual and societal dynamics amidst the COVID-19 crisis. *European Political Science*. Available at: <https://europemc.org/article/PMC/PMC7579903> (accessed 30 December 2022)
- Ko YJ et al. (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). Available at: <https://doi.org/10.3346/jkms.2018.33.e163> (accessed 30 December 2022)
- Koprivnikar H et al. eds. (2021). *Tobacco-attributable mortality in Slovenia 1997–2019*. Ljubljana, NIJZ. Available at: <https://nijz.si/publikacije/tobaku-pripisljiva-umrljivost-v-sloveniji-1997-2019/> (accessed 20 December 2022) Kucharski AJ et al. (2021). Sharing, synthesis and sustainability of data analysis for epidemic preparedness in Europe. *The Lancet*. Available at: [https://www.thelancet.com/journals/lanepublications/article/PiIS2666-7762\(21\)00192-7/fulltext](https://www.thelancet.com/journals/lanepublications/article/PiIS2666-7762(21)00192-7/fulltext) (accessed 30 December 2022)
- Lee JL, Haupt JP (2021). Scientific globalism during a global crisis: research collaboration and open access publications on COVID-19. *Higher Education*, 81, 949–966. Available at: <https://doi.org/10.1007/s10734-020-00589-0> (accessed 12 March 2023)
- Lee D, Lee J (2020). Testing on the move: Republic of Korea's rapid response to the COVID-19 pandemic. *Transportation Research Interdisciplinary Perspectives*, 5(2020): 100111. Available at: <https://www.sciencedirect.com/science/article/pii/S2590198220300221> (accessed 30 December 2022)
- Lee J, Lee JS, Park SH, Shin SA, Kim K (2017). Cohort Profile: The National Health Insurance Service-National Sample Cohort (NHIS-NSC), Republic of Korea. *International Journal of Epidemiology*, 46(2):e15. Available at: <https://doi.org/10.1093/ije/dyw319> (accessed 30 December 2022)
- Lu E et al. (2014). Birth Outcomes In Newborns Fathered By Men With Multiple Sclerosis Exposed To Disease-Modifying Drugs (P4.155). *Neurology*, 82(10 Supplement). Available at: https://n.neurology.org/content/82/10_Supplement/P4.155 (accessed 18 November 2019)
- Majeed A et al. (2022). How data from the United Kingdom has guided covid-19 vaccine policies. *BMJ*, 376:o839 doi:10.1136/bmj.o839. Available at: <https://pubmed.ncbi.nlm.nih.gov/35354582/> (accessed 20 December 2022)
- Manitoba Centre for Health Policy (2020). *The Manitoba Population Research Data Repository*. Winnipeg, University of Manitoba (<https://umanitoba.ca/manitoba-centre-for-health-policy/data-repository>, accessed 7 June 2022)
- 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):4056–62. Available at: <https://doi.org/10.1007/s10803-018-3609-7> (accessed 30 December 2022)

- McGrail K (2022). Excess mortality. COVID-19 and health care systems in Canada. *Canadian Medical Association Journal*, 194(21): E741-E745. Available at: <https://doi.org/10.1503/cmaj.220337> (accessed 7 June 2022)
- Menni C et al. (2022). Symptom prevalence, duration, and risk of hospital admission in individuals infected with SARS-CoV-2 during periods of omicron and delta variant dominance: a prospective observational study from the ZOE COVID Study. *The Lancet*, 23;399(10335):1618–1624. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC899396/> (accessed 30 December 2022)
- Medicines & Healthcare products Regulatory Agency (2022). *CPRD linked data*. London, MHRA (<https://cprd.com/cprd-linked-data>), accessed 7 June 2022)
- MHRA. *CPRD, Bibliography*. London, MHRA (<https://cprd.com/bibliography>), accessed 30 December 2022)
- Milioris K, Papageorgiou K (2021). A Study of Healthcare ICT Systems and their Usefulness During CoVid-19 Focused in the European Environment. *Journal of Hospital and Health Care Administration*, 4:149. Available at: <https://www.gavinpublishers.com/article/view/a-study-of-healthcare-ict-systems-and-their-usefulness-during-covid19-focused-in-the-european-environment-1> (accessed 20 December 2022)
- Milne BJ et al. (2019). Data Resource Profile: The New Zealand Integrated Data Infrastructure (IDI). *International Journal of Epidemiology*, 48(3):677–677e. Available at: <https://doi.org/10.1093/ije/dyz014> (accessed 30 December 2022)
- Morgan A et al. (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. Available at: <https://doi.org/10.1371/journal.pone.0208824> (accessed 30 December 2022)
- Mues K et al. (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*, 9:267–77. Available at: <https://doi.org/10.2147/CLEP.S105613> (accessed 30 December 2022)
- Muthuppalaniappan M, Stevenson K (2021). Healthcare cyber-attacks and the COVID-19 pandemic: an urgent threat to global health. *International journal for quality in health care: journal of the International Society for Quality in Health Care*, 33(1):mzaa117. Available at: <https://doi.org/10.1093/intqhc/mzaa117> (accessed 30 December 2022)
- National Health Service (NHS) Scotland (2020). *How do we link the data?* Edinburgh, NHS Scotland (https://www.isdscotland.org/Products-and-Services/eDRIS/_docs/2b-Indexing-and-Linkage-Process-Diagram.pdf), accessed 6 June 2022)
- Negro-Caldwell E et al. (2021). Health Information Systems in the COVID-19 Pandemic: A Short Survey of Experiences and Lessons Learned from the European Region. *Frontiers in Public Health*, 9:676838. Available at: <https://doi.org/10.3389/fpubh.2021.676838> (accessed 30 December 2022)
- New Zealand Constitutional Law Resources (2017). *Additional Instructions from Lord Normanby to Captain Hobson 1839 [1839] NZConLRes 3 (15 August 1839)*. Dunedin, Christchurch, Wellington, New Zealand Legal Information Institute (<http://www.nzlii.org/nz/other/NZConLRes/1839/3.html>), accessed 6 July 2022)
- NHS Digital (2022). *Compliance with the national data opt-out*. London, NHS Digital (<https://digital.nhs.uk/services/national-data-opt-out/compliance-with-the-national-data-opt-out>), accessed 6 July 2022)
- Nickel NC et al. (2014). Data Resource Profile: Pathways to Health and Social Equity for Children (PATHS Equity for Children). *International Journal of Epidemiology*, 43(5):1438–49. Available at: <https://doi.org/10.1093/ije/dyu190> (accessed 30 December 2022)
- National Institute of Public Health (NIJZ), Slovenia (2016). *Health in the Municipality*. Ljubljana, NIJZ (<https://www.nijz.si/en/health-in-municipality>), accessed 6 June 2022)

- NIJZ (2018). *Slovene eHealth Solutions*. Ljubljana, NIJZ (<https://nijz.si/en/slovene-ehealth-solutions>, accessed 3 June 2022)
- Nyberg F et al. (2021). Swedish Covid-19 Investigation for Future Insights – A Population Epidemiology Approach Using Register Linkage (SCIFI-PEARL). *Clinical Epidemiology*, 13:649–59. Available at: <https://pubmed.ncbi.nlm.nih.gov/34354377/> (accessed 30 December 2022)
- Office of the Australian Information Commissioner (OAIC) (2022). *The COVID Safe app and my privacy rights*. Australian Government Office of the Australian Information Commissioner. Canberra, OAIC, Australian Government (<https://www.oaic.gov.au/privacy/covid-19/the-covidsafe-app-and-my-privacy-rights>, accessed 7 June 2022)
- Oderkirk J (2021). *Survey results: National health data infrastructure and governance. Readiness of electronic health record systems to contribute to national health information and research*. Paris, OECD Publishing (OECD Health Working Papers, No. 127). Available at: <https://doi.org/10.1787/18152015> (accessed 30 December 2022)
- Oderkirk J, Ronchi E, Klazinga N (2013). International comparisons of health system performance among OECD countries: Opportunities and data privacy protection challenges. *Health Policy*, 112(1–2):9–18. Available at: <https://doi.org/10.1016/j.healthpol.2013.06.006> (accessed 30 December 2022)
- Organisation for Economic Cooperation and Development (OECD) (1980, amended 2013). *Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*. OECD/LEGAL/0188. Paris, OECD (OECD Legal Instruments Series, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0188#dates>, accessed 7 June 2022)
- OECD (2013). *The OECD Privacy Framework*. Paris, OECD Publishing (https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf, accessed 7 June 2022)
- OECD (2015). *Health Data Governance: Privacy, Monitoring and Research*. Paris, OECD Publishing (OECD Health Policy Studies, <https://www.oecd.org/publications/health-data-governance-9789264244566-en.htm>, accessed 30 December 2022)
- OECD (2017). *OECD Recommendation of the Council on Health Data Governance*. OECD/LEGAL/0433. Paris, OECD Publishing (OECD Legal Instruments Series, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433>, access 30 December 2022)
- OECD (2019). *Using Routinely Collected Data to Inform Pharmaceutical Policies Analytical Report for OECD and EU countries*. Paris, OECD Publishing (<https://www.oecd.org/health/health-systems/Using-Routinely-Collected-Data-to-Inform-Pharmaceutical-Policies-Analytical-Report-2019.pdf>, (30 December 202s)
- OECD (2022). *Glossary of Statistical Terms*. Paris, OECD (<https://stats.oecd.org/glossary/search.asp>, accessed 30 December 2022)
- Office of the National Coordinator for Health Information Technology (ONC), US (2019). *What are the differences between electronic medical records, electronic health records, and personal health records?* Washington D.C., ONC (<https://www.healthit.gov/faq/what-are-differences-between-electronic-medical-records-electronic-health-records-and-personal>, accessed 18 November 2019)
- Official Journal of the European Communities (2000). *Charter of Fundamental Rights of the European Union (2000/C 364/01)*. Brussels, European Parliament (https://www.europarl.europa.eu/charter/pdf/text_en.pdf#text=Article%208%20Protection%20of%20personal%20data%20Everyone%20has,protection%20of%20personal%20data%20concerning%20him%20or%20her, accessed 6 June 2022)
- Open Access Government (2019). *Data protection in the healthcare sector. Open Access Government* (Cheshire, UK), 11 December 2019 (<https://www.openaccessgovernment.org/data-protection-in-theHealthcare-sector/79169/>, accessed 30 December 2022)
- Padmanabhan S et al. (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):91–9. Available at: <https://doi.org/10.1007/s10654-018-0442-4> (accessed 30 December 2022)

- Panteli D et al. (2015). Continuity of care in the cross-border context: insights from a survey of German patients treated abroad. *European Journal of Public Health*, 25(4):557–63. Available at: <https://pubmed.ncbi.nlm.nih.gov/25667154/#:~:text=Continuity%20of%20care%20in%20the%20cross-border%20context%3A%20insights,need%20to%20be%20addressed%20in%20the%20cross-border%20context> (accessed 30 December 2022)
- Paprica PA et al. (2020). Prospective data linkage to facilitate COVID-19 trials – A call to action. *International Journal of Population Data Science*, 5(2):1383. Available at: <https://pubmed.ncbi.nlm.nih.gov/32935062/> (accessed 30 December 2022)
- Park J et al. (2017). Antibiotic use in South Korea from 2007 to 2014: A health insurance database-generated time series analysis. *PLOS ONE*, 12(5):e0177435. Available at: <https://doi.org/10.1371/journal.pone.0177435> (accessed 30 December 2022)
- Park S, Choi GJ, Ko H (2020). Information Technology-Based Tracing Strategy in Response to COVID-19 in South Korea-Privacy Controversies. *JAMA*, 323(21):2129–2130. Available at: <https://pubmed.ncbi.nlm.nih.gov/32324202/> (accessed 30 December 2022) Park YJ et al. (2020). Development and Utilization of a Rapid and Accurate Epidemic Investigation Support System for COVID-19. *Osong Public Health Res Perspect*, 11(3):118–127. Available at: <https://pubmed.ncbi.nlm.nih.gov/32528817/> (accessed 30 December 2022)
- Peralta-Santos A, Perelman J (2018). Who wants to cross borders in the EU for healthcare? An analysis of the Eurobarometer data in 2007 and 2014. *European Journal of Public Health*, 28(5):879–84. Available at: <https://pubmed.ncbi.nlm.nih.gov/29697799/> (Accessed 6 June 2022)
- Perlman C et al. (2019). Access to Psychiatrist Services for Older Adults in Long-Term Care: A Population-Based Study. *Journal of the American Medical Directors Association*, 20(5):610–616.e2. Available at: <https://doi.org/10.1016/j.jamda.2019.01.121> (accessed 30 December 2022)
- Pilcher D et al. (2021). A national system for monitoring intensive care unit demand and capacity: the Critical Health Resources Information System (CHRIS). *Medical Journal of Australia*, 214(7): 297–298.e1. Available at: https://www.mja.com.au/system/files/issues/214_07/mja250988.pdf (accessed 15 December 2022)
- Pooransingh S et al. (2021). COVID-19 Highlights a Critical Need for Efficient Health Information Systems for Managing Epidemics of Emerging Infectious Diseases. *Frontiers in Public Health*, 9:767835. Available at: <https://pubmed.ncbi.nlm.nih.gov/34957023/> (accessed 30 December 2022)
- Population Data BC (2022). *Research projects*. Vancouver, PopData (<https://www.popdata.bc.ca/ria/projects>), accessed 30 December 2022)
- Population Health Research Network (PHRN), Publications, Research Publications [web site]. Canberra, Australia Government (<https://www.phrn.org.au/publications/research-publications>), accessed 30 December 2022)
- Pottegård A et al. (2020). Existing Data Sources in Clinical Epidemiology: The Danish COVID-19 Cohort. *Clinical Epidemiology*, 12:875–81. Available at: <https://pubmed.ncbi.nlm.nih.gov/32848476/> (accessed 20 December 2022)
- Prieto-Alhambra D et al. (2021). Unraveling COVID-19: a large-scale characterization of 4.5 million COVID-19 cases using CHARYBDIS. *Res Sq* [Preprint]. Available at: <https://pubmed.ncbi.nlm.nih.gov/33688639/> (accessed 30 December 2022).
- Rados S et al. eds. (2022). *Alcohol consumption and impact on health of alcohol use in Slovenia in the period 2013–2018, trends*. Ljubljana, NIJZ. Available at: <https://nijz.si/publikacije/poraba-alkohola-in-zdravstvene-posledice-rabe-alkohola-v-sloveniji-v-obdobju-2013-2018-trendi/> (accessed 30 December 2022)
- Reeves JJ et al. (2021). The Clinical Information Systems Response to the COVID-19 Pandemic. *Yearbook of Medical Informatics*, 01: 1–348. Available at: <https://www.thieme-connect.com/products/ejournals/html/10.1055/s-0041-1726513> (accessed 20 December 2022)

- Republique Francaise (2022). *Direction de la recherche, des études, de l'évaluation et de statistiques*, Paris, Republique Francaise (https://drees.solidarites-sante.gouv.fr/recherche?f%5B0%5D=content_type%3A1, accessed 7 June 2022)
- ResDAC (2022) [web site]. Minneapolis, Research Data Assistance Center, University of Minnesota (<https://resdac.org/>, accessed 7 June 2022)
- Rice T et al. (2020). United States: Health system review. *Health Systems in Transition*, 22(4):1–479. Available at: <https://eurohealthobservatory.who.int/publications/i/united-states-health-system-review-2020> (accessed 20 December 2022)
- RKKP (2022) [web site]. Aarhus, Regionernes kliniske kvalitetsudviklingsprogram. (<https://www.rkkp.dk/>, accessed 7 June 2022)
- Robbin A (2001). The loss of personal privacy and its consequences for social research. *Journal of Government Information*, 28(5):493–527. Available at: [https://doi.org/10.1016/S1352-0237\(02\)00343-X](https://doi.org/10.1016/S1352-0237(02)00343-X) (accessed 30 December 2022)
- Rosker S, Paska AV. eds. (2021). *Suicide in Slovenia and around the world*. Ljubljana, NIJZ. Available at: https://njjz.si/wp-content/uploads/2022/07/samomor_v_sloveniji_elektronska_izdaja_25_10_21.pdf (accessed 12 March 2023)
- Rumbold JM, Pierscionek B (2017a). The Effect of the General Data Protection Regulation on Medical Research. *Journal of Medical Internet Research*, 19(2):e47. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5346164/> (accessed 07 June 2022)
- Rumbold JM, Pierscionek B (2017b). A critique of the regulation of data science in healthcare research in the European Union. *BMC Medical Ethics*, 18:27. Available at: <https://doi.org/10.1186/s12910-017-0184-y> (accessed 30 December 2022)
- Sadeghieh T et al. (2022). Real-Time Forecasting of COVID-19 Cases Using Human Mobility in Ontario, Canada. *International Journal of Infectious Diseases* (Italic), 116: S110–S111. Available at: <https://www.sciencedirect.com/science/article/pii/S120197122101153X> (accessed 06 December 2022)
- Safer Care Victoria (2021). *Data linkage plays a key role in COVID-19 pandemic response*. Melbourne, Safer Care Victoria (<https://www.safercare.vic.gov.au/news-and-media/data-linkage-plays-a-key-role-in-covid-19-pandemic-response>, accessed 30 December 2022)
- Salerno S et al. (2021). Comprehensive evaluation of COVID-19 patient short- and long-term outcomes: Disparities in healthcare utilization and post-hospitalization outcomes. *PLOS ONE*, 16(10): e0258278. Available at: <https://doi.org/10.1371/journal.pone.0258278> (accessed 30 December 2022)
- Sax Institute (2019). *SURE Users – Summary Steps*. Glebe, NSW, Sax Institute (https://www.saxinstitute.org.au/wp-content/uploads/SURE_flow_diagram.pdf, accessed 18 November 2019)
- Scheibner J et al. (2020). Data protection and ethics requirements for multisite research with health data: a comparative examination of legislative governance frameworks and the role of data protection technologies. *Journal of Law and the Biosciences*, 7(1). Available at: <https://academic.oup.com/jlb/article/7/1/lssa010/5825716> (accessed 20 December 2022)
- Schmidt AE, Abboud LA, Bogaert P (2021). Making the case for strong health information systems during a pandemic and beyond. *Archives of Public Health*, 79:13. Available at: <https://doi.org/10.1186/s13690-021-00531-5> (accessed 30 December 2022)
- Schønning K et al. (2021). Electronic reporting of diagnostic laboratory test results from all health care sectors is a cornerstone of national preparedness and control of COVID-19 in Denmark. *APMIS Acta Pathologica, Microbiologica et Immunologica Scandinavica*, 129(7):438–51. Available at: <https://doi.org/10.1111/apm.13140> (accessed 30 December 2022)

