CDA-AMC
Canada's Drug Agency
L'Agence des médicaments du Canada

Methods and Guidelines

Linking Patient and Disease Registry Data With Administrative Health Services Data

Table of Contents

Authors 5

Contributors 5

Pan-Canadian Health Organizations 5

Patient Registries 5

Health Data Research Network Canada 5

Glossary of Terms 6

Introduction 7

Which Registries Are the Focus of the Roadmap? 8

Why Link Registry Data With Administrative Health Services Data? 9

What Are the Challenges? 9

Could the Roadmap Be Used When Linking Other Types of Health Data? 10

Roadmap for Linking Registry Data With Administrative Health Services Data 10

Step 1: Articulate the Registry Purpose 11

Step 2: Translate the Purpose Into Clearly Defined Objectives 13

Step 3: Understand the Goal of Record Linkage and Plan the Process 13

Step 4: Confirm and Document Approval of the Registry Governance Committee 15

Step 5: Confirm That You Have Complied With All Linkage Requirements 16

Step 6: Prepare Documents for Submission (or Resubmission) to Research Ethics Board(s) 16

Step 7: Initiate the Consent Process According to an Established Communications Plan 17

Step 8: Transfer the Registry Data for Linkage 19

Key Considerations and Best Practices When Linking Registry Data for the Purpose of Decision-Making 20

Key Questions 20

Why Do You Need a Registry? 20

What Consent Do You Need to Obtain? 22

Could the Information You Collect Be Used to Identify Participants? 23

How Will You Ensure That Data Entries Are Accurately Linked? 24

How Will You Protect Information so It Cannot Be Used to Identify Participants? 26

How Will You Protect Fully Identifiable Data or Information? 26
How Will You Protect De-Identified Data? 27
How Will You Protect Coded Data? 28
What Do I Need to Know About Sharing Anonymized Data? 28
What Do I Need to Know About Sharing Aggregate Data? 29
Are There Specific Legal Obligations in the Jurisdiction Where Your Registry Operates? 29
What Governance Roles Are Needed to Manage a Data Registry? 30

References 32

List of Tables

Table 1: HDRN Canada Member Organizations That Conduct Record Linkage and/or Support Access to Linked Data 10
Table 2: Fully Identified Dataset 26
Table 3: De-identified Dataset 27
Table 4: Coded Dataset 28
Table 5: Anonymized Dataset 29

List of Figures

Figure 1: Roadmap for Linking Registry Data With Administrative Health Services Data 12

Authors

We would like to thank Magda Nunes de Melo, Anne Hayes, and Charles Victor from ICES who Canada's Drug Agency contracted to lead this work and who developed the contents of this report and Considerations When Developing Consent Forms for Registry Participants.

Contributors

Pan-Canadian Health Organizations

We would like to thank the following pan-Canadian health organizations for reviewing different versions of the 2 reports. The content of these reports does not necessarily always reflect the views of the organizations acknowledged in the following list:

- Health Canada
- Canadian Institute for Health Information (CIHI)
- Health Statistics Branch, Statistics Canada.

We would like to acknowledge the important contributions of the following individuals. The content of these reports does not necessarily always reflect the views of the individuals acknowledged in the following lists:

Patient Registries

- Beth Potter, Inform Rare
- Jay Onysko, Cancer in Young People in Canada, Public Health Agency of Canada
- Katherine McKenzie, Cancer in Young People in Canada, Public Health Agency of Canada
- Paul Gibson, Cancer in Young People in Canada, C17 Council
- Randy Barber, Cancer in Young People in Canada, C17 Council
- Stephanie Cheng, Cystic Fibrosis Canada
- Victoria Hodgkinson, Canadian Neuromuscular Disease Registry

Health Data Research Network Canada

- Charles Burchill, Manitoba Centre for Health Policy
- Donna Curtis Maillet, New Brunswick Institute for Research, Data and Training
- Jeffrey Bakal, Alberta Health Services
- Jodi Gatley, ICES
- Katelyn Frizzell, Health Data Nova Scotia
- Kim McGrail, Population Data BC
- Lindsay Stewart, Health Data Nova Scotia
- Mahmoud Azimaee, ICES

- Michael Schull, ICES
- Michael Smith, ICES
- Nicole Yada, ICES (*at the time of contribution*)
- Ted McDonald, New Brunswick Institute for Research, Data and Training

Glossary of Terms

Administrative health services data: Information collected by organizations, in the public or private sector, as part of their day-to-day operations as part of the process of providing health services to individuals. Examples include records of hospital admissions, physician billings, drugs dispensed at a pharmacy, and billings for diagnostic tests performed at a clinic.

Crosswalk file: A file that maps encrypted codes or study IDs with identifying information (such as health card numbers).

Forward sortation unit: Represents a geographic unit based on the first 3 characters in a Canadian postal code.

Patient-Powered Registry or Patient-Powered Network: A registry (or network) and the research it yields that is managed by patients and family members themselves, often through a disease advocacy organization or a network of organizations that receives advice and input from a scientific board of advisors.[1]

Patient or disease registry: An organized system that uses observational study methods to collect uniform data (clinical and other) to conduct specified evaluations for a population defined by a particular disease, condition, or exposure, and that serves 1 or more stated scientific, clinical, or policy purposes.[2]

Primary stakeholder: Primary stakeholders are usually responsible for creating and funding the registry. They are involved from the onset of the registry planning, and their perspectives are considered when determining which outcomes are most relevant to capture.[2]

Record linkage: The process in which records from different data sources are joined together into a single file using non-unique identifiers, such as names, date of birth, addresses, and other characteristics.[3]

Registry participant: Someone with a specific disease, condition (e.g., risk factor), or exposure whose health information is being collected for the purpose of addressing the registry questions. Because not all these individuals will necessarily be receiving active care, we do not refer to them as “registry patients.”

Registry team: The group of individuals managing a registry.

Secondary stakeholder: Secondary stakeholders are those who will use the knowledge gained from data collected in the registry or who will be impacted by the results but who are not critical for the setup of the registry.[2]

Introduction

In March 2023, the Government of Canada announced the National Strategy for Drugs for Rare Diseases to help increase the access to, and affordability of, effective treatments for these conditions. Part of the strategy includes funding for Canada's Drug Agency-led initiatives to improve the quality and use of evidence to inform and support decision-making. One of these initiatives is to **support efforts for improved evidence generation and the value of real-world data from rare disease registries.**

The term *registry* can be defined as both "an official record of people or things relating to a business or other activity," and "a place where official records are kept."[6] There is no standard definition of *patient registry* used in the health research field, and the term is often used interchangeably with *disease registry, clinical registry*, *clinical data registry*, and *outcomes registry*.[1]

In this document, we use the term "patient registry" or "disease registry" to refer to "an organized system that uses observational study methods to collect uniform data (clinical and other) to conduct specified evaluations for a population defined by a particular disease, condition, or exposure, and that serves one or more stated scientific, clinical, or policy purposes. A registry database is a file (or files) derived from the registry."[2] A registry may have more than 1 goal, and goals may evolve over time. Registry data can be used to describe the natural history of a disease or portrait of care; to help assess clinical effectiveness, cost-effectiveness, or safety; to determine the appropriateness of use of health care products and services; or to measure quality of care.

