Accessing Health and Health-Related Data in Canada

Executive Summary

Council of Canadian Academies
Conseil des académies canadiennes

Science Advice in the Public Interest

ACCESSING HEALTH AND HEALTH-RELATED DATA IN CANADA

The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation

THE COUNCIL OF CANADIAN ACADEMIES
180 Elgin Street, Suite 1401, Ottawa, ON, Canada K2P 2K3

Notice: The project that is the subject of this report was undertaken with the approval of the Board of Governors of the Council of Canadian Academies. Board members are drawn from the Royal Society of Canada (RSC), the Canadian Academy of Engineering (CAE), and the Canadian Academy of Health Sciences (CAHS), as well as from the general public. The members of the expert panel responsible for the report were selected by the Council for their special competencies and with regard for appropriate balance.

This report was prepared for the Government of Canada in response to a request from the Canadian Institutes of Health Research (CIHR). Any opinions, findings, or conclusions expressed in this publication are those of the authors, the Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation, and do not necessarily represent the views of their organizations of affiliation or employment.

Library and Archives Canada Cataloguing in Publication

Accessing health and health-related data in Canada : the Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation.

Includes bibliographical references.

Electronic monograph in PDF format.

Summary: "The Canadian Institutes of Health Research (CIHR) has asked the Council of Canadian Academies to assess timely access to health and social data for health research and health system innovation in Canada." – Provided by publisher.

ISBN 978-1-926522-05-0 (pdf)

1. Medical records–Access control–Canada. 2. Medical records–Law and legislation–Canada. 3. Data protection–Law and legislation–Canada. 4. Medical records–Management–Canada. 5. Privacy, Right of–Canada. 6. Health–Research–Canada. I. Council of Canadian Academies, issuing body II. Council of Canadian Academies. Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation, author

R864.A24 2015 651.5'04261 C2015-901476-X

This report should be cited as: Council of Canadian Academies, 2015. *Accessing Health and Health-Related Data in Canada*. Ottawa (ON): The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation, Council of Canadian Academies.

Disclaimer: The internet data and information referenced in this report were correct, to the best of the Council's knowledge, at the time of publication. Due to the dynamic nature of the internet, resources that are free and publicly available may subsequently require a fee or restrict access, and the location of items may change as menus and webpages are reorganized.

Printed in Ottawa, Canada

Council of Canadian Academies
Conseil des académies canadiennes

Canada This assessment was made possible with the support of the Government of Canada.

The Council of Canadian Academies

Science Advice in the Public Interest

The Council of Canadian Academies (the Council) is an independent, not-for-profit organization that supports independent, science-based, authoritative expert assessments to inform public policy development in Canada. Led by a 12-member Board of Governors and advised by a 16-member Scientific Advisory Committee, the Council's work encompasses a broad definition of *science*, incorporating the natural, social, and health sciences as well as engineering and the humanities. Council assessments are conducted by independent, multidisciplinary panels of experts from across Canada and abroad. Assessments strive to identify emerging issues, gaps in knowledge, Canadian strengths, and international trends and practices. Upon completion, assessments provide government decision-makers, researchers, and stakeholders with high-quality information required to develop informed and innovative public policy.

All Council assessments undergo a formal report review and are published and made available to the public free of charge in English and French. Assessments can be referred to the Council by foundations, non-governmental organizations, the private sector, or any level of government.

The Council is also supported by its three founding Member Academies:

The Royal Society of Canada (RSC) is the senior national body of distinguished Canadian scholars, artists, and scientists. The primary objective of the RSC is to promote learning and research in the arts and sciences. The RSC consists of nearly 2,000 Fellows — men and women who are selected by their peers for outstanding contributions to the natural and social sciences, the arts, and the humanities. The RSC exists to recognize academic excellence, to advise governments and organizations, and to promote Canadian culture.

The Canadian Academy of Engineering (CAE) is the national institution through which Canada's most distinguished and experienced engineers provide strategic advice on matters of critical importance to Canada. The Academy is an independent, self-governing, and non-profit organization established in 1987. Fellows are nominated and elected by their peers in recognition of their distinguished achievements and career-long service to the engineering profession. Fellows of the Academy, who number approximately 600, are committed to ensuring that Canada's engineering expertise is applied to the benefit of all Canadians.

The Canadian Academy of Health Sciences (CAHS) recognizes individuals of great achievement in the academic health sciences in Canada. Founded in 2004, CAHS has approximately 400 Fellows and appoints new Fellows on an annual basis. The organization is managed by a voluntary Board of Directors and a Board Executive. The main function of CAHS is to provide timely, informed, and unbiased assessments of urgent issues affecting the health of Canadians. The Academy also monitors global health-related events to enhance Canada's state of readiness for the future, and provides a Canadian voice for health sciences internationally. CAHS provides a collective, authoritative, multidisciplinary voice on behalf of the health sciences community.

www.scienceadvice.ca
@scienceadvice

Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation

Andrew K. Bjerring (Chair), Former President and CEO, CANARIE Incorporated (London, ON)

Marni Brownell, Senior Research Scientist, Manitoba Centre for Health Policy; Associate Professor, Department of Community Health Sciences, College of Medicine, Faculty of Health Sciences, University of Manitoba (Winnipeg, MB)