- Schweikardt C et al.(2016). The historical development of the Dutch Sentinel General Practice Network from a paper based into a digital primary care monitoring system. *Journal of Public Health*, 24(6):545–62. Available at: <https://link.springer.com/article/10.1007/s10389-016-0753-4> (accessed 20 December 2022)
- Shin S et al. (2019). Ambient Air Pollution and the Risk of Atrial Fibrillation and Stroke: A Population-Based Cohort Study. *Environmental Health Perspectives*, 127(8):087009. Available at: <https://doi.org/10.1289/EHP4883> (accessed 30 December 2022)
- Shull JG (2019). Digital Health and the State of Interoperable Electronic Health Records. *JMIR Medical Informatics*, 7(4):e12712. Available at: <https://doi.org/10.2196/12712> (accessed 30 December 2022)
- Siegal R et al. (2021). Using geographic information systems to inform the public health response to COVID-19 and structural racism: The role of place-based initiatives. *Journal of Community Psychology*, 10.1002/jcop.22771. Advance online publication. Available at: <https://doi.org/10.1002/jcop.22771> (accessed 30 December 2022)
- Smeets HM et al. (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). Available at: <https://doi.org/10.1186/s12913-018-3528-5> (accessed 30 December 2022)
- Smith CD, Mennis J (2020). Incorporating Geographic Information Science and Technology in Response to the COVID-19 Pandemic. *Preventing Chronic Disease*, 17:E58. Available at: <https://doi.org/10.5888/pcd17.200246> (accessed 30 December 2022)
- Song J et al. (2021). Individuals with lived experience of homelessness and their COVID-19 outcomes in Wales: a data linkage study. *The Lancet*, 398:S81. Available at: <https://www.sciencedirect.com/science/article/pii/S0140673621026246> (accessed 20 December 2022)
- Song S (2017). HIRA under fire for 'selling personal data' to private insurers. *Korea Biomedical Review*. Available at: <http://www.koreabiomed.com/news/articleView.html?idxno=1745> (accessed 18 November 2019)
- Spilsbury K et al. (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. Available at: <https://doi.org/10.3389/fpubh.2017.00013> (accessed 30 December 2022)
- SPIRE (2019). *Your questions answered*. Edinburgh, Scottish Primary Care Information Resource, NHS Scotland (<https://spire.scot/your-questions-answered/>, accessed 18 November 2019)
- Spittal MJ et al. (2014). Suicide in adults released from prison in Queensland, Australia: A cohort study. *Journal of Epidemiology and Community Health*, 68(10):993–8. Available at: <https://doi.org/10.1136/jech-2014-204295> (accessed 30 December 2022)
- Statistics Canada (2020). *2020 submissions*. Ottawa, StatCan, Government of Canada (<https://www.statcan.gc.ca/en/record/2020>, accessed 30 December 2022)
- Statistics New Zealand (Stats NZ) Tatauranga Aotearoa (2014). *Linking methodology used by Statistics New Zealand in the Integrated Data Infrastructure project*. Wellington, Statistics New Zealand (<https://www.stats.govt.nz/assets/Uploads/Retirement-of-archive-website-project-files/Methods/Linking-methodology-used-by-Statistics-New-Zealand-in-the-Integrated-Data-Infrastructure-project/linking-methodology-IDId-project.pdf>, accessed 18 November 2019)
- Stats NZ (2020). *Data in the IDI*. Wellington, Statistics New Zealand (<https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/data-in-the-idi/>, accessed 30 December 2022)
- Stats NZ (2020). *Te Whata storehouse of iwi data welcomed by Stats NZ*. Wellington, Statistics New Zealand (<https://www.stats.govt.nz/news/te-whata-storehouse-of-iwi-data-welcomed-by-stats-nz>, accessed 12 March 2023)

- 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*, 155(6):1158–65. Available at: <https://pubmed.ncbi.nlm.nih.gov/30922950/> (accessed 30 December 2022)
- Sundhed.dk (2016). *Background*. Copenhagen, Sundhed.dk (<https://www.sundhed.dk/borger/service/om-sundheddk/om-organisationen/ehealth-in-denmark/background/>, accessed 6 June 2022)
- Sundhed.dk (2022). *e-Health in Denmark*. Copenhagen, Sundhed.dk (<https://www.sundhed.dk/borger/service/om-sundheddk/om-organisationen/ehealth-in-denmark/>, accessed 30 December 2022)
- Szocska M et al. (2021). Countrywide population movement monitoring using mobile devices generated (big) data during the COVID-19 crisis. *Scientific Reports*, 11(1):5943. Available at: <https://doi.org/10.1038/s41598-021-81873-6> (accessed 30 December 2022)
- Tanke MAC et al. (2019). A challenge to all. A primer on inter-country differences of high-need, high-cost patients. *PLoS One*, 14(6):e0217353. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6583982/> (accessed 20 December 2022).
- Tew M et al. (2017). Growth of linked hospital data use in Australia: a systematic review. *Australian Health Review* 41(4):394–400. Available at: <https://www.publish.csiro.au/AH/AH16034> (accessed 30 December 2022)
- Trudeau R (2017). Social Data Linkage Environment. *International Journal of Population Data Science*, 1(1). Available at: <https://doi.org/10.23889/ijpds.v1i1.76> (accessed 30 December 2022)
- Tumiene B et al. (2021). European Reference Networks: challenges and opportunities. *Journal of Community Genetics*, 12:217–29. Available at: <https://doi.org/10.1007/s12687-021-00521-8> (accessed 30 December 2022)
- UK Government (1948). *Guidance, The NHS Constitution updated 1 January 2021*. London, UK Government, Department of Health & Social Care (<https://www.gov.uk/government/publications/the-nhs-constitution-for-england/the-nhs-constitution-for-england>, accessed 7 June 2022)
- Understanding Patient Data (2018). *What are the risks around patient data?* London, NHS Confederation (<https://understandingpatientdata.org.uk/weighing-up-risks>, accessed 20 February 2022)
- UNESCO (1997). Universal Declaration on the Human Genome and Human Rights. Paris, United Nations Educational Scientific and Cultural Organisation (UNESCO) (<https://en.unesco.org/about-us/legal-affairs/universal-declaration-human-genome-and-human-rights>, accessed 20 February 2022)
- UNESCO (2005). Universal Declaration on Bioethics and Human Rights. Paris, UNESCO (<https://en.unesco.org/about-us/legal-affairs/universal-declaration-bioethics-and-human-rights>, accessed 30 December 2022)
- United Nations (1948). Universal Declaration of Human Rights. Geneva, United Nations (<https://www.un.org/sites/un2.un.org/files/2021/03/udhr.pdf>, accessed 6 June 2022)
- United States (US), Centers for Disease Control (CDC) (2021). *CDC's COVID-19 Data Improvement*. Atlanta, US Department of Health and Human Services, CDC (<https://www.cdc.gov/coronavirus/2019-ncov/science/data-improvements.html>, accessed 30 December 2022)
- Verheij RA et al. (2018). Possible Sources of Bias in Primary Care Electronic Health Record Data Use and Reuse. *Journal of Medical Internet Research*, 20(5):e185. Available at: <https://pubmed.ncbi.nlm.nih.gov/29844010/> (Accessed 30 December 2022)
- Verra SE, Kroeze R, Ruggeri K (2016). Facilitating safe and successful cross-border healthcare in the European Union. *Health Policy*, 120(6):718–27. Available at: <https://pubmed.ncbi.nlm.nih.gov/27142178/> (accessed 20 December 2022)

- Vezyridis P, Timmons S (2016). Evolution of primary care databases in UK: a scientometric analysis of research output. *BMJ Open* 6:e012785. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5073525/> (accessed 30 December 2022)
- Vos T et al. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258):1204–22. Available at: [https://doi.org/10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9) (accessed 30 December 2022)
- Waitangi Tribunal (2019). *Report on stage one of health services and outcomes released*. Wellington, New Zealand Ministry of Justice, Waitangi Tribunal (<https://waitangitribunal.govt.nz/news/report-on-stage-one-of-health-services-and-outcomes-released/>, access 30 December 2022)
- Wammes J, et al. (2019). *The Dutch Healthcare System*. New York, The Commonwealth Fund (International Health Care System Profiles). Available at: <https://international.commonwealthfund.org/countries/netherlands/> (accessed 18 November 2019)
- Warren N, Zhang Y (2018). Analysis of correlation between prevalence of certain health conditions and ownership of private health insurance [CURF2018–17]. Wellington, StatsNZ (<https://cdm20045.contentdm.oclc.org/digital/collection/p20045coll17/id/738>, accessed 18 November 2019)
- Wilkinson MD et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:160018. Available at: <https://pubmed.ncbi.nlm.nih.gov/26978244/> (accessed 7 June 2022)
- WMA (2002). *Declaration of Taipei, Research on Health Databases, Big Data and Biobanks. Revised 2016*. Ferney-Voltaire, World Medical Association (<https://www.wma.net/what-we-do/medical-ethics/declaration-of-taipei/#:~:text=The%20Declaration%20of%20Taipei%20tries%20to%20achieve%20a,become%20a%20very%20powerful%20tool%20for%20increasing%20knowledge>, accessed 7 June 2022)
- Wolf A et al. (2019). Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *International Journal of Epidemiology*, 48(6):1740–1740g. Available at: <https://pubmed.ncbi.nlm.nih.gov/30859197/> (accessed 7 June 2022).
- Wood A et al. (2021). Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. *BMJ*, 373:n826. Available at: <https://www.bmj.com/content/373/bmj.n826.long> (accessed 30 December 2022)
- Young A, Flack F (2018). Recent trends in the use of linked data in Australia. *Australian Health Review*, 42(5):584. Available at: <https://doi.org/10.1071/AH18014> (accessed 30 December 2022)
- Yuan B, Li J (2019). The Policy Effect of the General Data Protection Regulation (GDPR) on the Digital Public Health Sector in the European Union: An Empirical Investigation. *International Journal of Environmental Research and Public Health*, 16(6):1070. Available at: <https://pubmed.ncbi.nlm.nih.gov/30934648/> (accessed 30 December 2022)
- Zoe Health Study (2022). London, Zoe Health Study (<https://health-study.joinzoe.com/the-health-study>, accessed 30 December 2022)

HiT methodology and production process

HiTs are produced by country experts in collaboration with the Observatory's research directors and staff. They are based on a template that, revised periodically, provides detailed guidelines and specific questions, definitions, suggestions for data sources and examples needed to compile reviews. While the template offers a comprehensive set of questions, it is intended to be used in a flexible way to allow authors and editors to adapt it to their particular national context. The most recent template is available online at: <https://eurohealthobservatory.who.int/publications/i/health-systems-in-transition-template-for-authors>.

Authors draw on multiple data sources for the compilation of HiTs, ranging from national statistics, national and regional policy documents to published literature. Furthermore, international data sources may be incorporated, such as those of the OECD and the World Bank. The OECD Health Data contain over 1200 indicators for the 34 OECD countries. Data are drawn from information collected by national statistical bureaux and health ministries. The World Bank provides World Development Indicators, which also rely on official sources.

In addition to the information and data provided by the country experts, the Observatory supplies quantitative data in the form of a set of standard comparative figures for each country, drawing on the European Health for All database. The Health for All database contains more than 600 indicators defined by the WHO Regional Office for Europe for the purpose of monitoring Health in All policies in Europe. It is updated for distribution twice a year from various sources, relying largely upon official figures provided by governments, as well as health statistics collected by the technical units of the WHO Regional Office for Europe. The standard Health for All data

have been officially approved by national governments. With its summer 2007 edition, the Health for All database started to take account of the enlarged EU of 27 Member States.

HiT authors are encouraged to discuss the data in the text in detail, including the standard figures prepared by the Observatory staff, especially if there are concerns about discrepancies between the data available from different sources.

A typical HiT consists of nine chapters.

1. Introduction: outlines the broader context of the health system, including geography and sociodemography, economic and political context, and population health.
2. Organisation and governance: provides an overview of how the health system in the country is organised, governed, planned and regulated, as well as the historical background of the system; outlines the main actors and their decision-making powers; and describes the level of patient empowerment in the areas of information, choice, rights, complaints procedures, public participation and cross-border health care.
3. Financing: provides information on the level of expenditure and the distribution of health spending across different service areas, sources of revenue, how resources are pooled and allocated, who is covered, what benefits are covered, the extent of user charges and other out-of-pocket payments, voluntary health insurance and how providers are paid.
4. Physical and human resources: deals with the planning and distribution of capital stock and investments, infrastructure and medical equipment; the context in which IT systems operate; and human resource input into the health system, including information on workforce trends, professional mobility, training and career paths.
5. Provision of services: concentrates on the organisation and delivery of services and patient flows, addressing public health, primary care, secondary and tertiary care, day care, emergency care, pharmaceutical care, rehabilitation, long-term care, services for informal carers, palliative care, mental health care, dental care, complementary and alternative medicine, and health services for specific populations.
6. Principal health reforms: reviews reforms, policies and organisational changes; and provides an overview of future developments.

7. Assessment of the health system: provides an assessment based on the stated objectives of the health system, financial protection and equity in financing; user experience and equity of access to health care; health outcomes, health service outcomes and quality of care; health system efficiency; and transparency and accountability.
8. Conclusions: identifies key findings, highlights the lessons learned from health system changes; and summarises remaining challenges and future prospects.
9. Appendices: includes references, useful web sites and legislation.

The quality of HiTs is of real importance because they inform policy-making and meta-analysis. HiTs are the subject of wide consultation throughout the writing and editing process, which involves multiple iterations. They are then subject to the following.

- A rigorous review process.
- There are further efforts to ensure quality while the report is finalized that focus on copy-editing and proofreading.
- HiTs are disseminated (hard copies, electronic publication, translations and launches). The editor supports the authors throughout the production process and in close consultation with the authors ensures that all stages of the process are taken forward as effectively as possible.

One of the authors is also a member of the Observatory staff team and they are responsible for supporting the other authors throughout the writing and production process. They consult closely with each other to ensure that all stages of the process are as effective as possible and that HiTs meet the series standard and can support both national decision-making and comparisons across countries.