Research institutions, academic hospitals, or individual research teams may establish patient and disease registries using public or private funds. Recently established Patient-Powered Registries or Patient-Powered Research Networks differ from traditional approaches in that the registry (or network) and the research results are managed by patients living with the particular disease and their caregivers, often through an advocacy organization or a network of organizations with input from a scientific board of advisors.[1]

Other patient and disease registries are compiled and maintained by organizations that are either mandated by legislation or have authority under specific pan-Canadian or provincial legislation to compile and maintain the registry. Patient and disease registries enacted under legislation operate under different rules than those not named in the legislation, and they are out of the scope of this report. For example, under the authority of the Statistics Act, Statistics Canada maintains the Canadian Cancer Registry, which holds data supplied by each province and territory on tumours and cancer patients diagnosed within their jurisdiction. The Manitoba Cancer Registry is operated by CancerCare Manitoba, an organization legally mandated under the province's Public Health Act to collect, classify, and maintain information on all cancer cases.

Health administrative data are information related to the provision of health care and collected by public or private health system organizations. This data includes records of hospital admissions, physician billings, drugs dispensed at a pharmacy, and diagnostic tests performed at a clinic.

Combining registry and health administrative data may provide a more fulsome and/or longitudinal view of the impact of a disease or health technology on patients and families, outcomes, and the health system, which could add to the evidence needed to support decision-making. However, given that registries and

administrative databases may not have aligned mandates, standards, or governance, linkage or “record linkage” can be challenging. As such, there is a need to understand how best to link evidence sources to provide robust information on patient demographics, disease burden, drug or other health technology use, health system use, and clinical outcomes.[4,5]

To facilitate linkage between registry and administrative data, we developed a roadmap to guide this process. We have prepared 2 documents to help registry teams in Canada maximize the benefits, address the challenges, and ensure the privacy and security of the data and informed patient consent during the record linkage process:

- A roadmap for registry teams to use when planning to link registry and administrative health services data followed by key considerations and best practices that registry teams should keep in mind when preparing to link registry data.
- An outline of what to consider when developing or revising informed consent forms is provided separately in Considerations When Developing Consent Forms for Registry Participants.

Which Registries Are the Focus of the Roadmap?

The roadmap focuses on patient and disease registries in which:

- The data are collected longitudinally in the context of the real-world health care setting.
- At least 1 component of the registry is active. That is, some of the data are collected specifically for the purpose of the registry (usually collected from the patient or care provider) rather than inferred from other sources in which the data are collected for another purpose (administrative, electronic health records, and so forth). However, it is possible that some data included in the registry may be collected from other sources.
- Data collection processes are designed to fulfill specific purposes, which are defined before collecting or aggregating the data (*purpose-driven* data collection).
- The collected information is used by clinicians for the diagnosis or management of patients (e.g., clinical data, laboratory tests, drug utilization, patient-reported outcomes).
- The information is collected in a standardized manner for every patient, and the data elements have specific and consistent definitions.
- The legal authority for creating the registry is based on patient consent. In Canada, under most privacy laws, organizations are generally required to obtain consent for the collection, use, and disclosure of personal information, although exceptions do exist.

Note: Whenever we reference specific privacy legislation, this should not be construed as legal advice. Registry teams should consult with legal counsel, privacy experts, and/or ethics boards for advice on legal requirements and specific challenges pertaining to their specific registry.

Why Link Registry Data With Administrative Health Services Data?

From the registry team perspective, there are several potential benefits to linking patient and disease registry data with administrative health services data available in Canada. For example, it can:

- improve insights into disease burden and patient and health system outcomes by combining the use of patient and health system information
- reduce duplication and minimize or eliminate the need to collect information that is already available in administrative datasets, allowing registry teams to avoid the time, costs, and other burdens associated with primary data collection, including the burden on patients
- facilitate assessment of the vital status of patients included in the registry and outcomes related to their use of health care resources, such as frequency of medical consultations, visits to the emergency department, admission to hospital, use of homecare services, and use of specialized care
- allow for the assessment of medication adherence by monitoring what prescriptions patients filled and how often they filled them.

More information on types of administrative health services data available in Canada is available in the report Drug Safety and Effectiveness Data Access.

What Are the Challenges?

Challenges that registry teams can encounter include the following:

- Information collected through a patient or disease registry is considered personal information and/or personal health information. So, it is subject to privacy legislation and ethical considerations, which may vary depending on the jurisdiction in which the data are being collected, used, and disclosed.
- The funding agreements, policies, and procedures of the institution housing the registry, and professional regulatory obligations of the registry team, may further impact how patient or disease registry data can be disclosed.
- The process for acquiring patient consent must ensure that the registry team or the supporting institution has a genuine and legitimate rationale for collecting and holding data (right to hold data).
- Substantial time and effort are required to navigate the record linkage process, especially when linking to data located in different jurisdictions.
- The registry team may need to negotiate multiple data-sharing agreements and amend and resubmit related documents to different ethics review boards and obtain additional consent from patients. One related challenge is ensuring that there are comparable identifiers in the data files that will be linked.
- There are fees associated with record linkage as well as costs related to cleaning, standardizing, and normalizing the data to prepare it for linkage. There are also potential fees for analyzing the new data received through the linkage process and for extracting data from unstructured fields or from images so it can be linked.

Could the Roadmap Be Used When Linking Other Types of Health Data?

While the roadmap was developed considering patient and disease registries, much of the content of this roadmap may equally apply to linking patient cohort data, clinical-trial data, or certain other patient-level health data systematically gathered by a researcher or clinician on a population of patients. Each dataset may have its own specific considerations and constraints, and registry teams are encouraged to engage early with Health Data Research Network (HDRN) Canada, CIHI, or Statistics Canada for planning and advice when a linkage with administrative health data is being considered.

HDRN Canada, through its Data Access Support Hub, offers navigation services for access to multijurisdictional data. Refer to Table 1 for a list of HDRN Canada member organizations that conduct record linkage and/or manage the provision of linked data for research.