Brent Diverty, Vice President of Programs, Canadian Institute for Health Information (Ottawa, ON)

Khaled El Emam, Associate Professor, Faculty of Medicine, University of Ottawa; Senior Investigator, Children's Hospital of Eastern Ontario Research Institute; Canada Research Chair in Electronic Health Information, University of Ottawa; CEO, Privacy Analytics Inc. (Ottawa, ON)

Isabel Fortier, Researcher, Maelstrom Research Institute, McGill University Health Centre (Montréal, QC)

David Henry, Professor, Dalla Lana Faculty of Public Health, University of Toronto; Senior Scientist and former CEO, Institute for Clinical Evaluative Sciences (Toronto, ON)

Bartha Maria Knoppers, O.C., O.Q., FCAHS, Canada Research Chair in Law and Medicine; Director, Centre of Genomics and Policy, Faculty of Medicine, McGill University (Montréal, QC)

Graeme Laurie, Professor of Medical Jurisprudence, University of Edinburgh; Founding Director, JK Mason Institute for Medicine, Life Sciences and the Law (Edinburgh, United Kingdom)

Trudo Lemmens, Professor and Scholl Chair in Health Law and Policy, Faculty of Law; Professor, Joint Centre for Bioethics and Faculty of Medicine, University of Toronto (Toronto, ON)

Matthew Morgan, Vice President, Patient Experience and Outcomes, Mount Sinai Hospital; Assistant Professor, Division of General Internal Medicine, Department of Medicine, Faculty of Medicine, University of Toronto (Toronto, ON)

Thomas William Noseworthy, C.M., Professor, Health Policy Management, Department of Community Health Sciences and O'Brien Institute for Public Health, University of Calgary (Calgary, AB)

Stephen Saunders, Senior Executive Consultant and Chief Architect, Healthcare, CGI Group (Eganville, ON)

Michael Wolfson, FCAHS, Canada Research Chair, Population Health Modelling/Populomics, Faculty of Medicine, University of Ottawa (Ottawa, ON)

Jennifer Zelmer, Executive Vice President, Canada Health Infoway (Toronto, ON)

Message from the Chair

The objectives of improving the health and well-being of Canadians and of the health system both require ongoing research and innovation. One of the major requirements for addressing these challenges is the availability of high-quality data, including data on individuals and their encounters with service providers in the health system as well as social data on factors that affect health outcomes. At the same time, individuals have a right to privacy; there is a clear obligation that personal health-related data are kept confidential. Striking an appropriate balance between these two imperatives is of fundamental importance. It is also of great concern to numerous organizations and individuals in every jurisdiction in the world, perhaps none more than those who have a responsibility to act as custodians of the data involved.

Ideally, the organizations and individuals who contribute to this collective effort, whether within a single province or territory or at the national level in a federated jurisdiction like Canada, would constitute a coherent and smoothly operating system with well-defined governance principles and efficient operating procedures that, among other things, would support timely access to health and social data for research and system innovation. This tends not to be the case in Canada. Indeed, those who need access to data must navigate a "complex environment of heterogeneous entities," often including numerous data custodians, privacy offices, and research ethics boards, whose collective governance and operational practices fall short of constituting a well-defined and coherent system.

To address the challenge of providing timely access to health and social data within this context, the Expert Panel was asked, among other things, to identify where the provision of such access could be seen as constituting a "best practice." One particularly noteworthy finding of this report is that many of the "best practice entities" identified here were themselves created as a result of a review of the collective behaviour of the complex environment existing in their particular jurisdiction. In other words, the undertaking of a review by a provincial, territorial, or federal jurisdiction of how well its complex environment addresses *collective* governance responsibilities itself constitutes a best practice.

On behalf of the Panel I would like to thank those who met with us early in the process to help us tackle our charge. I would also like to acknowledge the significant contribution of Council staff to the Panel's work, which would have been impossible without their professionalism, patience, and insight into how we might best make a contribution. Finally, I would like to personally thank the Panel members for their dedication and hard work. I cannot imagine any group of individuals better positioned to help the cause of providing timely access to data for health research and system innovation in Canada. Their report deserves careful consideration.

Andrew K. Bjerring,
Chair, Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation

Acknowledgments

Over the course of its deliberations, the Panel sought assistance from many individuals and organizations that provided valuable evidence, information, and assistance in the development of the report. The report also benefitted from numerous interactions with participants at the International Health Data Linkage Conference held in April 2014 in Vancouver, Canada. Special thanks go to the following: Carolyn Adams, Macquarie University; Judy Allen, University of Western Australia; Jane Badets and Lynn Barr-Telford, Statistics Canada; Daniel Bedard, BORN Ontario; Christiane Bétie, Commission d'accès à l'information du Québec; Pam Bjornson, National Research Council Canada; Charles Burchill, University of Manitoba; Ann Cavoukian, Former Information and Privacy Commissioner, Ontario; Geoff Davis, Department of Health Western Australia; Felicity Flack, Population Health Research Network; David Ford, Swansea University; Karey Iron, Institute for Clinical Evaluative Sciences; Patricia Kosseim, Office of the Privacy Commissioner of Canada; Nancy Meagher, University of British Columbia; Peter Morrison, Statistics Canada; Stephen Pavis, Farr Institute @ Scotland; Daryl Pullman, Memorial University; Parminder Raina, McMaster University; Diana Rosman, Department of Health Western Australia; Mark Smith, University of Manitoba; Merran Smith, Population Health Research Network; and Alan Winter, Genome British Columbia.