Appendix I – Country profiles

AUSTRALIA

AUSTRIA

CANADA

DENMARK

FINLAND

FRANCE

NETHERLANDS (KINGDOM OF THE)

NEW ZEALAND

REPUBLIC OF KOREA

SLOVENIA

SWEDEN

UNITED KINGDOM

UNITED STATES

AUSTRALIA

Population	25 687 041 (2020)	1
Administrative structure of country	Federal parliamentary democracy under a constitutional monarchy	2
Jurisdiction of health policy	Three levels of government are collectively responsible for providing universal healthcare: 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 (current prices, current PPPs)	<ul style="list-style-type: none"> • 9.4% of GDP (2019) • 4 919.20 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	<p>Healthcare financing in Australia is a combination of public government (at federal and state levels) funding and private funding.</p> <p>Public healthcare insurance: Medicare, a 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 levels of government.</p> <p>Private health insurance (PHI): There are two types of PHI in Australia: (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 non-emergency services. PHI 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>Out-of-pocket (OOP) payments: OOP payments are increasing in Australia for services and medicines that do not fall under the Medicare Benefits Schedule (MBS) or the Pharmaceutical Benefits Schedule (PBS).</p>	3
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>PHI: Around 46% of the population has some form of private-patient hospital cover; about 55% of the population has a form of general treatment cover.</p>	5, 6
Organization of health service delivery		
Primary care – gatekeeper function?	Yes; a referral is needed from a general practitioner (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

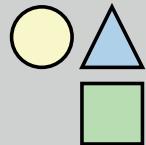
Payment mechanisms	<p>Hospital care: There are both public and private hospitals in Australia. Public hospital payment comes through global budgets from states and case-based payment (including physician costs). States finance budgets from a combination of tax revenues. Private hospitals are paid via fee-for-service (FFS) 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 and work in joint practices. They are primarily paid FFS, with fees set by the federal health minister through MBS; there is some use of quality-of-care activity incentive payments.</p> <p>Specialist care: Specialists are paid on FFS 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>	3
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
Health data strategy	National Digital Health Strategy Data Governance Framework 2021	7
Index of technical infrastructure	N/A	—
Relevant actors	<p>At the Commonwealth level:</p> <p>Department of Health: Develops and delivers policies and programmes 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 the Australian population, including Medicare and the PBS.</p> <p>Australian Institute for Health and Welfare (AIHW): the 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:</p> <ul style="list-style-type: none"> • Population Health Research Network • Centre for Victorian Data Linkages (VDL) • SA-NT Datalink • Data Linkage Queensland (DLQ) • Tasmanian Data Linkage Unit (TDLU) • Western Australia Data Linkage 	7

Sources:

- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SPOP.TOTL?view=chart> (accessed 4 January 2023).
- [2] Australia – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/australia/#government> (accessed 4 January 2023).
- [3] Commonwealth Fund (2020). International Health System Profile – Australia. Available at: <https://www.commonwealthfund.org/international-health-policy-center/countries/australia> (accessed 4 January 2023).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] Department of Health (2019). The Australian health system. Australian Government. Available at: <https://www.health.gov.au/about-us/the-australian-health-system> (accessed 4 January 2023).
- [6] Australian Institute of Health and Welfare (2018). Australia's health 2018. Australia's health series no.16. AUS 221. Canberra. Available at: <https://www.aihw.gov.au/getmedia/63fe0895-b306-4375-95ff-162149ffc34b/aihw-aus-221-chapter-2-1.pdf.aspx> (accessed 4 January 2023).
- [7] Office of the Australian Information Commissioner (2019). Health and medical research. Australian Government. Available at: <https://www.oaic.gov.au/privacy/the-privacy-act/health-and-medical-research/> (accessed 4 January 2023).

My Health Record

Electronic patient file introduced to improve care coordination and promote patient empowerment; secondarily, to enable research in the public interest with the aim of improving care.



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.



Is the data structured?

No



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 data linkable?

Data is linkable through a patient identifier

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 the data?

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



LEGEND

Sociodemographic/economic data

Healthcare data

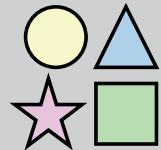
Health data

Environmental data

Population Health Research Network (PHRN)

National network of regional linking regional linking agencies that aggregates and prepares data from around Australia for research.

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.



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 data linkable?

Yes, data is linked through patient identifiers



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 the data?

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



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

LEGEND

● Sociodemographic/economic data
■ Healthcare data

▲ Health data
★ Environmental data

Medicare Benefit Schedule (MBS)

MBS data collection contains information on those health professional services subsidized by the Australian Government.

The MBS data collection 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.



Is the data structured?

Yes



At what level is the data aggregated?

Data is aggregated at national level but varies slightly by jurisdiction

Data collection

Who collects the data?

The Australian Government collects the data

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



Where is the data located?

The data is located with the Australian Government

Data storage

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



What is the process for acquiring access?

Application to the Department of Human Services' External Request Evaluation Committee (EREC) and AIHW Ethics

Research use

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

LEGEND

- Sociodemographic/economic data
- Healthcare data

- Health data
- Environmental data

Pharmaceutical Benefit Schedule (PBS)

PBS data collection maintains bills on all billed prescriptions paid for by the Australian Government.

The PBS data system (PBDS) comprises current and historical record of all prescriptions that are: (1) subsidised by the Government under the PBS and (2) priced below the patient co-payment and not subsidised (co-payment prescriptions).



Data collection

Is the data structured?

Yes



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

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 Committee for prospective studies

How long does it take for data to become available?

Data \leq 5 years old is easily accessible. Data \geq 5 years is archived and the retrieval process more lengthy and costly. The latest available data dates back to December 2018



How is the data used for research purposes?

Information is used for essential analysis of the PBS: modelling and monitoring of the budget process, monitoring of Community Pharmacy Agreements and of risk sharing arrangements

LEGEND

- Sociodemographic/economic data
- Healthcare data

- Health data
- Environmental data

AUSTRIA

Population	8 917 205 (2020)	1
Administrative structure of country	Federal parliamentary republic	2
Jurisdiction of health policy	Responsibilities around healthcare and policy are shared between the federal and state levels and some have been delegated to self-governing bodies. The federal level is primarily responsible for regulating social insurance and healthcare provision, except hospital care in which the federal level gives basic guidance and the states define the specifics of legislation and implementation.	3
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 11.5% of GDP (2020) • 5 899.10 US\$ PPP 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	16.9% (2020)	4
Type of health insurance and main sources of financing	<p>Healthcare financing in Austria is primarily financed through a mix of income-related statutory health insurance (SHI) contributions, general taxation and private sources.</p> <p>SHI: Contributions are wage-based; for the majority of the contribution population, this amounts to around 7.65% of income, shared nearly equally between employer and employee. Contributions for unemployed and pensioners are financed through general tax revenues; federal states cover contributions for those on social welfare benefits.</p> <p>Voluntary health insurance (VHI): VHI is offered by the SHI system and is mainly supplementary to provide (1) cover for extra amenities in the hospital sector and (2) cover for more choice of providers and medicines in ambulatory care. Premiums are risk-based.</p> <p>Out-of-pocket (OOP) payments: OOP costs, including direct payments, user charges (cost-sharing) and informal payments, have increased in the last ten years in terms of absolute spending.</p>	3, 5, 6
Extent of coverage by insurance type	<p>SHI: Through the mandatory SHI, there is near universal health coverage, with 99.9% of the population covered.</p> <p>VHI: More than 33% of the population is covered by some form of VHI.</p> <p>Uninsured: a limited number of people are not insured due to employment situations; they have the opportunity to register for VHI.</p>	3, 5, 6
Organization of health service delivery		
Primary care – gatekeeper function?	No	3, 5, 6
Specialist outpatient care	Specialist outpatient care is provided by office-based doctors, outpatient clinics and in hospitals.	3, 5, 6

Payment mechanisms	<p>Hospital care – inpatient: Payment for acute services is regulated by the national <i>Leistungsorientierte Krankenanstaltenfinanzierung</i> (LKF framework) in two ways: (1) DRG-based payment system (Austrian DRG) and (2) state-specific steering, which includes a fixed budget. Public hospitals also sometimes receive a per diem fee for patients with VHI. VHI also covers fee-for-service (FFS) 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 are financed by 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 general practitioners (GPs) generate revenue by FFS and capitation. Non-contracted GPs are paid FFS and may set their own fees.</p> <p>Specialist care: Ambulatory services are provided by GPs, specialist physicians, dentists and outpatient clinics, etc. Payment mechanisms include FFS and capitation, informed by whether services are included in the SHI catalogue of reimbursable services or what kind of contract the provider has with SHI fund(s). If no requirement is met, direct payment is required, which can be reimbursed up to 80%. PHI may cover the difference.</p> <p>Other: Long-term care institutions are financed by states and local authorities.</p>	3, 5, 6, 7
Health data protection laws	<ul style="list-style-type: none"> • EU Regulation 2016/679 (General Data Protection Regulation) • The Data Protection Act • The E-Government Act • The Health Telematics Act, 2012 • The Hospitals and Sanatoriums, Act • The Patient Charter • The Data Protection Amendment Act 	8, 9, 10, 11, 12, 13
Health data strategy	<ul style="list-style-type: none"> • The Research Organization Act • The Federal Statistics Act, 2000 • The Health Telematics Act, 2012 • The Privacy Amendment Act • Austria Health Targets 	14, 15, 16
Index of technical infrastructure	N/A	—

Relevant actors	<p>Federal Ministry for Social Affairs, Health, Care and Consumer Protection (Bundesministerium – Soziales, Gesundheit, Pflege und Jonsumentenschutz): Ministry responsible for health policy and the performance-based hospital funding system; operating the DIAG data warehouse on this basis.</p> <p>Electronic health record (Elektronische Gesundheitsakte, ELGA) GmbH: Public limited company owned by the Federal Ministry, the social insurance funds and the Länder in charge of developing the ELGA infrastructure.</p> <p>Health Austria (Gesundheit Österreich GmbH, (GÖG): Responsible body for the Austrian health information system (<i>Österreichisches Gesundheitsinformationssystem; ÖGIS</i>), the regional health information system (<i>Regionale Gesundheitsinformationssystem; REGIS</i>), both making use of the Ministry-held DIAG health data warehouse (<i>Dokumentations- und Informationssystem für Analysen im Gesundheitswesen; DIAG</i>).</p> <p>Social insurance authorities: Responsible for submitting collected data from healthcare providers to the umbrella association of the social insurance authorities after recoding it.</p> <p>Umbrella association of the social insurance funds (Dachverband der Sozialversicherungsträger; DVSV): Data received in DIAG are pseudonymized by the pseudonymization body (<i>Pseudonymisierungsstelle</i>) within the DVSV. Geographic and structural data are added to the individual datasets. The main social insurance federation submits the outpatient performance data to the Ministry of Labour, Social Affairs, Health and Consumer Protection.</p> <p>State (Länder) and regional health funds and regional health funds: Responsible for the inpatient sector, submitting inpatient diagnosis and performance data to the Ministry for billing and documentation purposes.</p> <p>DIAG: Diagnosis and performance data warehouse and information system at the Ministry for Social Affairs, Health, Care and Consumer Protection. Data are provided from hospitals for which federal states are in charge.</p> <p>Statistics Austria (Statistik Austria): Provides additional data (e.g., basic demographic data) from various sub-sources; hosts the national cancer statistics registry and the hospital discharge statistics; makes microdata from its sources available via the Austrian Micro Data Centre.</p>	3, 5, 6, 12
------------------------	---	-------------------

Sources:

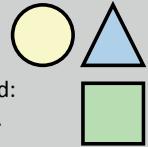
- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart> (accessed 4 January 2023).
- [2] Austria – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/austria/#government> (accessed 4 January 2023).
- [3] Bachner F, Bobek J, Habimana K, Ladurner J, Lepuschütz L, Ostermann H, Rainer L, Schmidt AE, Zuba M, Quentin W, Winkelmann J (2018). Austria health system review. Health systems in transition, 20(3). Available at: <https://eurohealthobservatory.who.int/publications/i/austria-health-system-review-2018> (accessed 4 January 2023).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] 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. Available at: <http://dx.doi.org/10.1787/9789264283268-en> (accessed 4 January 2023).
- [6] Gesundheit.gv.at (2019). Das Gesundheitswesen im Überblick [Healthcare at a Glance]. Available at: <https://www.gesundheit.gv.at/gesundheitsleistungen/gesundheitswesen/gesundheitssystem> (accessed 4 January 2023).

- [7] Kobel C, Pfeiffer K-P (2011). Austria: Inpatient care and the LKF framework. In Busse R, Geissler A, Quentin W, Wiley MM (eds). *Diagnosis-related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*. Available at: http://eurodrg.projects.tu-berlin.de/publications/DRGbook/Ch11_Kobel.pdf (accessed 4 January 2023).
- [8] European Commission (2018). EU data protection rules. Available at: https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en (accessed 4 January 2023).
- [9] ELGA (n.d.). Gesetzliche Grundlagen von ELGA [The legal basis of ELGA]. Available at: <https://www.elga.gv.at/faq/gesetzliche-grundlagen-von-elga> (accessed 4 January 2023).
- [10] Gesundheit.gv.at (2018). Patientencharta & Rechtsgrundlagen [Patient charter and legal bases]. Available at: <https://www.gesundheit.gv.at/gesundheitsleistungen/patientenrechte/patientencharta> (accessed 4 January 2023).
- [11] Bundesministerium für Digitalisierung und Wirtschaftsstandort (2019). Bundesgesetz über die Dokumentation im Gesundheitswesen. Rechtsinformationssystem des Bundes [Federal Law on Documentation in Healthcare. Federal legal information system]. Available at: <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10011011> (accessed 4 January 2023).
- [12] Bundesministerium für Gesundheit (2013). Wirkungsorientierte Folgenabschätzung. Verordnung des Bundesministeriums für Gesundheit zur ambulanten Dokumentation. Rechtsinformationssystem des Bundes [Outcome-oriented impact assessment. Ordinance of the Federal Ministry of Health on outpatient documentation. Federal legal information system]. Available at: https://www.ris.bka.gv.at/Dokumente/Begut/BEGUT_COO_2026_100_2_886724/COO_2026_100_2_892059.pdf (accessed 4 January 2023).
- [13] European Commission (2018). Regulation (EU) 2016/679 (General Data Protection Regulation; GDPR), OJ L 119, 04.05.2016; cor. OJ L 127, 23.5.2018. Available at: <https://gdpr-info.eu/> (accessed 4 January 2023).
- [14] Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz (2019). Austrian health targets. Available at: <https://gesundheitsziele-oesterreich.at/english-summary/> (accessed 4 January 2023).
- [15] Gesundheit Österreich (n.d.) Daten & Register [Data and register]. Available at: https://goeg.at/index.php/Daten_und_Register (accessed 4 January 2023).
- [16] Fülpö G (2017). Patientenorientierte Bedarfsplanung in Österreich [Patient-oriented demand planning in Austria]. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen, 125:60–9. Available at: <https://doi.org/10.1016/j.zefq.2017.07.003> (accessed 4 January 2023).

Electronic Health Records (ELGA)

ELGA (Elektronische Gesundheitsakte) supports medical treatment through improved availability of information.

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).



Is the data structured?
Yes

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



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?
There is no authority that checks the data for its accuracy and quality

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



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 the 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



What is the process of 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

LEGEND

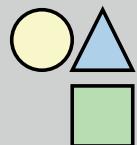
○ Sociodemographic/economic data
■ Healthcare data

△ Health data
★ Environmental data

Austrian health information system (ÖGIS)

ÖGIS (Österreiches Gesundheitsinformationssystem) is a geographic information system used for health reporting and planning.

ÖGIS stores individual level health data such as 1) diagnosis and performance documentation 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.



Is the data structured?
Yes

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



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)



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 care 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 the data?

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



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

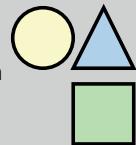
LEGEND

● Sociodemographic/economic data
■ Healthcare data

▲ Health data
★ Environmental data

Regional health information system (REGIS)

REGIS (Regionales Gesundheitsinformationssystem) is part of ÖGIS and provides regional health-related comparisons according to selected indicators.



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.