Table 1: HDRN Canada Member Organizations That Conduct Record Linkage and/or Support Access to Linked Data

Jurisdiction	Organization
British Columbia	Population Data BC
Alberta	Alberta SPOR SUPPORT Unit
Saskatchewan	Saskatchewan Health Authority
Manitoba	Manitoba Centre for Health Policy (MCHP)
Ontario	ICES
Quebec	Institut de la statistique du Québec (ISQ)
Newfoundland and Labrador	Newfoundland and Labrador Centre for Health Information (NLCHI)
New Brunswick	NB Institute for Research, Data and Training (NB-IRDT)
Nova Scotia	Health Data Nova Scotia (HDNS)
Prince Edward Island	Secure Island Data Repository (SIDR)
Northwest Territories	NWT Bureau of Statistics
Yukon	Yukon University
pan-Canadian	Canadian Institute for Health Information (CIHI)
pan-Canadian	Statistics Canada

HDRN = Health Data Research Network.

Roadmap for Linking Registry Data With Administrative Health Services Data

This roadmap is intended to help registry teams navigate the process of linking patient or disease registry data with administrative health services data (“record linkage”). It is a Canada’s Drug Agency-led initiative and is part of work being done to support the Government of Canada’s National Strategy for Drugs for Rare

Diseases to support efforts for improved evidence generation and the value of real-world data from rare disease registries.

Characteristics of patient and disease registries are shaped by the different stakeholders involved and the specific contexts in which they were established. For the purpose of decision-making, both primary and secondary stakeholders may want registry data to be linked to an administrative health services dataset or registry data that has already been linked with an external data source to support their information needs. In this roadmap, we refer to primary and secondary stakeholders as follows:

- Primary stakeholders: Those involved from the outset and responsible for creating and organizing funding for the registry.
- Secondary stakeholders: Those who will use the knowledge gained from the data collected in the registry or will be potentially affected by the results but who did not play a critical role in the registry setup.[2] Secondary stakeholders may be identified early in the registry planning phase or may emerge only after collected data.

Examples of health system stakeholders that may play a primary or secondary role in a specific registry include patients/caregivers, patient advocacy groups, care providers, researchers, public health units or agencies, regulatory or health technology assessment agencies, health technology manufacturers, public or private payers, health policy advisors or decision-makers, academic institutions or consortia, and funding agencies.

This roadmap can be used to:

- assess whether record linkage can occur
- understand how to move forward with a request for data linkage.

An outline of the roadmap is provided in Figure 1. There are 2 different starting points for the roadmap depending on the current stage of a registry:

- Registries in the planning stage should start at step 1.
- Teams managing existing registries, irrespective of whether they have been operating under the auspices of a research ethics board, should start at step 3.

Step 1: Articulate the Registry Purpose

The first step in planning a registry is to articulate its purpose. A clearly articulated purpose will help you define the objectives, provide the rationale for data collection, and clarify the need for certain data.

If your registry is in the planning stage, try to consider all possible future uses of the registry data and the associated costs and burden of data collection,[7] keeping in mind that new information needs may arise that require the collection of additional data.

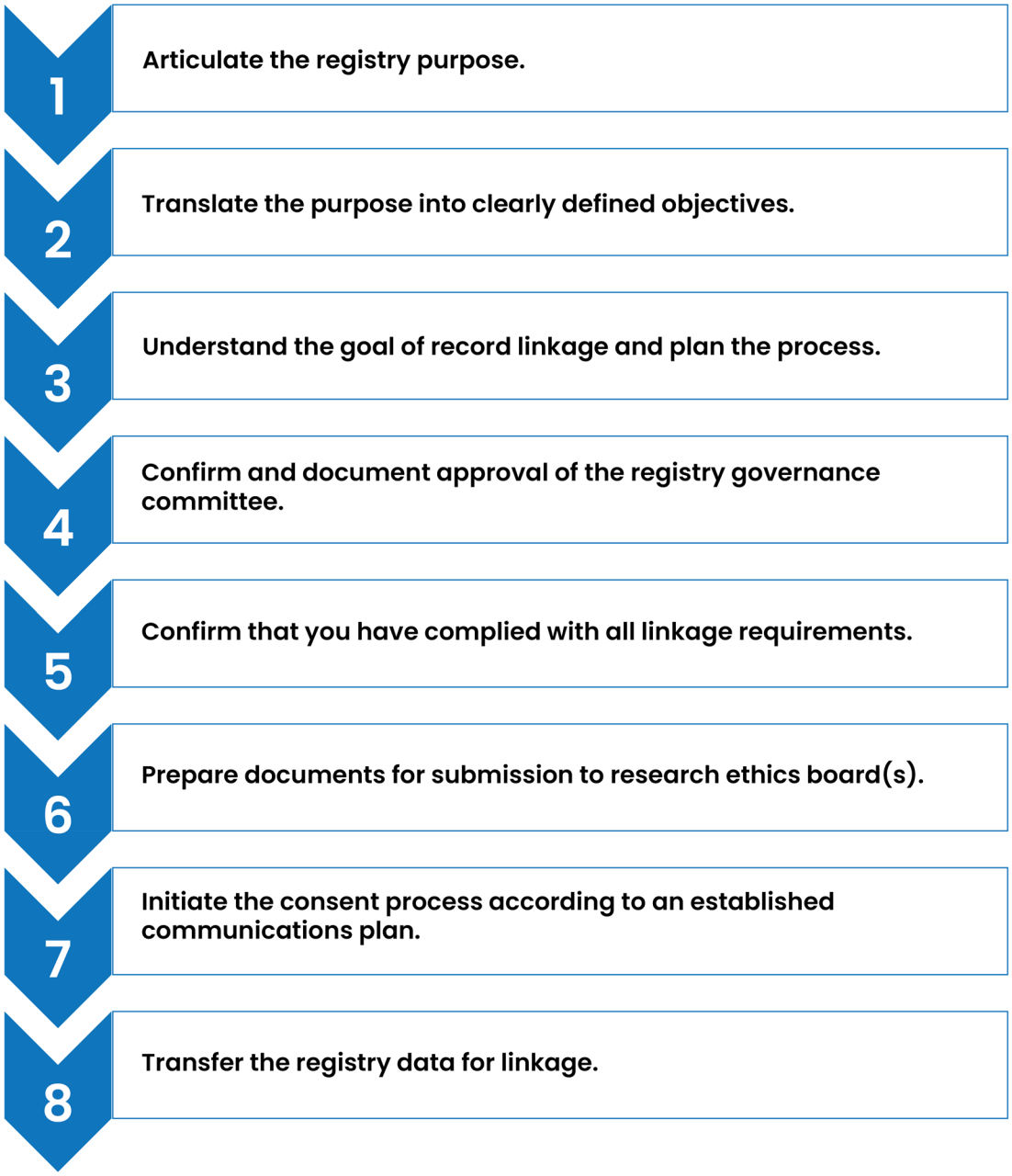
From the outset, establish a process for connecting with registry participants so that:

- additional data can be collected
- patients can participate in clinical trials

- patients can sign new consent forms (“reconsent”), and opt out of future studies
- patient engagement can be maintained.