The Council wishes to thank the Honourable Elizabeth Dowdeswell, O.C., O.Ont., Lieutenant Governor of Ontario, and Former President of the Council of Canadian Academies, for her guidance and support during the early stages of this assessment.

Project Staff of the Council of Canadian Academies

Assessment Team:	Christina Stachulak, Senior Program Director Aled ab Iorwerth, Research Associate Jennifer Bassett, Research Associate Weronika Zych, Program Coordinator
With assistance from:	Haryanto Darmawan, Consultant Thomas Archibald, Consultant Edward Dove, Consultant Carolyn Brown, Editor Clare Walker, Editor and Copyeditor Accurate Design & Communication, Report Design

Report Review

This report was reviewed in draft form by the individuals listed below — a group of reviewers selected by the Council of Canadian Academies for their diverse perspectives, areas of expertise, and broad representation of academic, industrial, policy, and non-governmental organizations.

The reviewers assessed the objectivity and quality of the report. Their submissions — which will remain confidential — were considered in full by the Panel, and many of their suggestions were incorporated into the report. They were not asked to endorse the conclusions, nor did they see the final draft of the report before its release. Responsibility for the final content of this report rests entirely with the authoring Panel and the Council.

The Council wishes to thank the following individuals for their review of this report:

Mark Dermer, Family Physician and Managing Associate, Central Ottawa Family Medicine Associates (Ottawa, ON)

David V. Ford, Professor of Health Informatics, College of Medicine, Swansea University (Swansea, Wales)

Patricia Kosseim, Senior General Counsel and Director General, Legal Services, Policy, Research and Technology Analysis Branch, Office of the Privacy Commissioner (Ottawa, ON)

Adrian Levy, Professor and Head, Department of Community Health and Epidemiology, Dalhousie University (Halifax, NS)

David Loukidelis, Q.C., Former Information and Privacy Commissioner for British Columbia; Former Deputy General of British Columbia; Former Adjunct Professor, Freedom of Information and Privacy Law, Faculty of Law, University of Victoria; Former Adjunct Professor, Freedom of Information and Privacy Law, Faculty of Law, Thompson Rivers University (Edmonton, AB)

Stan Matwin, Canada Research Chair and Professor, Director, Institute for Big Data Analytics, Faculty of Computer Science, Dalhousie University (Halifax, NS)

Ted McDonald, Professor, Department of Economics, University of New Brunswick (Fredericton, NB)

Nancy Meagher, Executive Director, Population Data BC (Vancouver, BC)

Noralou Roos, C.M., FRSC, FCAHS, Professor, Faculty of Health Sciences, University of Manitoba (Winnipeg, MB)

Charlie Schick, Director, Business Development, Healthcare, Atigeo LLC (Bellevue, WA)

Fiona Stanley, Patron, Telethon Kids Institute; Distinguished Research Professor, The University of Western Australia; Vice Chancellor's Fellow, The University of Melbourne (Perth, Australia)

Don Willison, Associate Professor, Institute for Health Policy, Management and Evaluation, University of Toronto (Toronto, ON)

Eric R. Wright, Professor, Department of Sociology, College of Arts and Science and School of Public Health, Georgia State University (Atlanta, GA)

Glenda Yeates, Former Deputy Minister, Health Canada (Ottawa, ON)

The report review procedure was monitored on behalf of the Council's Board of Governors and Scientific Advisory Committee by **Lorne Babiuk, O.C., FRSC, FCAHS,** Vice President (Research), University of Alberta (Edmonton, AB). The role of the report review monitor is to ensure that the Panel gives full and fair consideration to the submissions of the report reviewers. The Board of the Council authorizes public release of an expert panel report only after the report review monitor confirms that the Council's report review requirements have been satisfied. The Council thanks Dr. Babiuk for his diligent contribution as report review monitor.

Janet Bax,
Interim President Council of Canadian Academies

Executive Summary

INTRODUCTION

Canadians care deeply about health care — for themselves, their families, and their communities. Ensuring that the health-care system can deliver the best possible care depends fundamentally on research into system innovation and health and social well-being. This research depends on the availability of high-quality data.

The volume and variety of data relevant to such research have increased exponentially in recent years. Each patient interaction with a physician, a pharmacist, a laboratory technician, or hospital staff generates data. Social and environmental data are highly relevant to health research because they are vital for providing a complete picture about factors that affect the lives and health of Canadians. The research community, including health system innovators in hospital and government offices as well as academic researchers and clinicians, views these data as a critical resource. It recognizes the enormous potential of using health and health-related data in privacy-sensitive ways to reveal factors that can affect health and well-being, and discover interventions that can improve health outcomes.

Despite these benefits, working with the data on which the research is based can be challenging. Some challenges are technical, such as the use of different standards in different jurisdictions to record important data. Others are related to privacy concerns: access to health data for research carries the risk that personal data could be released, whether inadvertently or intentionally. The greatest challenges, which are indeed barriers to beneficial research, are institutional. These include the application of differing, and in some instances overly cautious, interpretations of privacy legislation, and complex and lengthy approval processes that impede researchers' access to data.