Is the data structured?
Yes

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



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?
Data is linked to bPK (unique healthcare identifier)



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 care data and 10 years for ambulatory data

Who owns the data?
Gesundheit Österreich GmbH (GÖG) as 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 the data?
Anybody can access REGIS



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

Research use

LEGEND

○ Sociodemographic/economic data
■ Healthcare data

△ Health data
☆ Environmental data

CANADA

Population	38 005 240 (2020)	1
Administrative structure of country	Federal parliamentary democracy (Parliament of Canada) under a constitutional monarchy, in which authorities and responsibilities are divided between the federal government and the provincial government and regulated by the constitution.	2, 3
Jurisdiction of health policy	Jurisdiction over healthcare governance and provision lies primarily with the individual provinces and territories.	2
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 10.8% of GDP (2019) • 5 370.40 US\$ PPP (2019) 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	14.9% (2019)	4
Type of health insurance and main sources of financing	<p>Healthcare financing in Canada is a mostly a combination of taxation and private sources.</p> <p>Public insurance programme: Canada has a regionally-administered universal public insurance programme, called 'Medicare'. Provinces/territories (P/T) are principally responsible for financing largely private provision and the federal government transfers funding. The main source of financing is general taxation at all levels of government, 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>Voluntary, private health insurance (PHI): PHI is chiefly employment-based; employees are usually 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>Out-of-pocket (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>	2, 4, 5, 6, 7, 8
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>	2, 7
Organization of health service delivery		
Primary care – gatekeeper function?	Yes; many provinces pay lower fees to specialists for non-referred consultations.	7
Specialist outpatient care	Hospitals provide the majority of specialist care; private, non-hospital facilities are also involved in specialist outpatient care provision.	2

Payment mechanisms	<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 (FFS).</p> <p>Primary care: Most physicians are self-employed and paid FFS across Canada's P/T; 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 with physician associations.</p> <p>Specialist care: Specialists are mostly financed through FFS, though there is some use of global budgets. Specialists have the same fee schedule as primary care physicians.</p> <p>Other: Mental healthcare 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>	3
Health data protection laws	<ul style="list-style-type: none"> The Personal Health Information Act, 2004 (Ontario). British Columbia's Freedom of Information and Protection of Privacy Act (FIPPA) 	7
Health data strategy	N/A	7
Index of technical infrastructure	Canada has a limited health information system, with medium coverage (60% of datasets cover 80% or more of the 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)	–
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 human health, well-being and development via access to individual-level, de-identified longitudinal data on BC's residents.</p> <p>Statistics Canada: The Canada federal government agency responsible for producing statistics on Canada'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 socioeconomic policy.</p> <p>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>	7

Sources:

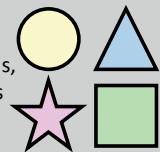
- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SPOP.TOTL?view=chart> (accessed 4 January 2023).
- [2] Marchildon GP, Allin S, Merkur S (2020). Canada: Health system review. Health systems in transition, 22(3):i–194.
- [3] Canada – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/canada/#government> (accessed 4 January 2023).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] WHO (2016). Current health expenditure as a share of GDP (%) and Current health expenditure per US \$ PPP. Global Health Expenditure Database. WHO, Geneva.
- [6] OECD (2011). OECD economic surveys: Canada. Paris, Organisation for Economic Co-operation and Development.
- [7] Allin S, Rudoler D (2017). The Canadian health care system. International health care system profiles. The Commonwealth Fund. Available at: <https://international.commonwealthfund.org/countries/canada/> (accessed 5 January 2023).
- [8] Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/_media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 5 January 2023).
- [9] ICES (2019). ICES Data. Available at: <https://www.ices.on.ca/Data-and-Privacy/ICES-data> (accessed 5 January 2023).
- [10] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Available at: <https://doi.org/10.1787/9789264244566-en> (accessed 5 January 2023).
- [11] Ishiguro L, Saskin R, Vermeulen MJ, Yates E, Gunraj N, Victor JC (2016). Increasing access to health administrative data with ICES Data & Analytic Services. *Healthcare Quarterly*, 19(1).
- [12] Population Data BC. (2019). About PopData. Available at: <https://www.popdata.bc.ca/about> (accessed 5 January 2023).
- [13] ICES (2019). Mission, vision & values. Available at: <https://www.ices.on.ca/About-ICES/Mission-vision-and-values> (accessed 5 January 2023).
- [14] Statistics Canada (2017). Social Data Linkage Environment. Available at: <https://www.statcan.gc.ca/eng/sdle/index> (accessed 5 January 2023).

Social Data Linkage Environment (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.

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.



Is the data structured?

Yes

Who collects the data?

The DRD database includes data 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



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 the 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 data source files and cannot see personal identifiers



What is the process of acquiring access?

Researchers can apply to access already linked records through Statistics Canada's 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 data used for research?

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

LEGEND

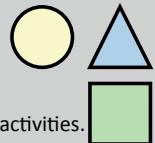
Sociodemographic/economic data
 Healthcare data

Health data
 Environmental 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.

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 is 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



Research use

What is the process of 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 for data 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 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.

LEGEND

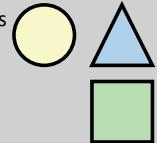
- Sociodemographic/economic data
- Healthcare data

- Health data
- Environmental data

Population Data British Columbia (PopData BC)

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

PopData manages 21 datasets from federal and provincial sources. External researchers also have the option of linking their own datasets to PopData. 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.



Is the data structured?

Yes



At what level is the data aggregated?

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

Data linkage and cleaning

Who collects the data?

Data is collected from federal and provincial sources

Is the data checked or cleaned?

Not for projects, but PopDataBC 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



Where is the data located?

A PopData Secure Research Environment (SRE), a central server accessible through an encrypted VPN firewall

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 the data?

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



What is the process of acquiring access?

Acquiring access follows a three-step process:

1. Submit data access and request form to PopData'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 with the SRE

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

Research use

LEGEND

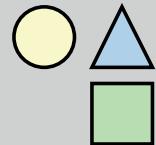
Sociodemographic/economic data
 Healthcare data

Health data
 Environmental data

Manitoba Population Data Research Data Repository

(a project at the Manitoba Centre for Health Policy 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 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.



Is the data structured?

Yes

Who collects the data?

A variety of sources including federal and government agencies, local health authorities, and indigenous groups



At what level is the data aggregated?

Data may be aggregated when sharing results with un-accredited analysts and researchers

Data linkage and cleaning

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



Where is the data located?

The data is housed at MCHP facilities at the University of Manitoba

Data storage

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 is able to access the data through a collection of data sharing agreements

Who performs the data linkage?

N/A

Who has access to the data?

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



What is the process of 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

Research use

How long does it take for data 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 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

LEGEND

● Sociodemographic/economic data
■ Healthcare data

▲ Health data
★ Environmental data

DENMARK

Population	5 831 400 (2020)	1
Administrative structure of country	Parliamentary constitutional monarchy	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, 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 healthcare services and coordination of specialized social care.	3
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 10.0% of GDP (2019) • 5 477.60 US\$ PPP (2019) 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	14.2% (2019)	4
Type of health insurance and main sources of financing	<p>Healthcare financing is mostly a combination of public financing and private sources.</p> <p>National healthcare system: the compulsory, national healthcare system in Denmark 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 healthcare 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' fulfilment 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/PHI): 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 from the statutory package and OOP payments; and (2) supplemental insurance to gain access to private providers.</p> <p>Out-of-pocket (OOP) payments: There is no cost-sharing for hospital and primary services, but there is for dental care, outpatient prescriptions and glasses. There is no cap on OOP costs, except for drug OOP spending.</p>	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>	5, 8
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.	5, 7

Specialist outpatient care	After referral from a GP, office-based specialists and hospital-based ambulatory clinics provide specialist outpatient care.	5, 6, 7
Payment mechanisms	<p>Hospital care: Hospitals in Denmark 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 (FFS), with some capitation. Regions are responsible for payment, though rates are set nationally in negotiations with doctors' associations.</p> <p>Specialist care: Services by private specialists are paid by the regions via FFS for referred patients. These fees are set in negotiation with regions. Private specialists and hospitals can also get OOP payments or funding through PHI. Specialists can set their own fees for patients without a referral. Specific agreements with the regions determine fees for services rendered at private institutions for patients referred from the public sector. Specialist outpatient care delivered in hospitals is paid via the general payment mechanisms for hospital care (see above).</p>	5, 7, 9
Health data protection laws	<ul style="list-style-type: none"> EU Regulation 2016/679 (General Data Protection Regulation) The Danish Data Protection Act, 2018 	10
Health data strategy	With Denmark's latest national "Digital Health Strategy 2018–2022", formulated by the Danish Ministry of Health, the Danish Ministry of Finance, the Danish Regions and Local Government Denmark, the country attempted to create a coherent and trustworthy health network for all. The strategy's focus is on digitization and use of health data regarding care, prevention, and treatment. It addresses five focus areas like trustworthy and secure data, and the patient as active partner, and includes 27 initiatives.	11, 12
Index of technical infrastructure	Denmark 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 (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 (Sundhedsstyrelsen, DHA): Responsible 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 and specific interventions among hospitals.</p> <p>Regional authorities and municipalities: The five regions are responsible for general practice, specialists, physiotherapists, dentists, chiropractors and pharmacists represented by their respective organizations. They collect data on healthcare services as well as demographic and economic data.</p> <p>Statistics Denmark: A state institution under the Ministry of the Interior and Housing – the central provider of statistics in Denmark analysing social and economic conditions.</p> <p>The Danish Clinical Quality Programm (Regionernes kliniske kvalitetsudviklingsprogramme, RKKP): Responsible for ensuring continuous improvement in the utilization of registers for clinical, managerial and research purposes. 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 healthcare. 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, 5, 8, 9

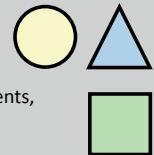
Sources:

- [1] World Bank (2021). Total population. Available at: https://data.worldbank.org/indicator/SPPOP_TOTL?view=chart (accessed 4 January 2023).
- [2] Denmark – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/denmark/#government> (accessed 6 January 2023).
- [3] Vrangbæk K (2020). The Danish health care system. International profiles of health care systems. The Commonwealth Fund. Available at: <https://www.commonwealthfund.org/international-health-policy-center/countries/denmark> (accessed 6 January 2023).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 6 January 2023).
- [5] Olejaz M, Nielsen AJ, Rudkjøbing A, Birk HO, Krasnik A, Hernández-Quevedo C (2012). Denmark health system review. Health systems in transition, 14(2).
- [6] Oresunddirekt (2019). Danish income tax – a short introduction. Available at: <https://www.oresunddirekt.se/en/working-in-denmark/taxes/danish-income-tax-a-short-introduction> (accessed 6 January 2023).
- [7] Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/_media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 6 January 2023).
- [8] WHO (2019). Global Health Expenditure Database, National Accounts Indicators. World Health Organization. Available at: <http://apps.who.int/nha/database/ViewData/Indicators/en> (accessed 6 January 2023).
- [9] Ministry of Health (2017). Healthcare in Denmark. Available at: <https://sum.dk/english>.
- [10] Federspiel G (2018). Data protected – Denmark. Linklaters. Available at: <https://www.linklaters.com/de-de/insights/data-protected/data-protected--denmark> (accessed 6 January 2023).
- [11] Sundhed.dk (2016). Background. Available at: <https://www.sundhed.dk/borger/service/om-sundheddk/om-organisationen/ehealth-in-denmark/background/> (accessed 6 January 2023).
- [12] Healthcare Denmark (2018). A Coherent and Trustworthy Health Network for All: Digital Health Strategy 2018–2022. Available at: <https://sum.dk/publikationer/2018/januar/a-coherent-and-trustworthy-health-network-for-all> (accessed 12 March 2023).
- [13] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Available at: <https://doi.org/10.1787/9789264244566-en> (accessed 6 January 2023).

Sundhed.dk

National eHealth portal that facilitates data flow and communication between patients, health professionals and improves transparency.

Sundhed.dk includes individual level health data such as 1) electronic patient record including clinical data, medical history, treatments, medications, prescriptions and laboratory results, 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



At what level is the data aggregated?

Not done within sundhed.dk but some included sources are aggregated (e.g., registries)

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

Data linkage and cleaning

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

Who owns the data?

Data ownership 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 the 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 is the data used for research purposes?

Sundhed.dk is not used for research purposes



LEGEND

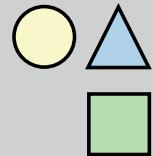
Sociodemographic/economic data
 Healthcare data

Health data
 Environmental data

Danish Clinical Registries (RKKP)

RKKP (Regionernes kliniskekvalitetsudviklingsprogram) manages a range of Danish clinical registries with different targets and populations.

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.



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 data linkable?

Yes, data is linkable to the Central Person Register



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 the 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



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 it 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



Research use

LEGEND

Sociodemographic/economic data

Healthcare data

Health data

Environmental data

FINLAND

Population	5 530 720 (2020)	1
Administrative structure of country	Parliamentary republic; unitary state organized on a decentralized basis, with central, regional and local levels of governance.	2, 3
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.	4
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 9.2% of GDP (2019) • 4 561.50 US\$ PPP (2019) 	5
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	17.4% (2019)	5
Type of health insurance and main sources of financing	<p>Healthcare financing in Finland 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/PHI): 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>Out-of-pocket (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>	4
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 have some type of VHI.</p> <p>Occupational healthcare: Employees are covered by occupational healthcare, the scope of which varies (about 33% of the total population).</p>	4
Organization of health service delivery		
Primary care – gatekeeper function?	Yes; patients can choose where to receive specialist services with a referral from a GP.	4

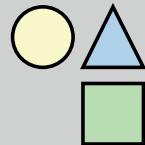
Specialist outpatient care	Hospital districts, comprising multiple municipalities and of which there are 20, are responsible for specialized care. Specialized care, both out- and inpatient, is provided by five different organization types (four are owned by hospital districts and one by one or more municipalities): (1) university hospitals (there are 5); (2) central hospitals; (3) university clinic departments; (4) psychiatric inpatient facilities; and (5) specialist-run health centre hospitals (there are 14). Private clinics can also offer services.	4
Payment mechanisms	<p>Hospital and specialist care: Hospital districts that provide inpatient and specialist care are mainly funded by their constituent municipalities. Funding mostly 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. Physicians in hospital districts are salaried employees, though many may work in private clinics on a fee-for-service (FFS) basis.</p> <p>Primary care: Municipalities prospectively fund budgets for health centres; costs are usually determined by previous years' budgets. In centres run by joint-municipal health authorities, budgets are developed similarly, but costs are determined by volume of services provided. GPs working in municipal centres have a variable salary scheme, including a monthly salary with some extra FFS for selective services. In centres where a "personal doctor system" has been introduced, payment is via basic salary, capitation, and FFS for visits.</p>	4
Health data protection laws	<ul style="list-style-type: none"> The Act on the Secondary Use of Health and Social Data, 2019 The Data Protection Act, 2018 	6, 7, 8
Health data strategy	<ul style="list-style-type: none"> The Health-Sector Growth Strategy for Research Innovation Activities and Information to Support Well-being and Service Renewal: eHealth and eSocial Strategy 2020 	8, 9, 10
Index of technical infrastructure	Finland 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).	11
Relevant actors	<p>Finnish Institute for Health and Welfare (Terveyden ja hyvinvoinnin laitos, THL): 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>Findata: The new data permit authority oversees data regulation under the Act on the Secondary Use of Health and Social Data. It is the single-point contact to access several different databases and safeguards secure and efficient procedures in utilizing the data in research, development and innovation. It operates in conjunction with the THL, but separately from the Institute's other activities.</p> <p>Kanta Services: Finland's eHealth 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 the Social Insurance Institution of (Kansaneläkelaitos, Kela).</p> <p>Statistics Finland: Large amounts of socioeconomic data are available through this institution.</p>	8, 9, 10, 12

Sources:

- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SPOP.TOTL?view=chart> (accessed 4 January 2023).
- [2] Finland – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/finland/#government> (accessed 7 January 2023).
- [3] Theodora (2019). Finland government 2019. Countries of the worlds. Available at: https://theodora.com/wfbcurrent/finland/_government.html (accessed 7 January 2023).
- [4] Keskimäki I, Tykkynen 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):1–166. Available at: <https://euroheatlhobservatory.who.int/publications/1/finland-health-system-review-2019> (accessed 4 January 2023).
- [5] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreId=9> (accessed 4 January 2023).
- [6] Seppänen J (2019). Findata – Secondary use of Finnish social and health data – a new act and Data permit authority. [Presentation] Available at: <https://www.sitra.fi/app/uploads/2019/05/a-finnish-model-for-the-secure-and-effective-use-of-data.pdf> (accessed 12 March 2023).
- [7] Dammart 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. Available at: <https://blogs.dlapiper.com/privacymatters/finland-parliament-approves-new-act-on-the-secondary-use-of-social-and-health-care-personal-data/> (accessed 7 January 2023).
- [8] Ministry of Employment and the Economy (2016). Innovating together – Health sector growth strategy for research and innovation activities. Roadmap for 2016–2018. Available at: https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/75145/MEE_guidelines_8_2016_Health_sector_growth_strategy_17062016_web.pdf (accessed 7 January 2023).
- [9] Ministry of Social Affairs and Health (2015). Information to support well-being and service renewal – eHealth and eSocial strategy. Available at: http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/74459/URN_ISBN_978-952-00-3575-4.pdf?sequence=1&isAllowed=y (accessed 7 January 2023).
- [10] Malkamäki S, Hämäläinen H (2019). New legislation will speed up the use of Finnish health data. Available at: <https://www.sitra.fi/en/blogs/new-legislation-will-speed-use-finnish-health-data/> (accessed 7 January 2023).
- [11] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Available at: <https://doi.org/10.1787/9789264244566-en> (accessed 7 January 2023).
- [12] Finnish Institute for Health and Welfare (2019). Data permit authority Findata. Available at: <https://thl.fi/en/web/thlfi-en/statistics/data-and-services/data-permit-authority-findata> (accessed 7 January 2023).

Findata

The Findata authority grants permission for the use of data in research, and collates and links samples of existing datasets as requested for research purposes.