This process is particularly relevant for rare disease registries with long-term follow-up, and for registries that may enrol participants as children who will transition to adult care and, therefore, move from giving assent

Figure 1: Roadmap for Linking Registry Data With Administrative Health Services Data



through adult family members to providing individual consent. Similarly, adults who are temporarily unable to make decisions for themselves may be able to consent once their health improves.

If you know from the outset that the registry data will be linked with administrative health services data, then you should describe the purpose for record linkage, the organization linking the data, and how often record linkage will occur in the registry protocol, the consent and assent forms, and any other materials you share with registry participants.

Step 2: Translate the Purpose Into Clearly Defined Objectives

You should translate the registry's purpose into clearly defined objectives so that you can identify the core dataset and clarify that it fulfills those objectives.

If your registry is in the planning phase, once data needs are identified, you must determine:

- Where you will store the data.
- Where you will capture the patient-identifying information:
  - in a central (web-based) registry database?
  - at registry recruiting sites?

To prepare for capturing and storing the data, you should consider:

- the software system you will use to capture the data and where you will store it (e.g., in the cloud, in a separate server behind firewalls at 1 of the participating sites, with desktop systems equipped with antivirus software, servers running the most recent security patches, and so forth)
- whether you can keep the patient-identifying information separately from the registry data in an encrypted format
- whether you can restrict access to the identifying information to only those with a need to know (defining access levels)
- whether the security and privacy policies and procedures of the organization managing the registry database are adequate.

Step 3: Understand the Goal of Record Linkage and Plan the Process

You should understand the goal of record linkage and plan what you will need to be able to carry out the process.

If you are managing a registry that is already in operation, you will need to consider whether a request for linkage to administrative health services data is permitted under your governance model. Even if it is permitted, you will need to consider whether the request is also permitted under the current ethics-approved protocol or whether you must seek approval for a new use of the registry data.

Whether planning a new registry, or managing an existing registry, you must have a good understanding of the process of linking registry data to administrative health services data and know the answers to the questions posed in the following sections.

What Is the Purpose of the Data Linkage?

In the registry protocol and any materials shared with participants whose data are held in the registry, clearly document the purpose for linking registry data with administrative health services data.

Which Data Elements From the Registry Will Be Linked With Administrative Health Services Data?

Consult with the organization linking the data to confirm what linking variables they require (e.g., first name, last name, date of birth and sex, health card number, date of diagnosis, date of hospital admission). The health card numbers and province of residence may be acceptable and sufficient for linkage to administrative health services datasets. However, 100% completeness and 100% accuracy of a direct identifier across all participants' records is rarely achieved. Additional information (e.g., patient name and date of birth) may be required to link some participants' data. It is also possible for no match to be found for a complete record during the linkage process due to errors in the captured information. For these reasons, it is often useful for you to provide additional participant identifiers for record linkage.

To ensure that the maximum number of registry records can be linked with the administrative health services dataset(s), it is important to collect, from the outset, the linking variables for all registry participants in a standardized way. For instance, you should always use:

- the legal/official names of registry participants
- the same format for collecting dates
- the same number of digits for health card numbers ensuring that the province or territory of residence is specified.

If you are managing an existing registry, assess the completeness and quality of all the identifying information that will be used for linkage and consider the feasibility of collecting the required information.

Errors in documented health care numbers might not be noticed until the record linkage stage. So, it is recommended that you include when necessary additional identifying variables such as first name, last name, sex, and date of birth for the linkage process.

To Which Party(ies) Will the Data Be Disclosed?

The organization linking the records may have requirements around disclosing data. For instance, organizations like Statistics Canada may fall under specific legislation that warrants that requirements be met before data can be disclosed to them and before they can link registry data.[8]

In some jurisdictions, approval is required by the appropriate Ministry of Health or equivalent for certain disclosures of personal health information. In addition, many jurisdictions require a separation of duties where the organization that links the records is different from the organization that is ultimately responsible

for the management and provision of linked data. When a third party is linking the records, you will need to prepare and send:

- a file containing only the identifying information (e.g., health card number) and the encrypted code or study ID in the registry database to the third party who will create the crosswalk file (once created, the crosswalk file will be shared with the party conducting the linkage)
- a separate registry dataset(s) that contains the clinical data and the encrypted code or study ID to the party conducting the linkage (i.e., the holders of the administrative health services data).

Refer to Table 1 for a list of HDRN Canada member organizations that conduct record linkage and/or manage the provision of linked data for research.

How Often Will We Need to Disclose Data for Linkage?

You must determine if data needs to be disclosed 1 time only or at ongoing regular intervals.

What Administrative Health Services Data Elements Must Be Linked With the Registry Data?

Choose data elements from the administrative health services dataset based on the purpose and specific objectives.

Before starting the process of record linkage, the administrative health services data holder must confirm that the data of interest can be linked and subsequently shared, or can be linked and analyzed before being shared, if necessary.

There are risks associated with disclosing registry data to the organization linking the records, such as:

- interception of personal information during the transfer of data
- inappropriate handling of the identifying information by the recipient organization
- inappropriate use of the disclosed data by the recipient organization for purposes that were not part of the agreement
- inappropriate storage of the disclosed identifying information by the recipient organization.

The recipient organization must therefore have policies and procedures around receiving and handling personal identifying information that minimize the risks and mitigate any negative consequences in the event of a breach. You must review the policies and procedures of the recipient organization before deciding whether to disclose registry data for record linkage.

Step 4: Confirm and Document Approval of the Registry Governance Committee

Prior to proceeding with a request for record linkage, ensure you have the support of the primary registry stakeholders and documented approval from the registry's governance committee.

Step 5: Confirm That You Have Complied With All Linkage Requirements

To ensure that you can meet or already have met all technical, operational, and legal requirements to be able to disclose the registry data for linkage with administrative health services datasets, consult with your legal counsel /privacy expert /ethics board (academic or independent, as applicable).

If your registry is in the planning phase:

Consider the legal or ethical implications related to the planned registry software platform and participant consent.

A research ethics board may have the authority to waive the need for participant consent if certain conditions are satisfied. Please refer to Considerations When Developing Consent Forms for Registry Participants for more information.