The primary, overarching challenge in Canada, as in other jurisdictions, is to meet two fundamental goals at the same time: to enable access to health and health-related data for research that is in the public interest, on the one hand, and to respect Canadians' privacy and maintain confidentiality of their information when it is used for research, on the other. Innovative organizations and less formal collaborative undertakings are finding ways to meet these goals. They are instituting governance models and practices that further scientifically sound, ethically robust research and respect privacy, while using technology in innovative ways to provide data access in a timely and confidentiality-preserving manner.

Charge to the Panel

In 2013, the Canadian Institutes of Health Research (CIHR; the Sponsor) asked the Council of Canadian Academies (the Council) to respond to the following charge:

> *What is the current state of knowledge surrounding timely access to health and social data for health research and health system innovation in Canada?*

The charge also included five sub-questions:

- *What is known about how to address technological and methodological challenges (such as variable data quality and comparability) associated with linkage of health and social data from various sources and across jurisdictions?*
- *What is known about the benefits, risks and barriers to timely access to health and social data for health research and health system innovation in Canada?*
- *What are the ethical, legal, and social implications of timely access to such data?*
- *What are best practices for improving access to such data for researchers while ensuring appropriate privacy safeguards and also taking full advantage of the digital data revolution?*
- *What are best practices in Canada and internationally for governance frameworks that facilitate access to such data and maintain public trust in the research enterprise?*

To address these questions, the Council formed the Expert Panel on the Timely Access to Health and Social Data for Health Research and Health System Innovation (the Panel), which comprised 14 Canadian and international experts from the health-care sector, academia, and industry. Panel members had experience as data custodians, researchers, managers of health research organizations, or in legal aspects of health research.

At the outset of the assessment, the Sponsor gave further direction on interpreting and refining the charge. First, the Sponsor defined *timely access* as access granted within four months of submitting a data request to an organization responsible for providing the data. Second, the assessment should concern only public interest research (i.e., research conducted by public bodies and/or supported by public funds). Thus, health and health-related data used by private, commercial

companies were excluded. Third, the assessment should identify best practices in Canada and other countries for timely access to data that can be linked and integrated for research purposes, rather than exploring barriers to accessing data in general. Finally, the Panel's work should encompass all types of health data related to publicly funded research, ranging from administrative health data to genomic data. The Panel understood the inclusion of social data as data on non-medical determinants of health such as health behaviours, living and working conditions, personal resources, and environmental factors, and hence uses the term *health-related data*.

Methodology for Identifying Best Practices

The identification of best practices was the first issue addressed because of broad implications for the Panel's overall approach. The Panel looked for organizations, institutions, programs, or other entities that had been especially successful in meeting the twin goals of enabling timely access and protecting privacy. It selected six entities, three from Canada and three from other jurisdictions with similar legal and social systems:

- Manitoba Centre for Health Policy (MCHP)
- Ontario – Institute for Clinical Evaluative Sciences (ICES)
- Ontario – Better Outcomes Registry and Network (BORN)
- Wales Secure Anonymised Information Linkage Databank (SAIL)
- Data Linkage Western Australia (Data Linkage WA)
- Farr Institute @ Scotland[1]

These "best practice entities" are mandated, in some cases under legislation, to receive data from encounters in the health-care system and to provide access for public interest research. They all succeed in providing access within a four-month timeframe and share four common principles:

- **Enabling appropriate use of data** to enhance public well-being;
- **Managing risk** by identifying the range of risks involved in providing data access and minimizing those risks where possible, while acknowledging that risks cannot be entirely eliminated;
- **Respecting privacy** to reassure citizens that risks to their core personal interests are kept to an absolute minimum; and
- **Maintaining public trust** by providing evidence of trustworthiness, including using data appropriately and demonstrating the social value of the resulting research.

1 The Farr Institute @ Scotland builds on the success of the Scottish Informatics Programme (SHIP), which ran from 2009 to 2013.

The practices highlighted by the Panel reflect both a literature review as well as the practices of the six entities. The Panel found many examples of *good practice*, including approaches for dealing with legal and ethical considerations. These are highlighted in Finding 4 below. However, in accordance with the charge, the Panel only identified *best practice* related to governance that could be put in place to enable access to health data. Best practices for governance are discussed in Finding 5. *Good practice* and *best practice* are defined in the glossary that accompanies the full report.

The Panel also examined many other organizations that provide access to health and health-related data, or that play a special role in analyzing such data, including, among others, Statistics Canada, Statistics Netherlands, the U.S. National Institutes of Health, Population Data BC, the Canadian Network for Observational Drug Effect Studies (CNODES), and the Canadian Institute for Health Information (CIHI). Many insights can be drawn from these innovative organizations about striking an appropriate balance between respecting privacy and providing timely access to data.

In addition to identifying best practice organizations, the Panel reviewed evidence on how other organizations in Canada and around the world enable access to health and health-related data for research. It drew from published literature, conference proceedings, and online reports, including a key OECD report (*Strengthening Health Information Infrastructure for Health Care Quality Governance*) that summarizes the use of health data in a range of countries.

KEY FINDINGS

The Panel's findings fall into five categories that roughly correspond to the five sub-questions of the charge: technological and methodological challenges, benefits, risks, legal and ethical considerations, and governance.