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



Is the data structured?
Yes (presumed)

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



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?
Data is linked using individual identification numbers, which are then removed.
Researchers have access to pseudonymized data.



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

Who has access to the data?

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



What is the process for acquiring access?
Free access to aggregate data will be subject to an information request. A data utilization plan is required for access to data sets. Users can be granted access to anonymized or pseudonymized 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.

LEGEND

○ Sociodemographic/economic data
■ Healthcare data

▲ Health data
★ Environmental data

FRANCE

Population	67 391 582 (2020)	1
Administrative structure of country	Semi-presidential republic, with a ceremonial president. France is divided into 18 regions as well as five overseas collectivities, one overseas territory, one special collectivity and an uninhabited island.	2
Jurisdiction of health policy	Jurisdiction of health policy and regulation of the healthcare system is divided among the state (parliament, government), SHI and local authorities, particularly at the regional level.	2
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 11.1% of GDP (2019) • 5 274.30 US\$ PPP (2019) 	3
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	9.3% (2019)	3
Type of health insurance and main sources of financing	<p>Healthcare financing in France is a combination of public (government and statutory health insurance, SHI) funding and private funding.</p> <p>SHI: from social security contributions paid by employers and employees, and earmarked income taxes, including the "general social contribution" (CSG). Additional revenue comes from taxes on tobacco and alcohol, on pharmaceutical companies and state subsidies.</p> <p>Voluntary health insurance (VHI): complementary insurance exists to mainly cover co-payments and for better coverage of medical goods and services poorly covered by SHI, e.g., dental and optical care. Supplementary insurance provides access to private amenities not in the benefits basket. VHI is financed through VHI premiums that are not income-related and are not related to age or risk.</p> <p>Out-of-pocket (OOP) payments: OOP expenditure is the lowest in the EU and mainly applies to co-payments, bill balancing, optical and dental care, pharmaceuticals and LTC.</p>	5
Extent of coverage by insurance type	<p>SHI: provides universal, compulsory coverage through three main types of compulsory SHI, incorporated into a single national exchange. The National Sickness Insurance Fund (<i>Caisse Nationale d'Assurance Maladie des Travailleurs Salariés</i>, CNAMTS), covers 92% of the population; the agricultural fund (<i>Mutualité Sociale Agricole</i>) covers 7%; other smaller schemes cover the remainder.</p> <p>VHI: Around 90% of the population has complementary health insurance (as of 2010).</p>	5, 6
Organization of health service delivery		
Primary care – gatekeeper function?	There is a voluntary gatekeeping system for adults over 16 years that provides incentives for registering with and visiting a general practitioner (GP) prior to consulting a specialist.	5, 6
Specialist outpatient care	Delivered by self-employed doctors in their own private practices or by doctors on staff at hospitals and health centres.	5, 6

Payment mechanisms	<p>Hospital care: Public and private hospitals are paid via DRGs within the medical activity-based payment system (T2A), according to homogeneous hospital stay groups. Block grants are also used. In private for-profit hospitals, doctors' procedures and services are paid separately, fee-for-service (FFS).</p> <p>Primary care: Self-employed GPs are paid FFS; some are salaried at hospitals and health centres; some have mixed-income. Some use of pay-for-performance.</p> <p>Specialist care: Patients in ambulatory care are expected to pay healthcare providers themselves and then claim reimbursement. However, direct payment by a health insurance fund is becoming more common in ambulatory care.</p>	5, 6
Health data protection laws	<ul style="list-style-type: none"> • EU Regulation 2016/679 (General Data Protection Regulation) • New French Data Protection Act, 2018 • New Decree No. 2019-536 (the "Implementing Decree"), 2019 • National Health Data System (SNDS) security referential • Act n°2002-303 dated 4 March 2002 on the rights of the sick and the quality of the health system • Relevant articles within other regulations: Articles L. 1461-1 and L.1111-8 of the Public Health Code 2017 	7, 8, 9
Health data strategy	<ul style="list-style-type: none"> • Law No. 2019-774 of 24 July 2019 on "the organisation and transformation of the healthcare system" • Order of 29 November 2019 approving an amendment to the constitutive agreement of the public interest group "National Institute of Health Data" creating the public interest group "Health Data Platform" • National artificial intelligence plan "AI for Humanity" • Road map: Accelerating the Digital Shift • Digital health roadmap 	10, 11, 12, 13, 14
Index of technical infrastructure	N/A	—

Relevant actors	Ministry of Solidarity and Health (Ministère de la Santé et de la Prévention) : responsible for defining the national health strategy, and sets and implements government policy for public health, and the organization and financing of the healthcare system. The Ministry is represented in the regions by the Regional Health Agencies . National Health Insurance Fund for Salaried Workers (Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, CNAMTS) : the general SHI scheme, CNAMTS, covers employees in commerce and industry and their families and individuals eligible for UHC basic coverage. Agricultural fund : covers farmers and agricultural employees and their families. National Institute of Health and Medical Research (Institut national de la santé et de la recherche médical, Inserm) : a public scientific and technological institute operating under the Ministries of Health and of Research dedicated to biomedical research and human health and international collaboration. National Commission for Data Protection and Liberties (Commission nationale de l'informatique et des libertés, CNIL) : an independent administrative authority responsible for ensuring the protection of personal data in computer files and processing operations. French National Health Data System (Système national des données de santé français, SNDS) : co-managed by the CNAMTS and the Health Data Hub (HDH), SNDS links many databases to make data available to promote studies, research or evaluation of public interest to analyse and improve population health and health system efficiency and innovation. Ethical and Scientific Committee for Research, Studies and Evaluations in the Field of Health (Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé, CESREES) . National Agency for the Security of Information Systems (Agence nationale de la sécurité des systèmes d'information, ANSSI) .	5, 15 22
------------------------	--	----------------

Sources:

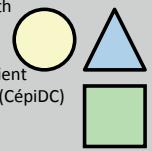
- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart> (accessed 4 January 2023).
- [2] France – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/france/#government> (accessed 7 January 2023).
- [3] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [4] Welcome to the English website of the French National Assembly – Assemblée nationale (assemblee-nationale.fr) (accessed 7 January 2023).
- [5] Chevreul K, Brigham KB, Durand-Zaleski I, Hernández-Quevedo C (2015). France: Health System Review. Health Systems in Transition, 17(3). Available at: https://www.euro.who.int/__data/assets/pdf_file/0011/297938/France-HiT.pdf (accessed 7 January 2023).
- [6] Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/__media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 6 January 2023).
- [7] European Commission (2018). Regulation (EU) 2016/679 (General Data Protection Regulation; GDPR), OJ L 119, 04.05.2016; cor. OJ L 127, 23.5.2018. Available at: <https://gdpr-info.eu/> (accessed 7 January 2023).

- [8] Agence Nationale de La Securite des Systemes d'Information (2011). Information systems defence and security, France's strategy. Available at: https://www.ssi.gouv.fr/uploads/IMG/pdf/2011-02-15_Information_system_defence_and_security_-_France_s_strategy.pdf (accessed 7 January 2023).
- [9] Systeme national des donnees de sante. Referentiel de securite, Guide d'accompagnement. Available at: SNDS Référentiel de sécurité Guide pédagogique du référentiel de sécurité (accessed 7 January 2023).
- [10] LOI n° 2019-774 du 24 juillet 2019 relative à l'organisation et à la transformation du système de santé (1). Available at: <https://legifrance.gouv.fr> (accessed 7 January 2023).
- [11] Republique Francaise (2019). Arrêté du 29 novembre 2019 portant approbation d'un avenant à la convention constitutive du groupement d'intérêt public « Institut national des données de santé » portant reation du groupement d'intérêt public « Plateforme des données de santé ». Available at: <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000039433105> (accessed 7 January 2023).
- [12] AI for Humanity (2018). Available at: AI for humanity.fr (accessed 7 January 2023).
- [13] Republique Francaise (2019). Dossier D'Information. Conference ministre –Paris, le jeudi 25 avril 2019. Available at: https://190425_dossier_presse_masante2022_ok.pdf (solidarites-sante.gouv.fr) (accessed 7 January 2023).
- [14] Ministère des Solidarités et de la Santé and Agence du Numerique en Santé (2020). Available at: <https://esante.gouv.fr/virage-numerique/feuille-de-route> (accessed 7 January 2023).
- [15] Ministère des Solidarités et de la Santé (n.d.). Available at: Ministère des Solidarités et de la Santé (solidarites-sante.gouv.fr) (accessed 7 January 2023).
- [16] Republique Francaise. Inserm (n.d.). Available at: <https://www.inserm.fr/en/home/> (accessed 7 January 2023).
- [17] Commission Nationale Informatique & Liberte (2019). The CNIL in a Nutshell 2021. Available at: <https://www.cnil.fr/sites/default/files/atoms/files/the-cnil-in-a-nutshell.pdf> (accessed 12 March 2023).
- [18] European Health Information Portal (n.d.). The National Healthcare Data System (SNDS). Available at: <https://www.healthinformationportal.eu/node/1217> (accessed 7 January 2023).
- [19] Health Data Hub (n.d.). Available at: <https://www.health-data-hub.fr/> (accessed 7 January 2023).
- [20] Health Data Hub (2021). Brochure. Available at: <https://www.health-data-hub.fr/sites/default/files/2021-04/Brochure%20HDH-compressed.pdf> (accessed 7 January 2023).
- [21] Agence nationale de la sécurité des systèmes d'information (2022). Available at: <https://www.ssi.gouv.fr/> (accessed 7 January 2023).
- [22] Kaguelidou F, Sommet A, Lapeyre-Mestre M (2018). Use of French healthcare insurance databases in pediatric pharmacoepidemiology. *Therapie*, 73(2):127–33. doi:10.1016/j.therap.2017.11.008.

Health Data Hub (HDH)

Data access platform acting as unique gateway to health data for public interest purposes.

The HDH is authorized to make data available from the expanded SNDS (National Health Data System). With continuous data on about 99% of the French population, SNDS contains data on demographics, medical information, outpatient reimbursements, hospitalizations, medicines/devices, survey/interview data, registry data, biobank/sample/specimen data and customer record data. It links several databases: outpatient claims (SNIIRAM), public/private hospital discharge summaries (PMSI), death and medical causes of death (CépiDC) and disability (CNSA). The HDH has data from a catalogue of additional databases, incl. the emergency log (OSCOUR), information on victims of attacks/exceptional sanitary situations (SI-VIC) and SNDS Fast Track (COVID-19). Other databases are expected to enter HDH's catalogue, including cohorts and other studies.



Is the data structured or unstructured?

Both; upcoming unstructured data may include medical images and free text.

Who collects the data?

There are various data controllers. The HDH may only collect copies of existing databases, removing identifying information. A selection of databases may be replicated and updated regularly in The HDH's technology platform.



At what level is the data aggregated?

The HDH stores individual-level, pseudonymized data.



Is the data checked or cleaned?

Both; partly by data controllers, partly by the HDH.

Is the data linkable?

Yes; in SNDS individuals are represented by a pseudonymized NIR. Other databases from the HDH catalogue use an NIR or other unique identifier methods that can be linked to SNDS.



Where is the data located?

Data is stored on a Microsoft Azure Cloud, physically based in Paris region.

How long is the data stored?

The databases of the catalogues are stored on the HDH platform permanently. The non-catalogue databases are accessible only during the project time frame.

Who owns the data?

No one; data is a common good, according to GDPR.

Who performs the data linkage?

A third party (e.g., National Health Insurance Fund ((CNAM))) and then sent to the HDH.

Who has access to the data?

Any person/institution working in public interest that obtains authorization, e.g., Researchers, policymakers (national, European, international) and individuals.



What is the process for acquiring access?

Data access is granted to project coordinators following approval by Ethics and Scientific Committee (CESREES) and the National Commission for Data Protection and Liberties (CNIL). The process follows: (1) request made to the HDH; (2) request sent to CESREES; (3) CNIL authorization; (4) the HDH consolidates data and prepares secure project environment on its platform; (5) remote access given to process (not retrieve) data; and (6) results made public on the HDH website. Access is free of charge for public actors; private project coordinators must pay a fee.

How long does it take for data to become available?

The HDH sends requests to CESREES regularly; CESREES approval takes one month; CNIL review takes 2 months renewable.

What is latest available data? N/A



How is the data used for research purposes?

The HDH enables project coordinators to easily access pseudonymized data securely for research purposes to improve the quality of care and patient support. The HDH supports 27 pilot projects that access data through the HDH and from which data will be made available.

LEGEND

Sociodemographic/economic data
 Healthcare data

Health data
 Environmental data

NETHERLANDS (KINGDOM OF THE)

Population	17 441 140 (2020)	1
Administrative structure of country	Parliamentary constitutional monarchy	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
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 11.2% of GDP (2020) • 6 298.50 US\$ PPP (2020) 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	9.5% (2020)	4
Type of health insurance and main sources of financing	<p>Healthcare financing in the Kingdom of the Netherlands 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.</p> <p>Voluntary, private health insurance (VHI/PHI): Private health plans 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>Out-of-pocket (OOP) payments: Since 2016, every insured person must pay an annual deductible for healthcare costs. Besides this, OOP costs include cost-sharing of some services.</p>	5, 6
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 the population are insured.</p> <p>VHI: 84% of the population purchase some combination of complementary VHI.</p>	3
Organization of health service delivery		
Primary care – gatekeeper function?	Yes; referral by a general practitioner (GP) is required for hospital and specialist care.	3

Specialist outpatient care	Almost all specialists are hospital-based; there is a growing trend to work in multidisciplinary ambulatory centres. The Kingdom of the Netherlands has six types of institutions that provide specialist care: general hospitals; academic (university) hospitals; specialized hospitals; independent treatment centres (day care only); top clinical centres (both general hospital care and complex care); and trauma centres.	3
Payment mechanisms	<p>Hospital care: Most hospitals are private, non-profit and are paid through case-based per diem payments, negotiated with the insurer over price, quality and volume, using the case-based diagnosis-treatment combination system (DBCs). Fee-for-service (FFS) 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 paid through capitation fees per registered patient and consultation fees for GPs and ambulatory mental healthcare; 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 FFS or salaried. Specialist fees are freely negotiable as a part of hospital payment. Specialists in ambulatory centres are paid FFS; the fee schedule is insurer negotiated.</p> <p>Other: Long-term residential care is paid through care intensity packages. Of note, insurers in the Kingdom of the Netherlands are expected to engage in strategic purchasing, and contracted providers are expected to compete on both quality and cost.</p>	3, 7
Health data protection laws	<ul style="list-style-type: none"> EU Regulation 2016/679 (General Data Protection Regulation) Personal Data Protection Act (wet Bescherming persoonsgegevens) (2001) 	3, 8
Health data strategy	N/A	—
Index of technical infrastructure	The Kingdom of the Netherlands 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).	9
Relevant actors	<p>Statistics Netherlands: If government invests public funds in a dataset, Statistics Netherlands can access it. It conducts data linkage projects using its own data.</p> <p>Health Care TTP: This agency leads on other data linkage projects and provides access to de-identified data.</p>	3

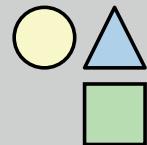
Sources:

- [1] World Bank (2021). Total population. Available at: https://data.worldbank.org/indicator/SPPOP_TOTL?view=chart (accessed 4 January 2023).
- [2] Netherlands – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/netherlands/#government> (accessed 7 January 2023).
- [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] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] OECD (2015). How does health spending in the Netherlands compare? OECD Health statistics 2015. Available at: <https://www.oecd.org/els/health-systems/Country-Note-NETHERLANDS-OECD-Health-Statistics-2015.pdf> (accessed 7 January 2023).
- [6] Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 7 January 2023).
- [7] Wammes J, Jeurissen P, Westert G, Tanke M (2017). The Dutch health care system. In Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 7 January 2023).
- [8] European Commission (2018). Regulation (EU) 2016/679 (General Data Protection Regulation; GDPR), OJ L 119, 04.05.2016; cor. OJ L 127, 23.5.2018. Available at: <https://gdpr-info.eu/> (accessed 7 January 2023).
- [9] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Available at: <https://doi.org/10.1787/9789264244566-en> (accessed 7 January 2023).

Nivel Primary Care Database

Primary care database operated by the Nivel research institute, bringing together routine data on health problems and the use of primary care services in order to monitor the health and healthcare of the population.

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.



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



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



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



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



LEGEND

Sociodemographic/economic data

Healthcare data

Health data

Environmental data

PHARMO Database Network

PHARMO is an independent research organization that provides rapid access to primary and secondary healthcare data for research purposes.

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?