If your registry is recruiting participants from multiple clinics or hospitals, you may be able to use a central research ethics board to minimize the need to make several submissions to different research ethics boards. Often, when recruiting participants from an academic site/hospital, the research ethics board at that specific site will need to review the registry protocol, unless there is a reciprocal agreement indicating that another academic research ethics board can take responsibility for the review process.[9]

If you are managing an existing registry:

Consider the need for and feasibility of obtaining reconsent of all registry participants for data linkage. When seeking reconsent, you should consider the potential effects of selection bias and the implications for external validity. Reconsented participants may be systematically different from non-reconsented participants.[7]

Therefore, you should carefully consider any decision to change the informed consent form and obtain the reconsent of participants. When the initial consent does not cover the proposed linkage to administrative health services data, you may seek a waiver of reconsent requirements from your ethics board(s).[7] Please refer to Considerations When Developing Consent Forms for Registry Participants for more information.

It is best practice to reconsent patients whenever you make significant changes to privacy practices, including new uses of data or disclosures to new stakeholders. You may also need to obtain reconsent if, as mentioned in step 1, the participants were below the age of consent when initially enrolled (i.e., if they only provided assent) but reach the age of majority when the registry is still active, or if participants were temporarily incapacitated and have subsequently recovered.[7] In addition, participants may no longer be able to consent (e.g., those enrolled in dementia studies) and may need someone to consent on their behalf.

Step 6: Prepare Documents for Submission (or Resubmission) to Research Ethics Board(s)

If your registry is in the planning phase, it is recommended that you pilot test the informed consent form, or assent form for children (if applicable), with a sample of the registry participants to ensure it is clear and understandable. If your registry includes long-term follow-up, it is advisable that you request permission

to contact a family member, or other alternate contact, in case you cannot reach the participant for any business related to participation in the registry or if other research opportunities arise.

It is also good practice to prepare a consent audit and documentation management plan to ensure that consent can be verified, monitored, and updated over time.

If you are managing an existing registry, if documents need to be revised and resubmitted to ethics review board(s), you should consider the impact of implementing those changes, including the effort of reconsenting or notifying participants. Changing documents may involve:

- revising the registry protocol
- revising the informed consent form
- pilot testing the new consent form before submitting it to ethics board(s)
- preparing and implementing a communications plan to inform registry participants of the new activity
- completing and submitting ethics forms at each of the sites (if not using a central research ethics board) and communicating with ethics boards during the review process (applicable to registries operating under the auspices of a research ethics board)
- effort with the reconsenting process
- updating any existing registry policies and procedures pertaining to data access.

The costs of implementing changes can be substantial.

Note: It is recommended, and may be a requirement, that the informed consent form be reviewed by the organization conducting the record linkage before it is finalized and administered to registry participants.

Step 7: Initiate the Consent Process According to an Established Communications Plan

Once the research ethics board(s) approves the registry documents (or revised documents), initiate the consent process according to an established communications plan. To engage with registry participants, you can use communication tools such as a registry website, newsletters, and posters or pamphlets in clinic waiting rooms.

Consider designating someone to be on-call during the consent process to address questions from participants.

If consent is not required (e.g., if consent is waived by a research ethics board or falls under 1 of the legal exceptions), it is still good practice, and the research ethics board may require it, to notify participants of new uses of their data and how it may impact future management of the disease.

For information about the considerations for waiver of consent, registry teams can consult the Tri-Council Policy Statement (TCPS-2) or the research ethics board from which they seek approval.

Case Scenario 1 – Disease Registry That Requires Reconsent Before Record Linkage Can Occur

A registry team that manages a disease registry including about 5,000 patients across Canada that has existed for 40 years is approached by a researcher for access to linked data to study the economic burden of the disease. The registry team discusses the request and approves it at the registry governance committee meeting. In the past, the registry data had been linked in 1 province (BC) and required that participants in that province sign new consent forms. The privacy expert at the pan-Canadian coordinating centre confirms that a new consent form will need to be administered to the remaining registry participants in the other provinces if linkage is required across the country. The researcher approaches HDRN Canada through its Data Access Support Hub (DASH) unit to initiate a multijurisdictional data access request. DASH assesses the consent form and determines that there are some important terms missing in the informed consent form as per HDRN Canada consent for linkage guidelines. The researcher works with the registry team to update the consent and assent documents as per HDRN consent for linkage guidelines. The registry team advises the researcher to apply for a grant through the disease foundation to support the reconsent process, development of necessary data-sharing agreements, record linkage fees, and cost of analyses.

Case Scenario 2 – Rare Disease Registry Without Consent for Record Linkage in Place

A registry team that manages a rare disease registry, which is sponsored by industry and has only 17 patients in Canada, is approached with a request for information from participants using an off-label product. The sponsor is considering the possibility of expanding that product's indication. As well as the clinical data available in the registry, the sponsor wants information on health care resource utilization, which requires registry data to be linked with administrative data. The registry team consults with its governance committee, which approves the request. The privacy officer of the organization housing the registry is consulted and determines that the signed consent form did not contemplate record linkage. Therefore, the informed consent/assent form will need to be updated together with the registry protocol and then approved by an ethics review board. All registry participants will need to sign the updated informed consent or assent form before their data can be disclosed for linkage. The registry team contacts Statistics Canada to initiate the record linkage request. Statistics Canada reviews the revised protocol and consent form and recommends that certain language pertaining to the record linkage process be included in the protocol and consent/assent forms. The registry team initiates amendment of the registry protocol and consent/assent forms and submits them to a private research ethics board. Approval is granted, and Statistics Canada initiates a service agreement that governs the transfer of the registry data, record linkage, and the preparation of the linked registry dataset.

Registries Operating Without Oversight of a Research Ethics Board

If your registry is not managed by an institution or by a researcher that receives funding from 1 of the Tri-Council agencies, or whose main goal is not research, you might not have pursued ethics review at registry inception. Examples include registries established under a Public Health Act, cancer registries prescribed under certain legislation (note that registries prescribed under legislation are out of scope for this report), or registries created by care providers of a certain condition. However, you still need to comply with legislation in the jurisdiction where they operate. In Canada, under most privacy laws, organizations are generally

required to obtain consent for the collection, use, and disclosure of personal information and personal health information, although exceptions exist. Please refer to Considerations When Developing Consent Forms for Registry Participants for more information.

If your registry is not operating under the auspices of a research ethics board, it is still good practice for you to have your documents reviewed by a privacy expert to ensure they meet the legal requirement of consent. Most organizations that conduct record linkage are permitted to do so under the auspices of a research project that has been reviewed and approved by a research ethics board. Therefore, if you have been operating outside the realm of research, it is recommended that you consult with the organization conducting record linkage or the organization responsible for providing access to linked data to confirm that their policies permit linking with registry data.