Overall, the Panel found data-intensive research has both clear benefits and risks. Striking the right balance can be achieved through good governance that demonstrates respect for legal and ethical considerations, and for the people whose data are being used.

1. Technological and Methodological Challenges of Access to Health Data

For effective research with health and health-related data, disparate sources of data must be brought together. Providing these data in an "analysis-ready" format, thereby allowing statistical relationships or patterns to be derived, is a central methodological challenge.

The full potential of Canada's health and health-related data can only be realized if the data are made ready for analysis. However, much of the data with the greatest potential for research are collected for other purposes, such as administration of health care services. To be used for research, these data need to be transformed into specific forms and formats — predominantly statistical ones. As electronic health records (EHRs) become increasingly prevalent, it will be more efficient to anticipate and design into these data the capacity to support secondary use rather than to retrofit after computer systems for EHR recording have already been designed.

EHRs and health-care encounter data inherently involve many disparate sources of data, from hospitalizations to lab tests. Thus, for research as well as effective patient care, it is necessary to bring different data sets together. The key difference is that for patient care, the focus is on a single patient, while for research, the focus is on large samples of individuals, where any given individual's identity is irrelevant. As a result, research-oriented data sets may be from the same province/territory, multiple provinces and territories, or multiple countries.

To be compared or combined and used meaningfully in statistical analysis, data elements must be *harmonized*. The best approach for harmonization involves the development of standard terminologies, questionnaires, measurements, and protocols (i.e., *prospective harmonization*). But this approach may be too challenging, time-consuming, or labour-intensive; or an underlying consensus on how to define or measure a given variable may be absent. In these situations, *retrospective harmonization* can be attempted. Tools are available to help determine whether similar inferences can be drawn from variables across different studies.

Data linkage allows different types of information for one individual to be brought together. It can be challenging if (i) unique identifiers are not available for all individuals in a data set, or (ii) data have been strongly de-identified.[2] To overcome the first challenge, probabilistic methods can be used to link records. The simplest solution to the second challenge is to link the data prior to de-identification, if possible. Databases do not always need to be linked permanently. The link can be destroyed after the research is completed, and/or kept completely separate by implementing the *separation principle*.

2 De-identification is the act of minimally perturbing individual-level data to decrease the probability of discovering an individual's identity. It involves masking direct identifiers (e.g., name, phone number, address) as well as transforming indirect identifiers that could be used alone or in combination to re-identify an individual (e.g., birth dates, geographic details, dates of key events).

Pooling of similar data from several populations is often used to increase the sample size for a study. *Bona fide* pooled data analysis involves physical transfer of individual-level data to a central server, where the data are then analyzed as they would be if they were from the same study (with statistical adjustments if needed). In many important cases in Canada, restrictive interpretations of privacy and other laws have hindered pooling of individual-level data from different provinces. Therefore, approaches that avoid the need to pool individual-level data have been developed. One of these approaches (used by CNODES) involves statistical analyses of harmonized, individual-level data at each study site, followed by pooling of the (non-confidential) summary statistics to obtain an overall result. Another, provided by DataShield, uses sophisticated iterative techniques to mimic a pooled analysis of data from individual participants, when, in reality, the data always remain with their original data custodian.

2. Benefits of Access to Health Data

Evidence shows that timely access to data enables significant high-quality research that can have far-reaching effects for health care and the overall health of Canadians.

Timely access to health and health-related data enables significant high-quality research, which identifies risk factors for various health and social outcomes, and determines health interventions with the most beneficial effects. The knowledge gained from this research is fundamental for improving health generally, and maintaining high quality health care. Recent Canadian studies with significant clinical or public health implications have demonstrated the benefits of research using health and health-related data. For example, analysis of data from the Canadian Community Health surveys by researchers at ICES and Public Health Ontario led to the development of a Life Expectancy Calculator that helps Ontarians understand the effect of certain behaviours on their life expectancy. Researchers at MCHP used record linkage to show that low socio-economic status affects educational achievement much more than previously thought. CNODES analysis of hospital data from across Canada showed that seniors over age 65 were five times more likely than the rest of the population to be hospitalized for adverse drug reactions due to specific risk factors such as drug interactions.

3. Risks of Access to Health Data

The risk of potential harm resulting from access to data is tangible but low. The level of risk can be further lowered through effective governance mechanisms.

While there are clear benefits of research using individual Canadians' personal health and health-related data, there are also risks. These can include accidental release of identifiable data, to the public or unauthorized researchers, when proper security and privacy protocols are not followed (e.g., through loss of computer equipment); illicit access to identifiable data (e.g., through hacking); and inadvertent access to identifiable data by those working inside data organizations.

While these types of breaches have occurred during research projects, breaches rarely happen at institutions with databases set up specifically for maintaining large volumes of health and health-related data for research and administrative purposes. They are much more likely to occur when researchers or employees are accessing data directly from health-care centres. Importantly, there are no examples of breaches at the six best practice entities identified by the Panel.

In many cases, the data that researchers access from secure facilities are de-identified. However, re-identification remains a concern. The Panel found that best practices in de-identification can lower the risk of re-identification to acceptable levels. Although health data breaches can cause serious harm, the risk of a breach actually occurring in the context of research is low, particularly if effective governance mechanisms and protocols are in place and respected by care providers, researchers, and data custodians.