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

How long does it take for data 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

LEGEND

- Sociodemographic/economic data
- Healthcare data

- Health data

- Environmental data

NEW ZEALAND

Population	5 084 300 (2020)	1
Administrative structure of country	Parliamentary democracy under a constitutional monarchy	2
Jurisdiction of health policy	The government with the Ministry of Health plays a central role in determining the healthcare policy agenda, service requirements and budget. The government dominates all aspects of healthcare as the primary funder and supplier of healthcare; it also sets regulations and monitors compliance. Te Whatu Ora/Health New Zealand is responsible for planning, purchasing and providing health services for the population, and disability support for those over age 65.	3
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 9.1% of GDP (2019) • 4 211.90 US\$ PPP (2019) 	4, 5
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	<p>Healthcare financing in New Zealand primarily comes from public, government funding, though PHI and OOP payments also play a role.</p> <p>National healthcare system: The national healthcare system in New Zealand is financed through general tax revenue.</p> <p>Private, voluntary health insurance (PHI): PHI 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>Out-of-pocket (OOP) payments: OOP costs include cost-sharing and other direct costs, such as co-payments for general practitioner (GP) visits and pharmaceuticals. Currently there are reduced fees after 12 doctor visits and for those aged under 14 years, and no drug co-payments after 20 prescriptions annually.</p>	5
Extent of coverage by insurance type	<p>Public healthcare system: New Zealand's public healthcare system provides universal healthcare for all permanent residents.</p> <p>PHI: About 33% of the population purchase some form of PHI.</p>	5
Organization of health service delivery		
Primary care – gatekeeper function?	Yes	5

Specialist outpatient care	Most specialists are employed by Te Whatu Ora – Health New Zealand and work in public hospitals; however, special care is also provided in private clinics and hospitals.	5
Payment mechanisms	<p>Hospital care: There is a mix of private and public hospitals in New Zealand. Public hospitals receive global budgets from Te Whatu Ora – Health New Zealand, which are 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 the budget within a public hospital.</p> <p>Primary care: General practitioners (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 co-payments for fee-for-service (FFS) 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 Te Whatu Ora – Health New Zealand are salaried for work in public hospitals; when specialists work privately or treat patients in private hospitals payment comes in the form of FFS, setting their own fees.</p>	5
Health data protection laws	<ul style="list-style-type: none"> • The Privacy Act, 1993 • The Health Information Privacy Code, 1994 • Health Information Governance Guidelines • Health Information Security Framework 	6
Health data strategy	• NZ Digital Health Strategy	7
Index of technical infrastructure	New Zealand 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>New Zealand implemented a substantial health system reform in July 2022. Data and health information sit within all of the entities below.</p> <p>NZ Ministry of Health: is the chief adviser to the government on health, policy and performance, and drives the strategy and monitoring of the health and disability system.</p> <p>Te Whatu Ora – Health New Zealand: established in 2022 to lead the day-to-day running of the health system across New Zealand, with functions delivered at local, district, regional and national levels. It is responsible for funding and commissioning health services and owns and operates public hospitals.</p> <p>Te Aka Whai Ora – Māori Health Authority: established in 2022 to lead and monitor the way the health system understands and responds to the health and well-being needs of Māori.</p> <p>Stats NZ Tatauranga Aotearoa: 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, a large research database.</p>	3, 10, 11

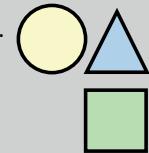
Sources:

- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SPOP.TOTL?view=chart> (accessed 4 January 2023).
- [2] New Zealand – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/new-zealand/#government> (accessed 7 January 2023).
- [3] The Commonwealth Fund (2020). International Health System Profile – New Zealand. Available at: <https://www.commonwealthfund.org/international-health-policy-center/countries/new-zealand> (accessed 7 January 2023).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] Gauld R (2017). The New Zealand health care system. In Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 7 January 2023).
- [6] Ministry of Health (2009). Data protection and privacy. New Zealand Government. Available at: <https://www.health.govt.nz/nz-health-statistics/access-and-use/data-protection-and-privacy> (accessed 7 January 2023).
- [7] Ministry of Health (2019). Digital health portfolio – February 2019. New Zealand Government. Available at: <https://www.health.govt.nz/system/files/documents/pages/digital-health-portfolio-report-feb-2019.pdf> (accessed 7 January 2023).
- [8] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Available at: <https://doi.org/10.1787/9789264244566-en> (accessed 7 January 2023).
- [9] Stats NZ Tatauranga Aotearoa (n.d.) About us. Available at: <https://www.stats.govt.nz/about-us> (accessed 7 January 2023).
- [10] Stats NZ Tatauranga Aotearoa (n.d.) Integrated Data Infrastructure. Available at: <https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/> (accessed 7 January 2023).
- [11] Te Whatu Ora – Health New Zealand (n.d.). About Us. Available at: <https://www.tewhatuora.govt.nz/about-us/who-we-are/our-work/> (accessed 7 January 2023).

Integrated Data Infrastructure (IDI)

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

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.



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



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 the 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



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

LEGEND

● Sociodemographic/economic data
■ Healthcare data

▲ Health data
★ Environmental data

REPUBLIC OF KOREA

Population	51 780 580 (2020)	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 (current prices, current PPPs)	<ul style="list-style-type: none"> • 8.4% of GDP (2020) • 3 493.70 US\$ PPP (2020) 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	29.2% (2020)	4
Type of health insurance and main sources of financing	<p>Healthcare in the Republic of Korea 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 Programme, 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>Out-of-pocket (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 the population).</p> <p>Medical Aid Programme: The Medical Aid Programme covers the remaining 3% of the population.</p> <p>VHI: About 75% of the population in the Republic of Korea purchase VHI.</p>	3
Organization of health service delivery		
Primary care – gatekeeper 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	Hospital care: Hospitals are not-for-profit; payment methods from NHIS insurance funds include fee-for-service (FFS) , 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 numbers and profits. Primary care: Healthcare facilities are paid FFS through health insurance funds; individual physicians are salaried. FFS is based on the Resource-based Relative Value (RBRVI) system ; fees are set by insurers. Specialist care: Healthcare providers are paid fee-for-service , based on fees set by insurers based on the RBRVI system; individual physicians are salaried.	3
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 • The Rare Disease Management Act, 2015 • Guidelines for the protection of biometric information, 2005 • Guidelines for Clinical Trial Management by the KGCP, 2005 • Guidelines for the Protection of Personal Information in Hospitals, 2012 • The Regulation on the Operation and the Management of the National Biobank of Korea, 2013 • The Cancer Control Act, 2003 • Guidelines for the Protection of Bio-information 	3
Health data strategy	<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, Centres for Disease Control and Prevention and hospitals 	—
Index of technical infrastructure	The Republic of Korea rates among the highest for health data availability, though less so for accessibility. Data are also not often used for research/monitoring purposes (OECD's Key national dataset availability, maturity and use score: 5.92/7.0).	5
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, ageing and child welfare.</p> <p>Ministry of Economy and Finance: This ministry oversees the financial policies of KOR, including the taxation system, management of public institutions, and planning and coordinating the mid- and long-term economic development goals of the country.</p> <p>NHIS and National Health Insurance Database (NHIIDB): NHIS includes data from all three types of health insurances in the Republic of Korea – NHIS, Medical Aid Programme and long-term care, as well as from several sub-databases.</p> <p>Health Insurance Review and Agency: The HIRA database contains medical claims data for the entire Korean population as a result of NHIS.</p>	3

Sources:

- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SPOP.TOTL?view=chart> (accessed 4 January 2023).
- [2] Republic of Korea – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/korea-south/#government> (accessed 7 January 2023).
- [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).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Available at: <https://doi.org/10.1787/9789264244566-en> (accessed 7 January 2023).

National Health Insurance Service (NHIS) and National Health Insurance Database (NHIDB)

A database which gathers beneficiary health and healthcare data.

NHIS and the NHIDB include data from all three types of health insurances in the Republic of Korea: NHIS, Medical Aid Programme, and LTIC. It includes data from several sub-databases, such as 1) patient sociodemographics, insurance, and health check-up info; 2) diagnosis, 3) encounters, and 4) medical utilization, including prescriptions.



Is the data structured?

Yes



At what level is the data aggregated?

At the regional and national level

Data linkage and cleaning

Is the data checked or cleaned?

N/A

Is data linkable?

Yes, using a unique ID on the NHID platform



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

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 the data?

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



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

LEGEND

- Sociodemographic/economic data
- Healthcare data

- Health data
- Environmental data

Health Insurance Review Agency (HIRA)

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

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	 At what level is the data aggregated? At the regional and national level
	 Who collects the data? HIRA	 Is the data checked or cleaned? N/A
Data storage	 Where is the data located? In a data warehouse (DW)	 Who owns the data? HIRA and NHIS Who performs the data linkage? The NHIS
	 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 has access to the 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 a big data open system	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
	How long does it take to become available? N/A	

LEGEND

 Sociodemographic/economic data
 Healthcare data

 Health data
 Environmental data

SLOVENIA

Population	2 100 126 (2020)	1
Administrative structure of country	Parliamentary republic, with ceremonial president	2, 3
Jurisdiction of health policy	Health policy and the management of the health system occur centrally at the national level through the Parliament and executive government, including the Ministry of Health. Compulsory social health insurance (SHI) is also administered centrally at the national level.	3
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 10.1% of GDP (2020) • 3 783.20 US\$ PPP (2020) 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	10.6% (2020, provisional)	4
Type of health insurance and main sources of financing	<p>Healthcare financing in Slovenia is a combination of public and private funding.</p> <p>SHI: contributions are employment-based and levied on gross income (13.45%: 6.36% from employees; 6.89% from employers). Pensioner contributions are paid by the Pension and Disability Insurance; National Institute for Employment covers contributions for unemployed; state/municipal taxes cover individuals without income, prisoners, and war veterans.</p> <p>Government: Health funds come from general taxation (national and local), non-tax revenue, capital revenue and grants.</p> <p>VHI: provides complementary insurance to cover co-insurance for services in benefits package and is paid via a flat rate premium.</p> <p>Out-of-pocket (OOP) payments: Slovenia has one of the lowest rates of OOP expenditure in the European Union. These payments go to co-insurance and direct payments for services not included in the benefits package.</p>	3, 5
Extent of coverage by insurance type	<p>SHI: provides compulsory coverage to more than 99% of all permanent residents. There are two main groups of insured: employees (and dependents) and the other employed (e.g., unemployed, others without fixed income, pensioners, farmers and self-employed).</p> <p>VHI: Around 73% of the population have complementary health insurance.</p>	3, 5
Organization of health service delivery		
Primary care – gatekeeper function?	Yes; primary care is provided mainly by a network of community-based primary healthcare centres owned and managed by municipalities. A range of public health and primary care services are delivered by a range of providers. Office-based physicians in private practice may also deliver publicly funded primary care services.	3, 5
Specialist outpatient care	Three types of settings: (1) hospitals (university, clinical or general), where most care is provided; (2) bigger CPHCs; and (3) individual or group practices of private specialists.	3

Payment mechanisms	A capped national annual budget for health leads to capped payments for providers contracted by ZZZS (see below). Hospital care: Acute care is paid by a case payment model based on DRGs . Other payment mechanisms for inpatient services include prospectively determined number of bed-days and fee-for-service (FFS) . Primary care: primary care services provided by personal physicians in health centres are financed through capitation and FFS . Flat-based payments have been used to incentivize chronic disease prevention. Specialist care: care provided in hospitals is paid FFS ; services in health centres are remunerated by a mix of capitation and FFS ; those in private practice are FFS .	3
Health data protection laws	<ul style="list-style-type: none"> • EU Regulation 2016/679 (General Data Protection Regulation) • The Data Protection Act (Zakon o varstvu osebnih podatkov, ZVOP-1) (O.G. 94/07 and 177/20) • Healthcare Databases Act (Zakon o zbirkah podatkov s področja zdravstvenega varstva, ZZPPZ, 65/00, 47/15, 31/18, 152/20 – ZZUOOP, 175/20 – ZIUOPDVE, 203/20 – ZIUOPDVE in 112/21 – ZNUPZ) • Patients' Rights Act, 2008 • Many sectoral laws define data filing systems for public and private sectors, e.g., around data storage 	6, 7, 8, 9
Health data strategy	<ul style="list-style-type: none"> • The Health Databases Act, 2000 (amended in 2018, 2020 and 2021) (Zakon o zbirkah podatkov s področja zdravstvenega varstva, Official Gazette of RS, no. 65/00, 47/15 and 31/18) • Resolution on the National Health Care Plan, 2016–2025 	3, 8, 9
Index of technical infrastructure	N/A	—
Relevant actors	<p>Ministry of Health (MoH): responsible for governance and leadership of the healthcare system. It determines health and healthcare policy and supervises big capital investments and medicines and medical devices.</p> <p>Health Insurance Institute of Slovenia (Zavod za zdravstveno zavarovanje Slovenije, ZZZS): single payer; administers the centralized compulsory SHI, collects healthcare contributions, contracts with providers, monitors health expenditures and negotiates prices of services. Along with NIJZ, ZZZS is a principal controller and processor of large health data repositories.</p> <p>National Institute of Public Health (Nacionalni inštitut za javno zdravje, NIJZ): responsible for essential public health functions and provides support for health system governance. Along with ZZZS, NIJZ is a principal controller and processor of large health data repositories and maintains patient and service registries. NIJZ is also an authorized producer of national statistics.</p> <p>Statistical Office of the Republic of Slovenia</p> <p>Information Commissioner of the Republic of Slovenia: responsible for data protection.</p>	3

Sources:

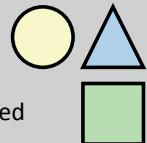
- [1] World Bank (2021). Total population. Available at: https://data.worldbank.org/indicator/SPPOP_TOTL?view=chart (accessed 4 January 2023).
- [2] Slovenia – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/slovenia/> (accessed 7 January 2023).
- [3] Albreht T, Polin K, Brinovec RP, Kuhar M, Poldrugovac M, Rehberger PO, Rupel VP, Vrakoc P (2021). Slovenia: Health system review. Health Systems in Transition, 23(1):i–188. Available at: <https://eurohealthobservatory.who.int/publications/i/slovenia-health-system-review-2021> (accessed 10 March 2022).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] OECD/European Observatory on Health Systems and Policies (2021). Slovenia: Country Health Profile 2021, State of Health in the EU, OECD Publishing, Paris/European Observatory on Health Systems and Policies, Brussels.
- [6] European Commission (2018). Regulation (EU) 2016/679 (General Data Protection Regulation; GDPR), OJ L 119, 04.05.2016; cor. OJ L 127, 23.5.2018. Available at: <https://gdpr-info.eu/> (accessed 7 January 2023).
- [7] Linklaters (2020). Data Privacy. Data Protected – Slovenia. Available at: <https://www.linklaters.com/en/insights/data-protected/data-protected--slovenia> (accessed 7 January 2023).
- [8] ICLG.com (2021). Data Protection Laws and Regulations Slovenia 2021. Available at: <https://iclg.com/> (accessed 7 January 2023).
- [9] Government of Slovenia (2016). Pravno-informacijski sistem (PIS). *Resolucija o nacionalnem planu zdravstvenega varstva 2016–2025 » Skupaj za družbo zdravja « (ReNPZV16–25)* (Uradni list RS, št. 25/16). Available at: <http://pisrs.si/Pis.web/preglejPredpisa?id=RESO102> (accessed 7 January 2023).

Health "All in One" Portal (zVEM Portal)

The zVEM Portal (Vdravje Vse Na Enem Mestu) is for users of the Slovenian healthcare system serving as a repository for all the latest healthcare data from the Central Registry of Patient Data (CRPD – Slovenian EHR), eReferrals, ePrescriptions, among other databases.

The zVEM Portal offers insight into the following data: health and healthcare utilization, e.g., exam reports, discharge letters, lab reports, patient summaries, COVID-19 test results, demographic data, etc.

E-Prescription and dispensations are also included as well as records of issued and used eReferrals for secondary appointments (eAppointments).



Data collection

Is the data structured?

Both structured and unstructured (in pdf form).



Data linkage
and cleaning

At what level is the data aggregated?

Data is collected for distinct patients and aggregated nationally.

Who collects the data?

NIPH in the capacity of eHealth authority. zVEM Portal does not collect any data; it provides access to accumulated data from other sources/services.

Is the data checked or cleaned?

No active checking/cleaning is performed. Certain validation is implemented at various points.

Is the data linkable?

Yes, via several identifiers, but subject to legal constraints and decisions made case by case.



Data storage

Where are the data located?

Data located in the two national eHealth data centres (in Ljubljana and Maribor).

How long are the data stored?

Records remain active at least 5 years; can be archived after that.

Who owns the data?

NIPH is the manager of eHealth databases.

Who performs the data linkage?

NIPH and institutional healthcare providers.



Who has access to the data?

Patients and healthcare providers have access to original records. Others have access to specific data in accordance with their legal basis (e.g., MoH, public health authority).



Research use

What is the process for acquiring access?

Access provided upon obtaining appropriate digital identity through the SI-PASS system and review from the ethics board or data commissioner.