To ensure that your registry is complying with the highest ethical standards, you may consider pursuing a review of its operations by an independent ethics review board and implementing any recommendations arising from that review. This will show your commitment to operating under the highest ethical standards and build the trust of the registry participants. In addition, it will facilitate record linkage in most jurisdictions.

Step 8: Transfer the Registry Data for Linkage

When transferring data electronically, it is good practice, and often required by the organization receiving the data, to use a secure channel such as a managed transfer file (MTF) software and to encrypt the identifying information. Organizations that conduct record linkage typically offer a platform where data can be uploaded securely. If your registry is housed at a hospital or large research organization, your IT department may want to know what security measures the organization receiving the data has in place before the registry data is disclosed. Examples of the type of information that the IT departments would want to know include:

- the cybersecurity framework being used
- whether the receiving organization has undergone a privacy/security assessment of their infrastructure by a third party
- how role-based accounts and access are implemented
- what encryption levels exist
- whether the MTF solution has the capability to audit administrative activities.

When possible, you should send identifying information (i.e., the crosswalk file with identifying information and the respective encrypted code or study ID) separately from the registry clinical information and study ID. Record linkage should be conducted by trained individuals who will also de-identify, code, or anonymize the data before making it available to the approved data users.

If the registry data required for linkage is to be transferred to an international registry or to be used by a government agency in another country, you should ensure that you understand the legal and ethical framework of the jurisdiction where the recipient entity operates and that there is alignment and compliance with the 3 core principles of the TCPS-2: respect for persons, concern for welfare, and justice.[9]

Refer to the next section — Key Considerations and Best Practices When Linking Registry Data — for details on what registry teams should ask themselves when preparing for record linkage.

Key Considerations and Best Practices When Linking Registry Data for the Purpose of Decision-Making

Key Questions

In this section, we present key questions that registry teams should ask themselves when linking registry data with administrative health services data. The relevance of each question to record linkage is presented and related key considerations and/or best practices are highlighted. Due to the specific challenges of rare disease registries, where warranted, information pertaining to rare disease registries is emphasized. The questions are as follows:

- Why do you need a registry?[a]
- What consent do you need to obtain?
- Could the information you collect be used to identify participants?
- How will you ensure that data entries are accurately linked?[a]
- How will you protect information so it cannot be used to identify participants?
- How will you protect fully identifiable data or information?
- How will you protect de-identified data?
- How will you protect coded data?
- What do I need to know about sharing anonymized data?
- What do I need to know about sharing aggregate data?[a]
- Are there specific legal obligations in the jurisdiction where your registry operates?[a]
- What governance roles are needed to manage a data registry?

[a] These subsections include information specific to rare disease registries.

Why Do You Need a Registry?

Why This Matters

The intended use or purpose of data collection will affect whether and how the data may be collected, accessed, used, and disclosed. If the primary reason for creating a registry is to conduct research, you will usually need to obtain ethics approval and informed consent. As new questions posed by registry stakeholders emerge, you will need to update data collection and ethics approval.

Why This Matters for Rare Disease Registries

Often rare disease registries have developed organically based on the need to understand the number of people affected, their geographical distribution, and basic demographic and clinical characteristics of the disease.[7] The scope of these registries may evolve over time as more knowledge about the disease, biomarkers, new therapies, and disease-specific patient-reported outcome measures become available.

Considerations

What Questions, Problems, or Decisions Is the Registry Intended to Help Address?

The registry should be an appropriate means to fulfill the purpose.

Are the Data Already Available?

Consider whether the data already exist elsewhere (e.g., as part of electronic health records or administrative health services data) and if so, if they are of sufficient quality to answer the purpose and are accessible. If the required data have not been sufficiently collected or are not accessible for the desired purpose, it is appropriate for you to consider creating a new registry.[10]

Might There Be Future Uses of Your Registry Data?

If you make any updates to the protocol, you will need to get approval from a research ethics board and likely reconsent all patients (unless consent is waived), which can be costly. To avoid the administrative burden of resubmitting the protocol to ethics boards, try to anticipate and plan some future uses of your registry data. Consider clarifying with your stakeholders what questions your registry may be able to address.

Although informed consent for registry research that allows broad data sharing is optimal for promoting innovation, there may be specific disease registries (e.g., HIV/AIDS research) where more specific consent may be more appropriate.

Best Practice

Establish a Process for Connecting With Registry Participants in the Future.

When planning a registry, keep in mind that as time passes new questions may arise that may require you to collect additional data from participants. So, at the onset, you should establish a process to connect with registry participants to facilitate additional data collection, involve patients in trials, obtain reconsent, or keep patients engaged.

This is particularly relevant if your registry has a long-term follow-up and if the participant is unable to speak for themselves, or if your registry enrolls participants as children, who may move from assent to consent, or who may move or transition to the care of adult specialists.

If you are planning a rare disease registry, you should establish a process for connecting with registry participants in the future to allow for long-term follow-up.

What Consent Do You Need to Obtain?

Why This Matters

In Canada, the collection, use, and disclosure of personal identifying information requires consent, unless exceptions apply. The law also regulates the decision-making capacity of the individual. There are requirements in the legislation for consent to be valid.

The registry team has an ongoing ethical and legal obligation to bring to participants' attention any changes to the research project that may affect them. These changes may have ethical implications, may be germane to their decision to continue research participation, or may be relevant to the particular circumstances of individual participants. Please refer to Considerations When Developing Consent Forms for Registry Participants for more information.

There are special instances when a research ethics board may grant the registry team a waiver of consent — for instance, for research conducted in specific types of emergency situations or if the data are collected under a public health act.

You may be required to reconsent participants when there are changes in the scope of the registry, substantive changes to the protocol, addition of procedures not previously addressed in the consent, changes in reasonably foreseeable risks or potential benefit, changes in data sharing or reporting procedures, or identified errors or omissions in the original consent document. Such revisions must be reviewed and approved by a research ethics board before you use the revised consent unless it is needed to eliminate apparent immediate hazards to subjects. Reconsent may also be necessary if the participants were below the age of consent when initially enrolled but reach the age of majority when the registry is still active.[7]

Consideration

Does Each Participant Have the Capacity to Provide Consent?

The ability to assess capacity poses a challenge when consenting children and youth, older individuals, and those with mental health conditions, and people living with disabilities (permanent or temporary). Capacity is the ability to understand information related to the decision to consent to the collection, use, or disclosure of personal information and to appreciate the reasonably foreseeable consequences of giving, not giving, or withdrawing consent.