4. Legal and Ethical Considerations of Access to Data

Timely access to data is hindered by variable legal structures and differing interpretations of the terms *identifiable* and *de-identified* across jurisdictions. Instead of rigidly classifying data as either identifiable or non-identifiable, it is useful to view de-identification as a continuum and to adjust access controls accordingly.

In enabling access to data for research, the benefits of research, as well as the range of risks, need to be weighed. Canadian research projects demonstrate that beneficial research can be advanced while maintaining confidentiality of sensitive personal information. Yet, access to data and successful data-based research is not uniform across Canada because of (i) the lack of consistency and clarity in Canada's ethical and legal framework, and (ii) differing interpretations of key terms and issues across the country.

While federal and provincial/territorial laws generally allow researchers to access data that do not include "identifiable information," this term is not always defined precisely. This makes it confusing to base data sharing guidelines on the notion that "non-identifiable data" can be used freely. As well, data custodians may interpret their legal duty to protect privacy as precluding access. Laws on sharing data across provinces/territories and countries differ or are lacking, which can also make researchers and research ethics boards (REBs) uncertain as to whether data can be shared.

This lack of legal clarity has contributed to cautious and conservative interpretations of allowable access in many Canadian organizations. While the law provides specific limits for data custodians, it is less specific in other areas. And although provincial and federal laws lay out broad rules about when and how data can be used or shared, often they are silent on specific questions about *whether* data should be so used in specific settings. This means that data custodians often face an asymmetry — there are clear sanctions if there is a data breach when they are in charge, but no benefit to them if their release of data for *bona fide* research generates important public benefits. This asymmetry supports a tendency to *not* grant access, even if access would be acceptable within their legal frameworks.

A number of good practices for addressing legal and ethical issues are summarized below.

Good Practices – Legal and Ethical Considerations

Appropriate access controls for differing levels of de-identification: Because data may be fully identifiable (i.e., no identifiers removed), mildly de-identified, or strongly de-identified, the Panel did not single out one specific process for dealing with de-identified data. Rather, a good practice is to use the degree of de-identification to determine the circumstances under which the data may be made accessible for research purposes (i.e., increase access control as identifiability increases).

Rules governing sharing of identifiable data for research purposes: Maintaining a set of rules that govern the sharing and use of fully identifiable or partially de-identified data for research purposes is a good practice. Examples of such rules are as follows:

- **Data are held at designated research entities:** In some provinces, the legislation designates specific entities that may receive health and health-related data without consent for research purposes, acknowledging that establishing such centres is in the public interest.
- **Research meets approval criteria:** To ensure privacy is respected, and to clearly delineate the requirements for access to identifiable data without consent, good practice suggests showing that the research serves the public interest; obtaining consent is impracticable; identifiable data are necessary to the research project; and physical, electronic, software, and all other security measures are appropriately calibrated to protect the data and to sanction any misuses.
- **Researchers sign researcher-custodian agreements:** To ensure that researchers are accountable for protecting data confidentiality, good practice suggests that full and explicit data transfer agreements between researchers and custodians are needed for each research project.

Risk management strategies: The Tri-Council Policy Statement, which governs ethical research in Canada, recognizes that risk cannot be eliminated but should be considered proportionately. Good practice suggests incorporating risk management in all aspects of governance, including ethical governance.

Establishment of dedicated governance: Whatever the applicable law in a given jurisdiction, it may be open to a considerable range of interpretation. A dedicated governing body is a good practice that could, for example, establish reasonable processes to de-identify data, as well as ensure respect for overall legal and ethical principles.

5. Governance

Evidence demonstrates that a shift is occurring among leading entities from a "data custodianship" model to a "data stewardship" model. Central to the success of this shift is the adoption of good governance practices, specifically in privacy governance, research governance, information governance, and network governance.

The Panel found a marked shift among the six best practice entities from a "data custodianship" model, in which holding and securing data are emphasized to the exclusion of other considerations, to a "data stewardship" model, in which enabling access is a core institutional objective proportionately balanced with protecting privacy. The balance is achieved through good governance, which encompasses the definition of an entity's purpose, objectives, values, and policies.

Addressing the question of providing timely access to health data for research is particularly challenging in Canada as the many institutions, organizations, programs, and activities that deal with health and health-related data are only loosely coordinated. They are best thought of as a "complex environment of heterogeneous entities," the parts of which were not designed to work in concert with one another as a system with a common overall purpose.

Over time, coordination, consistency, and overall effectiveness of the "complex environment" could be achieved through the adaptation of the pre-existing entities. Alternatively, the responsible governments could carry out a broad review and subsequent redesign of their system, comparable to that undertaken in Wales or Scotland.

Along with the other organizations in each of their jurisdictions, the six best practice entities share a *collective* responsibility for addressing several cross-cutting aspects of governance. The Panel found four particularly relevant aspects of governance: privacy governance, information governance, research governance, and network governance. When considered together, these four aspects provide a reasonable framework for examining how the complex environment as a whole governs access to health and health-related data for research.

Privacy Governance

Privacy governance involves monitoring the specific risk to privacy posed by data access by researchers and protecting data confidentiality. Such governance may involve specialized knowledge of technology, privacy law, ethics, and statistical methods.