How long does it take for data to become available?

Immediately after being entered.

What is latest available data?

Data is transmitted continuously after services are provided.



How is the data used for research purposes?

Researchers can submit their requests for data to the NIPH and then the appropriate procedures are carried out in accordance with the law. Results must be shared before dissemination.

LEGEND

Sociodemographic/economic data

Health data

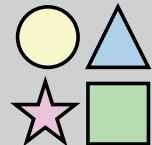
Healthcare data

Environmental data

Health in the Municipality (HiM)

HiM (Zdravje v občini) presents a variety of health indicators at municipal level.

HiM provides selected health and social indicators that describe socio-economic and demographic data, health data, incl. on risk factors and prevention measures, and some environmental data based on modelling.



Is the data structured?
Yes

Who collects the data?
Data is collected by seven institutions*

Data collection



At what level is the data aggregated? Indicators are available at municipal level and level of NUTS3 region.

Is the data checked or cleaned?
Both; checked and cleaned.

Is the data linkable?
Initial data for the HiM portal are linkable by unique identifiers.

Data linkage
and cleaning



Where is the data located?
Data is located at NIPH.

Data storage

How long is the data stored?
Indicators are stored permanently.

Who owns the data?

NIPH is controller of the data after receiving it from 6 other institutions.

Who performs the data linkage?

Data linkage is performed by each institution as necessary.



Who has access to the data?
This is open data.



What is the process for acquiring access?
It is an open data portal for anyone to access.

How long does it take for data to become available?

The portal is refreshed annually.

What is latest available data?

2021, but reference periods differ and the majority of data covers 2015-2020.



How is the data used for research purposes?

For analytical, decision-making and research to improve healthcare, design new services and understand differences among regions.

Research use

* Faculty of Sports; Institute of Oncology; Institute for Social Welfare; Road Safety Agency; Statistical Office; Ministry of Finance; NIJZ.

LEGEND

Sociodemographic/economic data
 Healthcare data

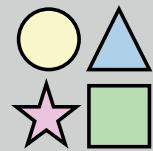
Health data
 Environmental data

Health Data Portal

The Health Data Portal (Podatkovni portal) is the main portal for health reporting and health system planning.

All health related databases in Slovenia are presented in the portal.

Data is available on health status, socio-economics, health service utilization, resources, prescribed medicines, and expenditures as well as links to HiM indicators and the environmental agency and cancer registry.



 Data collection	Is the data structured? Yes	At what level is the data aggregated? Data may be aggregated at different levels, e.g., age-gender groups, territory.
	Who collects the data? Most data is collected by NIPH, but links to data in eight other institutions provided.*	Is the data checked or cleaned? Data are checked and cleaned.
 Data storage	Where is the data located? Data located at NIPH.	Who owns the data? NIPH is controller of the data.
	How long is the data stored? Indicators are stored permanently.	Who performs the data linkage? Data linkage is performed by NIPH.
 Research use	What is the process for acquiring access? This is an open data portal, which anyone can access.	Who has access to the data? This is open data.
	How long does it take for data to become available? Portal is updated approx. 6 times/year.	
	What is latest available data? 2020	 How is the data used for research purposes? Data used for analytical, decision-making and research purposes at different territorial levels and other aggregated levels.

* Statistical Office; Institute of Macroeconomic Analysis and Development; Environment Agency; Ministry of the Environment and Spatial Planning; Institute of Oncology; HIIS; Radiation Protection Administration; Faculty of Sport, Traffic Safety Agency; Association of Social Institutions of Slovenia.

LEGEND

 Sociodemographic/economic data
 Healthcare data

 Health data
 Environmental data

SWEDEN

Population	10 323 440 (2020)	1
Administrative structure of country	Parliamentary constitutional monarchy	2
Jurisdiction of health policy	The healthcare system is organized into three levels: national, regional and local. The Health and Medical Services Act of 1982 specifies that the responsibility for ensuring access to good healthcare is that of the county councils/regions and municipalities.	3
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 11.4% of GDP (2020) • 5 753.60 US\$ PPP (2020) 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	13.7% (2020)	4
Type of health insurance and main sources of financing	<p>Healthcare financing in Sweden 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>Private health insurance (PHI): PHI is mainly supplementary coverage from employers for enhanced access to specialists and elective treatments.</p> <p>Out-of-pocket (OOP) payments: OOP costs are contained to a degree through caps on cost-sharing for health services (US\$123) and drugs (US\$246); most OOP spending goes to pharmaceuticals. County councils set co-payment 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>PHI: Around 10% of all employed individuals aged 15–74 have private insurance.</p>	3
Organization of health service delivery		
Primary care – gatekeeper function?	No	3, 5
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 Sweden 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: Sweden'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 for compensation.</p>	3

Health data protection laws	<ul style="list-style-type: none"> EU Regulation 2016/679 (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
Health data strategy	National eHealth Strategy 2010	8
Index of technical infrastructure	Sweden has very high overall data availability, sharing and accessibility, but reports less use of 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	National Board of Health and Welfare: Maintains health data registers and official statistics.	3

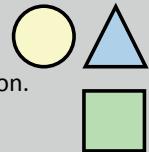
Sources:

- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart> (accessed 4 January 2023).
- [2] Sweden – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/sweden/#government> (accessed 7 January 2023).
- [3] Anell A, Glenngård AH, Merkur S (2012). Sweden health system review. Health systems in transition, 14(5).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/_media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 7 January 2023).
- [6] European Commission (2018). Regulation (EU) 2016/679 (General Data Protection Regulation; GDPR), OJ L 119, 04.05.2016; cor. OJ L 127, 23.5.2018. Available at: <https://gdpr-info.eu/> (accessed 7 January 2023).
- [7] 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. Available at: https://ec.europa.eu/health/sites/health/files/ehealth/docs/laws_sweden_en.pdf (accessed 7 January 2023).
- [8] International Society for Telemedicine & eHealth (2010). Sweden National eHealth Strategy 2010. Available at: https://www.isfteh.org/media/sweden_national_ehealth_strategy_2010 (accessed 7 January 2023).
- [9] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Available at: <https://doi.org/10.1787/9789264244566-en> (accessed 7 January 2023).

Registerforskning

Provides researchers with information on existing registers. The RUT (Register Utiliser Tool) is being developed to facilitate searching and matching of information among registers.

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; but RUT metadata is structured	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
	Who collects the data? Government agencies or other organizations, using various methods	Is the data checked or cleaned? N/A
Data storage	Where is the data located? Different register holders are responsible for different registers	Is the data linkable? Yes, unique personal identity numbers allow for data to be linkable at the source
	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? N/A 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

LEGEND

Sociodemographic/economic data
 Healthcare data

Health data
 Environmental data

UNITED KINGDOM

Population	67 215 290 (2020)	1
Administrative structure of country	Parliamentary constitutional monarchy	2
Jurisdiction of health policy	The UK is mainly a devolved nation, 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 UK government directly.	3
Current health expenditure (CHE) and per capita health expenditure (current prices, current PPPs)	<ul style="list-style-type: none"> • 12.8% of GDP (2020) • 5 267.70 US\$ PPP (2020) 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	13.8% (2020)	4
Type of health insurance and main sources of financing	<p>Healthcare financing in the UK is primarily derived from government funding, with some private sources.</p> <p>National Health Service (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. The NHS also receives income from co-payments, e.g., from drug prescription payments.</p> <p>Voluntary, private health insurance (VHI): There is private insurance in the UK, varying from coverage for specific conditions to broader packages. It is primarily employer-based private medical insurance, but individuals also purchase plans.</p> <p>Out-of-pocket (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>	5, 6, 7
Extent of coverage by insurance type	<p>NHS: Coverage is universal through the NHS for all legal residents.</p> <p>VHI: Approximately 11% of the population have VHI for quicker access to services.</p>	4, 5, 6
Organization of health service delivery		
Primary care – gatekeeper function?	Yes	5
Specialist outpatient care	Most outpatient specialist consultations are performed in hospitals; some take place in general practices.	5

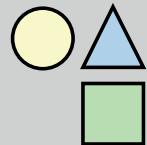
Payment mechanisms	<p>Hospital care: Hospitals in the UK are mostly public and payment is made mostly through case-based payments. Budgets are used to cover mental health, education, and research and training; these also include physician and drug costs.</p> <p>Primary care: Primary care is delivered primarily via private general practitioners (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 (FFS) 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 FFS basis.</p>	3, 5, 6
Health data protection laws	<ul style="list-style-type: none"> • The Data Protection Act, 2018 • Data Protection, Privacy and Electronic Communication • The Health Service (Control of Patient Information) Regulations, 2002 	7, 8, 9
Health data strategy	<ul style="list-style-type: none"> • Data saves lives: reshaping health and social care with data, 2022 • Better, broader, safer: using health data for research and analysis, 2022 • 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 	10, 11, 12, 13, 14, 15, 16
Index of technical infrastructure	<p>The UK overall has good health data availability and accessibility. 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>	16
Relevant actors	<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 England Transformation Directorate: The directorate has fully incorporated NHS X and will incorporate NHS Digital in coming months to drive the digital transformation of the NHS and social care.</p>	17, 18

Sources:

- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SPOP.TOTL?view=chart> (accessed 4 January 2023).
- [2] United Kingdom – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/united-kingdom/#government> (accessed 7 January 2023).
- [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).
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] Thorlby R, Arora S (2017). The English health care system. International health care system profiles. The Commonwealth Fund. Available at: <https://international.commonwealthfund.org/countries/england/> (accessed 7 January 2023).
- [6] Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/_media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 7 January 2023).
- [7] legislation.gov.uk (n.d.). The Health Service (Control of Patient Information) Regulations 2002. Available at: <http://www.legislation.gov.uk/uksi/2002/1438/contents/made> (accessed 7 January 2023).
- [8] NHS Digital (n.d.). Codes of practice for handling information in health and care. Available at: <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> (accessed 7 January 2023).
- [9] UK Government (October 2020). Data Protection, Privacy and Electronic Communication. Available at: <https://uk-gdpr.org/> (accessed 7 January 2023).
- [10] Department of Health and Social Care (2022). Data saves lives: reshaping health and social care with data. Available at: <https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data/data-saves-lives-reshaping-health-and-social-care-with-data#ministerial-foreword> (accessed 7 January 2023).
- [11] Department of Health and Social Care (2022). Independent report. Better, broader, safer: using health data for research and analysis. Available at: <https://www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis> (accessed 7 January 2023).
- [12] Department of Health and Social Care (2018). The future of healthcare: our vision for digital, data and technology in health and care. Available at: <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> (accessed 7 January 2023).
- [13] NHS Digital (2018). NHS Digital data and information strategy. Available at: <https://digital.nhs.uk/data-and-information/nhs-digital-data-and-information-strategy> (accessed 7 January 2023).
- [14] Scottish Government (2015). A Health and Biomedical Informatics Research Strategy for Scotland. Available at: <https://www.gov.scot/publications/health-biomedical-informatics-research-strategy-scotland/> (accessed 7 January 2023).
- [15] Scottish Government (2012). Joined up data for better decisions: A strategy for improving data access and analysis. Available at: <https://www.gov.scot/publications/joined-up-data-better-decisions-strategy-improving-data-access-analysis/> (accessed 7 January 2023).
- [16] Scottish Government (2012). Joined-up data for better decisions: Guiding principles for data linkage. Available at: <https://www.gov.scot/publications/joined-up-data-better-decisions-guiding-principles-data-linkage/> (accessed 7 January 2023).
- [17] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. Available at: <https://doi.org/10.1787/9789264244566-en> (accessed 7 January 2023).
- [18] UK Government (n.d.). NHS England. Transformation Directorate. Available at: <https://transform.england.nhs.uk/about-us/> (accessed 7 January 2023).

Clinical Practice Research Datalink (CPRD)

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



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.



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



At what level is the data aggregated?

N/A

Data linkage and cleaning

Is the data checked or cleaned?

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

Is data linkable?

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



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 the 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



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 for data 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 to use

How is 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

LEGEND

Sociodemographic/economic data

Health data

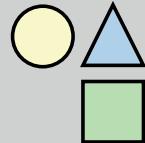
Healthcare data

Environmental data

Electronic Data Research and Innovation Service (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.

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 	<p>Is the data structured? Yes</p> <p>Who collects the data? Different actors collect the data depending on the dataset</p>	<p>At what level is the data aggregated? N/A; varies by dataset</p> <p>Is the data checked or cleaned? N/A; varies by dataset</p> <p>Is data linkable? Yes, data is linkable through the Community Health Index</p>
	<p>Where is the data located? ISD holds most source datasets centrally</p> <p>How long is the data stored? Data is stored as per study protocol, then kept in the eDRIS archive for the specified time, after which it is destroyed</p>	<p>Who owns the data? Shared between the NHS Boards and the ISD</p> <p>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.</p> <p>Who has access to the data? Approved applicants/researchers have to undergo information governance training and have the necessary approvals</p>
Data storage 	<p>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</p> <p>How long does it take for data to become available? ISD holds information centrally that is updated on a monthly basis and provides NHS Boards with local extracts</p>	<p>How is the data used for research purposes? N/A</p>
Research use 		

LEGEND

- Sociodemographic/economic data
- Healthcare data

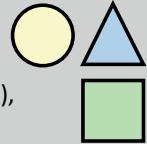
- Health data
- Environmental data

Scottish Primary Care Information Resource (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.

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; varies by dataset

Is 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 the 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 for data to become available?

An NSS analyst will contract the requestor for next step after receipt of the form, normally within two weeks

How is the data used for research purposes?

N/A



LEGEND

Sociodemographic/economic data
 Healthcare data

Health data
 Environmental data

UNITED STATES OF AMERICA (US)

Population	329 484 12	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 (current prices, current PPPs)	<ul style="list-style-type: none"> • 16.8% of GDP (2019) • 10 948.50 US\$ PPP (2019) 	4
Share of spending paid out-of-pocket (OOP) at point of use (% CHE)	11.3% (2019)	4
Type of health insurance and main sources of financing	<p>The US 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 US: There are also additional smaller public insurers for the military, veterans 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).</p> <p>Supplementary insurance for Medicare financed by individuals also exists.</p> <p>Out-of-pocket (OOP) payments: Cost-sharing varies from plan to plan. Co-payments are required for the most part for physician visits, hospital services and prescription drugs. There is a cap on OOP spending at US\$6600 annually for individuals with private insurance.</p>	5, 6
Extent of coverage by insurance type	<p>Public insurers: In total, public insurers in the US cover 37.7% of the total insured population, with Medicare covering 17.2% and Medicaid 19.3% of the population.</p> <p>Private insurers: 67.2% of the total insured population are covered by private insurances. Of this number, 56% come from employer-based insurance and 16% from direct-purchase coverage.</p> <p>Uninsured: 8.8% of the population of the US are uninsured.</p>	6, 7
Organization of health service delivery		
Primary care – gatekeeper function?	No, though some insurance programmes require general practitioners (GPs) to play a gatekeeping role. Private plans use narrow networks of providers, however, which limit patient choice.	5, 6
Specialist outpatient care	Specialist outpatient care is provided in both hospitals and private practices.	5

Payment mechanisms	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 (FFS) basis. Primary care: Primary care is delivered by private GPs and other physicians in the US. It is paid primarily via FFS . Capitation is used in some private plans. There is 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. Specialist care: Specialists in both hospitals and private practices are mostly paid FFS .	5, 6
Health data protection laws	The Health Insurance Portability and Accountability Act (HIPAA)	8
Health data strategy	<ul style="list-style-type: none"> The Health Data Initiative The Health Information Technology for Economic and Clinical Health (HITECH) Act, 2009 The Promoting Interoperability proposed rule 	8
Index of technical infrastructure	The US 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).	9
Relevant actors	US Department of Health and Human Services (DHHS) and associated agencies (e.g., Centers for Medicare and Medicaid Services, Centers for Disease Control and Prevention, National Institutes of Health, US Food and Drug Administration). Providers and commercial insurers also have a strong influence within the US's health system.	5

Sources:

- [1] World Bank (2021). Total population. Available at: <https://data.worldbank.org/indicator/SPOP.TOTL?view=chart> (accessed 4 January 2023).
- [2] United States – The World Factbook (2021). Available at: <https://www.cia.gov/the-world-factbook/countries/united-states/#government> (accessed 7 January 2023).
- [3] Rice T, Rosenau P, Unruh LY, Barnes AJ, van Ginneken E (2020). United States of America: Health system review. *Health Systems in Transition*, 22(4):i–441.
- [4] OECD Statistics (2021). Health Expenditure and Financing. Available at: <https://stats.oecd.org/Index.aspx?ThemeTreeId=9> (accessed 4 January 2023).
- [5] The Commonwealth Fund (2017). The U.S. health care system. International health care system profiles. Available at: https://international.commonwealthfund.org/countries/united_states/ (accessed 7 January 2023).
- [6] Mossialos E, Djordjevic A, Osborn R, Sarnak R (eds) (2017). International Profiles of Health Care Systems. The Commonwealth Fund. Available at: https://www.commonwealthfund.org/sites/default/files/documents/media_files_publications_fund_report_2017_may_mossialos_intl_profiles_v5.pdf (accessed 7 January 2023).
- [7] Berchick ER, Hood E, Barnett JC (2018). Health insurance coverage in the United States: 2017. United States Census Bureau, report no. 60–264. Available at: <https://www.census.gov/library/publications/2018/demo/p60-264.html> (accessed 7 January 2023).
- [8] US Department of Health & Human Services (n.d.). Laws & regulations. Available at: <https://www.hhs.gov/regulations/index.html> (accessed 7 January 2023).
- [9] OECD (2015). Health data governance: Privacy, monitoring and research. OECD Health policy studies, OECD Publishing, Paris. <https://doi.org/10.1787/9789264244566-en> (accessed 7 January 2023).