The age of consent varies by jurisdiction. For instance, in Ontario, a capable person may provide consent regardless of age. This means if a child is capable, the child may consent. If a child's decision conflicts with a parent's (even if the child is under 16 years of age), the child's decision prevails. However, the Office of the Privacy Commissioner of Canada takes the position that in all but exceptional circumstances, a child under 13 years of age is unable to provide meaningful consent.[11]

Best Practices

If You Make Significant Changes to Your Privacy Practices, Reconsent the Participants.

It is best practice for you to reconsent patients when making significant changes to privacy practices, including new uses of data or disclosures to new stakeholders. A research ethics board may waive the

need to reconsent if certain conditions are satisfied as per Article 5.5A of the Tri-Council Policy Statement 2(2022).[9] Please refer to Considerations When Developing Consent Forms for Registry Participants for more information on meaningful consent.

Give Special Consideration to Certain Elements of Informed Consent.

Elements of informed consent that should be given special consideration include the scope of the use of registry data, potential for recontact, withdrawal, and information regarding the electronic data security and management to be employed.

Recontact: At the outset, you should establish whether and how registry participants will be recontacted and you should include this information in the consent form.

Withdrawal of consent: You should provide instructions on how registry participants can withdraw from the registry at any time. You should also inform registry participants that it is impracticable, if not impossible, to withdraw results once they have been anonymized, or combined into aggregated results, published, or otherwise disseminated.

Electronic data security: Given the public concerns about electronic data security, you should clearly inform registry participants of where their data will be stored geographically and the physical security of their data and/or biospecimens, including methods of coding and removal of identifiers, encryption techniques, potential for cloud computing, and quality assurance policies. As well, participants should be informed about the process of releasing and transferring data to future registry teams as it relates to maintaining confidentiality and data protection.

Develop Jurisdictionally Specific Informed Consent Forms.

If you want to link your data with administrative health services datasets in most provinces and territories or nationally, it is suggested that you develop an informed consent form specific to each jurisdiction, including details about the organization within the jurisdiction linking the data. If the organization linking the data is a pan-Canadian organization, then details specific to that organization should be included. In addition, it is good practice to have the consent form reviewed by the organization linking the data to ensure the relevant detail is included, before the consent form is finalized.

Use Targeted Communication to Let Patients Know What Their Data Will Be Used For.

When educating the patients about what their data will be used for, consider leveraging the registry website, posters and pamphlets in doctors' offices, privacy ambassadors, contact hotlines, letters, and other forms of formal, targeted communication.

Could the Information You Collect Be Used to Identify Participants?

Why This Matters

The information collected for patient and disease registries may include direct identifiers (e.g., an individual's name, health card number) or indirect identifiers that, when taken as a whole or when combined with other available information, can be used to identify an individual.[12] Privacy risks in research relate to the identifiability or re-identification of participants and the potential harms they, or groups to which they belong,

may experience from the collection, use, and disclosure of personal information. Privacy risks arise at all stages of the research life cycle, including during the initial collection of information, use and analysis of the information to address research questions, dissemination of findings, storage and retention of information, and disposal of records or devices on which information is stored.[9]

Health-related data that contains identifying information is subject to privacy legislation. Please refer to Considerations When Developing Consent Forms for Registry Participants for more information.

Best Practices

Make Sure Your Registry Collects the Direct Identifiers Needed for Data Linkage.

When planning to link registry data with administrative health services datasets, ensure your registry collects direct personal identifiers (e.g., health card numbers) that are required to link with administrative health services datasets. In addition, to maximize linkage, ensure the collection is complete (across all participants), its format is consistent (e.g., same format, same number of digits, same length), and as much as possible free of manual entry errors.

Keep Direct Identifiers in a Separate and Secure Table Within the Registry Database or Externally.

Keeping direct identifiers separate and secure will require you to create an encrypted code or study ID to keep with the registry clinical data. The encrypted code or study ID and the registry clinical data can then be used for research studies. In addition, when transferring data for record linkage, you can generate a crosswalk file linking the encrypted code or study ID data with direct identifiers, which you can send separately from the registry clinical data to the organization conducting record linkage.

Comply With the Legal and Regulatory Requirements in Your Area.

In addition to following the guidance in the Tri-Council Policy Statement, you should be aware of legislation and regulatory requirements in the area where you operate. These requirements may vary by jurisdiction and, depending on who is funding or conducting the research, may include obligations under the Constitution (including the Canadian Charter of Rights and Freedoms), and pan-Canadian or provincial privacy legislation, among other legal and regulatory requirements.

You should consult with your legal counsel, privacy experts, and ethics boards to ensure you comply with legal and ethical requirements, and with your organizational policies and procedures.

How Will You Ensure That Data Entries Are Accurately Linked?

Why This Matters

It can be challenging to link records if there is a lack of consistent and shared variables or codes for you to use to match records across the datasets. In other words, it is helpful to have a clear, shared “key” so you can align the data entries from 1 dataset with those of another. If you can link with external data sources, you will need to collect less information from patients, thereby reducing the burden on the patient.

Why This Matters for Rare Disease Registries

This is particularly important if you are managing a rare disease registry since these tend to span the life of a patient. If a patient is lost to follow-up due to a move to attend school in another city for instance, outcomes like death and health care utilization that are captured in administrative health services data could be obtained through record linkage.[7]

If record linkage involves identifiable information, or where record linkage can foreseeably or potentially produce identifiable information, you are required to satisfy the conditions of your research ethics board. For example, you must ensure that:

- the record linkage is essential to the research; and
- you will implement appropriate security measures to safeguard information.[9]

Without common identifiers, it becomes challenging, if not impossible, to accurately link or merge datasets. This can result in missed matches, inaccurate data linkage, or even the introduction of errors. Ensuring proper linkage is vital for deriving accurate insights and drawing valid conclusions from the combined dataset.

Consideration

Should You Collect Unique Direct Identifiers to Facilitate Record Linkage?

If 1 of the goals is to generate evidence by combining information in a registry with external data sources (e.g., cancer registry, hospital data, and so forth), you should consider collecting unique direct identifiers such as the patient health card number and issuing province to facilitate record linkage.

Best Practices

Ensure the Variables Used for Linkage Are Complete and Consistent.

To maximize record linkage, you should ensure the variable(s) used for linkage is complete (collected for all participants in the registry) and consistent (collected the same way, same format, same length).

Keep the Direct Identifiers in a Separate and Secure Table and Create an Encrypted Code or Study ID That Is Maintained With the Registry Clinical Data.

This will make it easy for you to separately send a crosswalk file linking direct identifiers to the encrypted code or study ID to the organization conducting record linkage.