This aspect of governance ensures appropriate use of confidential data in carefully defined circumstances and under specific conditions. Principles may be put in place to guide access to and protection of personal confidential data.

The six best practice entities have dedicated processes to evaluate privacy concerns when enabling data access. For example, MCHP operates within the context of Manitoba legislation where the Health Information Privacy Committee is responsible for approving health research projects that use personal health information held by a government department or agency. In Wales, SAIL's Information Governance Review Panel (IGRP) is dedicated to privacy review, which ensures appropriate de-identification of data and addresses research ethics concerns. In Ontario, data from the BORN database are certified to indicate that they are de-identified in an approved way, and data from ICES are governed by internal procedures set in consultation with the Information and Privacy Commissioner of Ontario.

Best Practices – Privacy Governance

Dedicated Privacy Evaluation: The best practice entities have developed dedicated processes (parallel to REBs) that specifically evaluate privacy concerns when enabling data access.

Research Governance

The processes and entities that govern the research enterprise in Canada face special challenges in connection with research using health and health-related data. While research governance entails many aspects, the panel chose to focus on the REB process. Of particular importance are the requirements for research projects to be approved in advance by an REB, and for data access requests to be approved, often through a separate process. Timeframes for these approvals vary widely across organizations and jurisdictions in Canada, ranging from months to years. Ethics approval for research projects that involve more than one centre or more than one province/territory, in particular, can involve time-consuming (and duplicative) approval processes.

This issue has been addressed in New Zealand and Wales, as well as in two Canadian provinces, through a reduced number of REBs. Alberta decreased its REBs from six to three. Newfoundland and Labrador has created a central research ethics authority that oversees ethics review but can also approve reviews from other boards within and outside the province, thereby avoiding duplicate reviews.

Another challenge arises when REBs and other boards are inconsistent in interpreting ethical and legal guidelines, for example, regarding what constitutes identifiable information. To overcome this potential problem, many countries and Canadian provinces have established a separate review process for data access requests (e.g., HIPC in Manitoba).

Best Practices – Research Governance

Harmonized REB process: To minimize the number of approvals when performing cross-subject or cross-jurisdictional research — and therefore to improve timeliness — certain jurisdictions such as Alberta, New Zealand, and Wales have harmonized the REB process.

Information Governance

Information governance addresses how information is handled within an organization or among organizations. It covers data organizations and their employees, researchers accessing data, and public input. This aspect of governance is concerned with enabling access to data, and doing so within a reasonable timeframe. The best practice entities have made enabling access one of their central purposes, and, as a result, are moving towards a culture of data stewardship.

Physical and technical measures are also required to enable access to data. However, approaches to data access are on a spectrum, with progressively greater security and precautions as data are less aggressively de-identified. Some organizations allow researchers to access data sets containing identifiable information only at secure locations, often called "safe havens" (e.g., MCHP, ICES, Statistics Canada), or through secure internet connections (e.g., Statistics Netherlands, Population Health Research Network in Australia). For both identifiable and de-identified data, however, the researcher is typically bound by confidentiality agreements and/or the research is subject to pre-approval. Data that are very strongly de-identified may be made publically available by large entities. For example, Statistics Canada provides public-use files for data that are rendered non-identifiable within the meaning of the *Statistics Act*. In some cases, however, these highly de-identified data are much less valuable for research.

Linking data sets across organizations could raise the possibility that many employees can access large amounts of identifiable data. To address this, institutional structures can be established to minimize the risks. One way to manage employee access, referred to as the *separation principle*, is to separate data into a demographic component (with identifying information such as name, address, etc.) and a content component (with information such as medicines prescribed, test results, etc.). This prevents any given individual from seeing both components. The separation principle can be observed by using an external organization to deal with identifying information or by managing all data internally but ensuring that identifying data and content data are administratively — and sometimes physically — separated.

A critical element of any information governance model is the determination of an "acceptable" level of risk, which relies on the development of a method to characterize risk. To address this need in the context of product safety, the European Commission has developed a risk assessment matrix. The Farr Institute @ Scotland has adopted a "proportionate governance" approach in which the level of scrutiny for a data linkage request depends on the level of risk that it entails.

To analyze data across provincial or national boundaries, innovative methods are being undertaken. Through CNODES, data on drug effects are analyzed in each province using standardized methods, and a meta-analysis is conducted on a national level to determine the scale of effects for Canada. Other suggestions include encryption of raw data, and security of core identifiable data with release of summary statistics for analysis via the internet. Various techniques are used to ensure that the system works effectively and efficiently. Common features include (i) adoption of privacy management programs, (ii) adoption of an effective risk management framework, and (iii) adoption and documentation of a "reasonable" process of de-identification.

In summary, application of information governance practices can effectively deal with the public concern of risks such as inadvertent access to data and accidental release of data through loss or theft.

Best Practices – Information Governance

Data access: Certain entities successfully maintain data confidentiality through safe havens and/or encrypted access. Key features of a well-functioning safe haven include mechanisms to approve researchers, robust internal and external monitoring and oversight, and ongoing review of governance arrangements over time.

Enabling data use: Appropriate provision of data to researchers is central to the best practice entities. For example, the mission statements of the Farr Institute @ Scotland, SAIL, Data Linkage WA, ICES, and MCHP clearly lay out that enabling appropriate use of data is a core purpose of their organization.