CancerLinQ

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

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 the 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

LEGEND

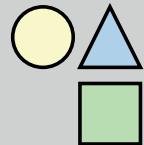
Sociodemographic/economic data
 Healthcare data

Health data
 Environmental data

Centers for Medicare and Medicaid Services (CMS Data)

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

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.



Is the data structured?
Varies; yes and no

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



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)



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

How long is the data stored?
Varies based on database

**Data linkage
and cleaning**



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



What is the process for acquiring access?
Many databases are available to anyone, on websites such as data.cms.gov. For identifiable data, CMS has a contract with the Research Data Assistance Center (ResDAC, hosted by the University of Minnesota) to facilitate data access requests. Researchers submit application, which is reviewed by CMS privacy board. After approval, researcher pays for data and CMS prepares it

How long does it take to become available?
Around 3-5 months



How is it 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

LEGEND

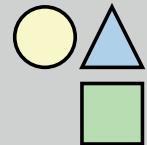
- Sociodemographic/economic data
- Healthcare data

- △ Health data
- ★ Environmental data

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.

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	At what level is the data aggregated? 18 instances of Epic and 8 regions aggregate to 1 fully integrated system
	Who collects the data? Kaiser Permanente staff and members	Is the data checked or cleaned? Yes, some quality checks include profiling tools such as SAS Dataflux
Data storage	Where is the data located? Kaiser Permanente data warehouses	Is the data linkable? Yes, data is linkable using a patient identifier
	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 for verifying insurance coverage, can request information

LEGEND

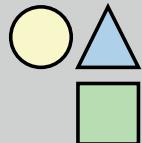
○ Sociodemographic/economic data
■ Healthcare data

△ Health data
★ Environmental data

National Patient-centered Clinical research Network (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 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.



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



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 provider

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



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, and sent back to the researcher, allowing them to ask the same question to all at the same time

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



How is it 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

LEGEND

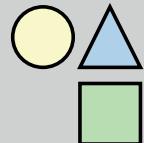
○ Sociodemographic/economic data
■ Healthcare data

▲ Health data
★ Environmental data

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.

The SEER (Surveillance, Epidemiology, and End Results) dataset has data from 18 registries across the US for persons with cancer and Medicare data contains enrolment 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 enrolment 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 request for restricted 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 it 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

LEGEND

○ Sociodemographic/economic data
■ Healthcare data

△ Health data
☆ Environmental data

Appendix II – Examples of privacy and data protection policies related to secondary use of health data in study countries

COUNTRY	PRIVACY AND DATA PROTECTION REGULATIONS AND POLICIES PERTAINING TO HEALTH INFORMATION	HEALTH DATA GOVERNANCE, LEGISLATION, STRATEGIES, AND GUIDELINES
Australia	<p>Commonwealth level:</p> <ul style="list-style-type: none">• The Privacy Act, 1988• The Privacy Regulation, 2013• The Privacy Amendment Act, 2017• My Health Records Act, 2012• Healthcare Identifiers Act, 2010• Health Insurance Act, 1973• National Health Act, 1953 <p>State/Territory level:</p> <ul style="list-style-type: none">• 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)	<ul style="list-style-type: none">✓ National Digital Health Strategy, 2020✓ Data Governance Framework, 2021
Austria	<p>National level:</p> <ul style="list-style-type: none">• The Data Protection Act• The Health Telematics Act, 2012• The Electronic Health Record (ELGA) Directive• The Hospitals and Sanatoriums Act• The Patient Charter• The Research Organization Act• The Federal Statistics Act, 2000• The Data Protection Amendment Act• Federal Act on Documentation in Healthcare <p>Supra-national level:</p> <ul style="list-style-type: none">• EU Regulation 2016/679 (General Data Protection Regulation)	<ul style="list-style-type: none">✓ Austria Health Targets since 2011

Canada	<p>Federal level:</p> <ul style="list-style-type: none"> • Privacy Act, 1985 • Personal Information Protection and Electronic Documents Act (PIPEDA), 2000 <p>Provincial level (selected examples):</p> <ul style="list-style-type: none"> • Health Information Act (Alberta) • Freedom of Information and Protection of Privacy Act (FIPPA) (British Columbia) • EHealth 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) • The Personal Health Information Act, 2004 (Ontario) • Personal Health Information Act (Ontario) • Freedom of Information and Protection of Privacy Act (Prince Edward Island) • Health Information Act (Prince Edward Island) • The Health Insurance Act (Quebec) • Health Information Protection Act (Saskatchewan) • Health Information Privacy and Management Act (Yukon) 	
Denmark	<p>National level:</p> <ul style="list-style-type: none"> • Danish Data Protection Act, 2018 • Health Act, 2005 <p>Supra-national level:</p> <ul style="list-style-type: none"> • EU Regulation 2016/679 (General Data Protection Regulation) 	✓ Digital Health Strategy, 2018–2022
Finland	<p>National level:</p> <ul style="list-style-type: none"> • The Data Protection Act, 2018 • The Act on the Secondary Use of Health and Social Data, 2019 • Act on Patient's Status and Rights • Decree on Patient Records <p>Supra-national level:</p> <ul style="list-style-type: none"> • EU Regulation 2016/679 (General Data Protection Regulation) 	✓ The Health-Sector Growth Strategy for Research ✓ eHealth and eSocial Strategy, 2020

France	<p>National level:</p> <ul style="list-style-type: none"> Act No. 78-17 on Information Technology, Data Files and Civil Liberties dated 6 January 1978, as amended by Act No. 2018-493 dated 20 June 2018 on Personal Data Protection Decree No. 2005-1309 of 20 October 2005, modified by Decree No. 2018-687 of 1 August 2018 (DPA Decree) Public Health Code (Articles L.1110-4 et seq, L.1111-8 et seq, L.1115-1 et seq, L.1122-1 et seq, L.1435-6, L.1460-1 et seq, R.1111-1 et seq) (on the processing of health data) Act of 28 January 2016 on the modernization of the healthcare system The 2019 Act on the Organization and Transformation of the Health System <p>Supra-national level:</p> <ul style="list-style-type: none"> EU Regulation 2016/679 (General Data Protection Regulation) 	<ul style="list-style-type: none"> ✓ Act on the Modernization of the Health Care System ✓ 2019 Act on the Organization and Transformation of the Health System
Netherlands (Kingdom of the)	<p>National level:</p> <ul style="list-style-type: none"> Personal Data Protection Act, 2000 Medical Treatment Agreement Act Implementation Act of EUGDPR <p>Supra-national level:</p> <ul style="list-style-type: none"> EU Regulation 2016/679 (General Data Protection Regulation) 	—
New Zealand	<p>National level:</p> <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) 	<ul style="list-style-type: none"> ✓ NZ Digital Health Strategic Framework ✓ Ngā Tikanga Paihere: a framework guiding ethical and culturally appropriate data use
Republic of Korea	<ul style="list-style-type: none"> The Bioethics and Safety Act, 2005 The Personal Information Protection Act (PIPA), 2011 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 	<ul style="list-style-type: none"> ✓ Guidelines for the Protection of Bio-information ✓ Guidelines for the protection of biometric information, 2005 ✓ Guidelines for Clinical Trial Management by the KGCP, 2005 ✓ Guidelines for the Protection of Personal Information in Hospitals, 2012

Slovenia	<p>National level:</p> <ul style="list-style-type: none"> The Data Protection Act (Zakon o varstvu osebnih podatkov, ZVOP-2) (O.G. 163/22) Healthcare Databases Act (Zakon o zbirkah podatkov s področja zdravstvenega varstva, ZZPPZ, 65/00, 47/15, 31/18, 152/20 – ZZUOOP, 175/20 – ZIUOPDVE, 203/20 – ZIUPOPDVE in 112/21 – ZNUPZ) Patients' Rights Act, 2008 Slovenian Consumer Protection Act (Zakon o varstvu potrošnikov, UL RS No. 98/2004, 126/2007, 86/2009, 78/2011, 38/2014 and 19/2015) Electronic Commerce Market Act (Zakon o elektronskem poslovanju na trgu, UL RS No. 96/2009 and 19/2015) (the "ZEPT") Slovenian Electronic Communications Act (Zakon o elektronskih komunikacijah, UL RS No. 109/2012, 110/2013) (the "ZEKom-1"), which implemented the Privacy and Electronic Communications Directive. Many sectoral laws define data filing systems for public and private sectors, e.g., around data storage <p>Supra-national level:</p> <ul style="list-style-type: none"> EU Regulation 2016/679 (General Data Protection Regulation) 	<ul style="list-style-type: none"> ✓ Resolution on the National Healthcare Plan, 2016–2025 ✓ National eHealth Strategy was presented January 2023
Sweden	<p>National level:</p> <ul style="list-style-type: none"> The Public Access to Information and Secrecy Act The Personal Data Act The Data Protection Act The Patient Data Act Health Data Register Act Public Access to Information and Secrecy Act <p>Supra-national level:</p> <ul style="list-style-type: none"> EU Regulation 2016/679 (General Data Protection Regulation) 	<ul style="list-style-type: none"> ✓ National eHealth Strategy, 2010
United Kingdom	<p>National level:</p> <ul style="list-style-type: none"> The Data Protection Act, 2018 The Health Service (Control of Patient Information) Regulations, 2002 <p>Supra-national level:</p> <ul style="list-style-type: none"> UK-GDPR 	<p>England:</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, 2016 <p>Scotland:</p> <ul style="list-style-type: none"> ✓ A Health and Biomedical Informatics Research Strategy, 2015 ✓ Joined up data for better decisions: a strategy for improving data access and analysis, 2012

United States	<p>National level:</p> <ul style="list-style-type: none">• The Health Insurance Portability and Accountability Act (HIPAA)• HIPAA Privacy Rule• HIPAA Security Rule• The 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act <p>State level: Around half the states in the US have enacted health privacy rules that apply in addition to HIPAA. These tend to be more protective of patient privacy than HIPAA and concern specific clinical conditions or circumstances (e.g., HIV/AIDS status, mental or reproductive health conditions, health information of teenagers).</p>	<ul style="list-style-type: none">✓ Privacy, Security, and Breach Notification Rules (implementing rules of HIPAA)✓ The Health Data Initiative✓ The Promoting Interoperability proposed rule
----------------------	---	--

The Health Systems in Transition Series

A series of the European Observatory on Health Systems and Policies

The Health Systems in Transition (HiT) country reports provide an analytical description of each health system and of reform initiatives in progress or under development. They aim to provide relevant comparative information to support policy-makers and analysts in the development of health systems and reforms in the countries of the WHO European Region and beyond.

The HiTs are building blocks that can be used:

- to learn in detail about different approaches to the financing, organization and delivery of health services;
- to describe accurately the process, content and implementation of health reform programmes;
- to highlight common challenges and areas that require more in-depth analysis; and
- to provide a tool for the dissemination of information on health systems and the exchange of experiences of reform strategies between policymakers and analysts in countries of the WHO European Region.

How to obtain a HiT

All HiTs are available as PDF files at www.healthobservatory.eu where you can also join our list for monthly updates of the activities of the European Observatory on Health Systems and Policies, including new HiTs, books in our co-published series with Cambridge University Press, Policy briefs, Policy Summaries, and the Eurohealth journal.

If you would like to order a paper copy of a HiT, please contact us at:
contact@obs.who.int



The publications of the
European Observatory on
Health Systems and Policies
are available at:
www.healthobservatory.eu

HiT Country Reviews Published to Date

Albania (1999, 2002 ^a)	Italy (2001, 2009, 2014, 2022)	Switzerland (2000, 2015)
Andorra (2004)	Japan (2009)	Tajikistan (2000, 2010 ^a , 2016)
Armenia (2001 ^a , 2006, 2013)	Kazakhstan (1999 ^a , 2007 ^a , 2012)	Türkiye (2002 ^a , 2011 ⁱ)
Australia (2002, 2006)	Kyrgyzstan (2000 ^a , 2005 ^a , 2011 ^a , 2022)	Turkmenistan (2000)
Austria (2001 ^a , 2006 ^e , 2013 ^e , 2018)	Latvia (2001, 2008, 2012, 2019)	Ukraine (2004 ^a , 2010 ^a , 2015)
Azerbaijan (2004 ^a , 2010 ^a)	Lithuania (2000, 2013)	United Kingdom of Great Britain and Northern Ireland (1999 ^a , 2015, 2022)
Belarus (2008 ^a , 2013)	Luxembourg (1999, 2015)	United Kingdom (England) (2011)
Belgium (2000, 2007, 2010, 2020)	Malta (1999, 2014, 2017)	United Kingdom (Northern Ireland) (2012)
Bosnia and Herzegovina (2002 ^a)	Mexico (2020)	United Kingdom (Scotland) (2012)
Bulgaria (1999, 2003 ^b , 2007 ^a , 2012, 2018)	Mongolia (2007)	United Kingdom (Wales) (2012)
Canada (2005, 2013 ^c , 2020)	Netherlands (2004 ^a , 2010, 2016)	United States of America (2013, 2020)
Croatia (1999, 2006, 2014, 2022)	New Zealand (2001*)	Uzbekistan (2001 ^a , 2007 ^a , 2014 ^a)
Cyprus (2004, 2012)	North Macedonia (2000, 2006, 2017)	Veneto Region, Italy (2012)
Czechia (2000, 2005 ^a , 2009, 2015, 2023)	Norway (2000, 2006, 2013, 2020)	
Denmark (2001, 2007 ^a , 2012)	Poland (1999, 2005 ^k , 2011, 2019)	
Estonia (2000, 2004 ^a , 2008, 2013, 2018)	Portugal (1999, 2004, 2007, 2011, 2017)	
Finland (2002, 2008, 2019)	Republic of Korea (2009*)	
France (2004 ^c , 2010, 2015)	Republic of Moldova (2002 ^a , 2008 ^a , 2012)	
Georgia (2002 ^a , 2009, 2017)	Romania (2000 ^f , 2008, 2016)	
Germany (2000 ^e , 2004 ^e , 2014 ^e , 2020)	Russian Federation (2003 ^a , 2011 ^a)	
Greece (2010, 2017)	Serbia (2019)	
Hungary (1999, 2004, 2011)	Slovakia (2000, 2004, 2011, 2016)	
Iceland (2003, 2014)	Slovenia (2002, 2009, 2016, 2021)	
Ireland (2009)	Spain (2000 ^b , 2006, 2010, 2018)	
Israel (2003, 2009, 2015)	Sweden (2001, 2005, 2012)	

All HiTs are available in English.

When noted, they are also available in other languages:

- ^a Albanian
- ^b Bulgarian
- ^j Estonian
- ^c French
- ^d Georgian
- ^e German
- ^k Polish
- ^f Romanian
- ^g Russian
- ^h Spanish
- ⁱ Turkish



The publications of the
European Observatory
on Health Systems
and Policies
are available at
www.healthobservatory.eu





The Observatory is a partnership, hosted by WHO/Europe, which includes other international organizations (the European Commission); national and regional governments (Austria, Belgium, Finland, Ireland, Kingdom of the Netherlands, Norway, Slovenia, Spain, Sweden, Switzerland, the United Kingdom and the Veneto Region of Italy (with Agenas)); other health system organizations (the French National Union of Health Insurance Funds (UNCAM), the Health Foundation); and academia (the London School of Economics and Political Science (LSE) and the London School of Hygiene & Tropical Medicine (LSHTM)). The Observatory has a secretariat in Brussels and it has hubs in London (at LSE and LSHTM) and at the Berlin University of Technology.

HiTs are in-depth profiles of health systems and policies, produced using a standardized approach that allows comparison across countries. They provide facts, figures and analysis and highlight reform initiatives in progress.

Print ISSN 1817-6119, ISBN 978 92 890 5942 8 Web ISSN 1817-6127, ISBN 978 92 890 5946 6