Privacy management: Entities have developed comprehensive researcher-custodian agreements to ensure that researchers maintain the confidentiality of the information that they receive.

Appropriate institutional structure, respecting separation principle: Entities that use the separation principle have minimized the risk of inadvertent and inappropriate access to data by staff.

De-identification of data: Robust de-identification techniques that have met legal standards (i.e., de-identification is "reasonable") have made it possible to reduce the risk of re-identification to a level that is appropriate for a given access mode (and its accompanying security controls). These include practices to ensure that de-identification is documented, transparent, and meets statistical thresholds for re-identification risk while maintaining data utility.

Technology's role in enabling access to and safeguarding data: New technologies can be adopted and developed to improve the safeguards on confidentiality. Given the central importance of technology, it is critical to have individuals with knowledge of its importance involved in governance.

Acceptable level of risk: The European Commission has developed a systematic method for characterizing risk. Scotland has integrated a proportionate approach to risk in its governance system.

Network Governance

The creation of collaborative research networks, potentially involving not just a circle of researchers but also other stakeholders such as data custodians and funding agencies, has the potential to maximize social benefits flowing from data-oriented research. Given Canada's complex and heterogeneous set of actors and stakeholders, governance to create and maintain these networks is vital for standardizing data collection and developing policies for national and international data sharing.

Among the benefits of building a research network is that it may be the only way to amass enough data to conduct a study. A by-product of network-driven collaborations is that definitions and standards must be defined in advance to make the data involved comparable. Thus, networks are a central contributor to standardization and harmonization. Standardization has been a core function of CIHI and the WHO, whose boards and committees represent another type of network, composed of individuals who may have different research interests and diverse professional backgrounds but who share the common goal of developing national or international standards. Standardization is also a main objective of Statistics Canada.

It is also important that networks develop standardized data security protocols. Genetics initiatives such as the International Cancer Genome Consortium (ICGC) are among the most advanced in their successful development of policies for international sharing of individual-level data.

Networks may play a role in mitigating inaccuracy of research results. Analysis of large data sets involves complex statistics, and results can be erroneous if there is a lack of expertise in this area. Networks can put in place standards for statistical analysis and share information about issues in statistical methods. If incorrect research conclusions are publicly released, networks can act to address these both within the scientific community and vis-à-vis the public.

Best Practices – Network Governance

Data harmonization: To enable prospective data harmonization, entities such as the WHO and CIHI have put standards in place prior to data collection.

Distributed analysis: When it is not possible to pool individual-level data, other models, such as CNODES in Canada and DataShield in Europe, have been successful in enabling statistical analysis across jurisdictions.

Multinational sharing: When legal systems differ, methods have been developed to further research by multinational consortia such as the International Cancer Genome Consortium.

A "Best Practice" Governance Model

The jurisdictions involving some of the best practice entities chosen by the Panel, in particular Wales and Scotland, consciously decided to redesign their entire complex environment of entities involved in health and health-related data for research. Their aim was to prevent overlap, duplication, and confusion, and more effectively address the challenges of privacy, information, research, and network governance. For example, one element of Scotland's good governance framework is a mechanism based on proportionate governance to ensure that data access requests with lower risks receive lighter touch governance. Another element is an "account of responsibilities" of key actors and decision-makers. In contrast, the new system in Wales incorporates all governance into a single governance review panel. Clarifying the responsibilities of key entities in Canada's complex environment could be a positive step in enabling timely access to health and health-related data for research. Chapter 5 of the report summarizes the roles of different groups (e.g., researchers, data custodians, policy makers) and governing bodies (e.g., REBs, privacy monitoring boards) in overseeing various aspects of governance and provides examples of entities that are following best practice by successfully performing these roles.

CONCLUSION

To ensure that Canadians continue to have access to high-quality health care, and benefit from effective health policies, the country's health researchers and system innovators need to make effective use of health and health-related data, including administrative health and social data. This need will increase in the future as technology continues to develop and digitized data such as EHRs become ever more abundant.

However, timely access to health and health-related data for research varies across Canada. While some jurisdictions have developed processes that provide access to data within four months, the target provided to the Panel, others can take a year or longer. The reasons for delays are multifold, such as concerns over data quality, lack of a roadmap on how to access data, limited budgets for supporting research, fear of potential legal liabilities in the case of data breaches, or broader fears that the research may generate embarrassing results (e.g., evidence of poor performance).

The Panel found that legal definitions and interpretations differ across provinces/territories and countries, which can lead to confusion or overly cautious interpretations of whether data can be accessed or shared. As a result, careful ethical judgments must be taken sometimes in the absence of specific laws. However, good governance ensures that data can be accessed while respecting ethical principles and the law. In searching for models of good governance, the Panel found that successful entities in Canada and abroad have developed systems of governance incorporating four cross-cutting aspects — namely privacy, information, research, and network governance — that achieve this goal. The Panel has identified specific "best practices" within these aspects of governance that can provide the necessary guidance to help transform what is known as a culture of caution to a culture of trust.

The Panel concluded that, although access to health and health-related data vary across Canada, the exemplary practices identified in this report clearly indicate the feasibility of an elevated standard of appropriate data access for *bona fide* public interest research.