Know the Legislation That Applies to the Organization Holding the External Data Sources.

Certain organizations that hold data holdings for the province or for Canada may fall under specific legislation that warrant requirements to be in place before they can receive and link your registry data.[8] Therefore, it is crucial that you consult with your legal counsel or the privacy expert of the organization chosen to do the linkage to ensure you understand the requirements and use the right language in both the registry protocol and consent form. If an amendment is required, it is recommended that you get sign-off

from legal counsel or the privacy expert of the organization doing the record linkage before submitting the changes to the research ethics board.

HDRN Canada has developed guidelines for informed consent for record linkage.[13]

Inform Registry Participants of the Data Linkage and Potential Risks.

If consent is waived, it is still best practice to inform registry participants of this activity and the potential risks. You can do this by providing information on the registry website or through a newsletter. Having a "hotline" and designated contact person to address any concerns participants may have with respect to record linkage is also recommended.[10]

How Will You Protect Information so It Cannot Be Used to Identify Participants?

How you protect information will differ depending on the type of information and is described in the following paragraphs.

How Will You Protect Fully Identifiable Data or Information?

Why This Matters

Under Canadian privacy laws, fully identifiable data or information (also known as personally identifying information) must be collected, used, and disclosed with the consent of the subject it refers to, and it can only be used for the purpose for which it was collected.[14] Examples of personally identifying information include name, address, telephone number, email address, and social insurance number or other identifying number or code, which identifies an individual directly, as well as gender, race, birth date, and geographic indicator, which can be used in combination with other data elements to indirectly identify individuals. Table 2 provides an example of a fully identified dataset.

Safeguarding fully identifiable information is a critical responsibility that must be always taken seriously.

Table 2: Fully Identified Dataset

Name	Address	Postal code	Date of birth	Gender	Salary ($)
Sally Ford	123 City Ave, Vancouver, BC	M5M 1P2	1970-01-02	Female	89,000
Sam Smith	4567 Town Rd, Smalltown, BC	M5B 1A5	1980-03-03	Male	67,000

Best Practices

- Implement policies and procedures around safeguarding the identifying information collected from registry participants.
- Ensure that all staff are trained in privacy and follow the policies and procedures.
- When transferring fully identifying information to a third party, assess risks that could adversely impact the protection of that information.

- Make sure that the party to which you intend to disclose the information will provide a sufficient level of protection for the identifying information.

You must, through contractual or other means, ensure that this third party will protect the personal information at a level that is comparable to that required in legislation and will limit its use of the personal information to the purposes specified to fulfill the contract.

How Will You Protect De-Identified Data?

Why This Matters

The deidentification process removes any information that could identify an individual or for which there is a reasonable expectation that the information could be used, either alone or with other information, to identify an individual.[15] (This is the definition under the Ontario Privacy legislation. Other jurisdictions may have slightly different definitions.) However, de-identified data exists in the spectrum between fully identified data (data with direct identifiers) and anonymized data (data that cannot be traced back to the individual).

In a de-identified dataset, direct identifiers are transformed (either through encryption or replaced by pseudonyms) and/or other quasi-identifiers (date of birth, age, postal code, and so forth) are suppressed or generalized to minimize the risk of re-identification. For example, date of birth could be presented as YYYY as opposed to its full date representation (DDMMMYYYY). Table 3 provides an example of de-identified data.

However, when de-identified information is combined with other types of personal information, there may be enough similar data elements between the 2 sets of information to allow a user to re-identify the subject.[12] De-identified information, therefore, has some residual risk of re-identification. In addition, while de-identification techniques help to protect against the disclosure of individuals' identifiers, they do not protect against other risks, including the disclosure of stigmatizing group attributes.[15]

Table 3: De-identified Dataset

ID	FSA	Year of birth	Gender	Salary range ($)
TS01	M5M	1970	Female	80,000-99,999
TS02	M5B	1980	Male	60,000-79,999

FSA = forward sortation area, which represents a geographic unit based on the first 3 characters in a Canadian postal code.

An ID was assigned, some direct identifiers were removed, and certain fields were generalized such that it is harder to re-identify the individuals without having the crosswalk file to link the ID to the direct identifiers.

Best Practices

- Consult with legal counsel, a privacy expert, or an ethics board to ensure you are compliant with legal and ethical requirements and with organizational policies and procedures.
- Before disclosing de-identified data, consider the risk of re-identification that could occur when releasing the dataset and have measures in place to remedy a potential breach if it is identified.
- If the records are to be linked and the identifiers in the registry were encrypted or kept separately from the rest of the information (as is good practice), supply the organization doing the linking with a crosswalk file between the identifiers and the encrypted code or study ID.

How Will You Protect Coded Data?

Why This Matters

Coded data have unique codes that link them to individual subjects' identifiers (e.g., name, medical record number, email address, or telephone number). Generally, data are collected with a Study ID, and a linkage file is maintained where the Study ID is associated with the subjects' identifiers.[16] Table 4 provides an example of coded data.

Table 4: Coded Dataset

ID	Postal code	Date of birth	Gender	Salary ($)
TS01	M5M 1P2	1970-01-02	Female	89,000
TS02	M5B 1A5	1980-03-03	Male	67,000

An ID was assigned, and direct identifiers were removed.

For security reasons, you might keep identifying information about a participant (direct identifiers) isolated from other data collected about that participant. If so, you would assign a code to the record and keep a crosswalk file connecting the codes and identifiers locally at the site/hospital where the patients are enrolled and followed.

Considerations

- Is there other available information or a key (crosswalk file) that would enable re-identification of the coded data?
- What protections could be applied to prevent or limit re-identification?
- Will you supply the crosswalk file to the organization doing the linkage?

  If the interest is record linkage, you will need to supply the crosswalk file to the organization doing the linkage with administrative health services data.

Best Practices

- When disclosing coded data to a researcher, ensure the existing key code is accessible only to a custodian or trusted third party who is independent of the researcher.[9]
- Do not use a code that contains information related to the participant.

  The code should not be a combination of information related to the individual, such as initials, date of birth, and so forth. It can be sequential numbers and/or letters, such as ST01, ST02, ST03, and so on.[16]
- If you keep identifiers at a local site, consider using Common Data Elements (CDE) when collecting direct identifiers in case there is ever a need to link data from multiple sites.

What Do I Need to Know About Sharing Anonymized Data?

Why This Matters

There are currently no legal requirements around sharing anonymized data.

In anonymized data, direct and in some instances indirect identifiers have been removed. This results in an irreversible and permanently modified record from which no individual can be identified, even when used in combination with other data.[17] Table 5 provides an example of anonymized data.

Table 5: Anonymized Dataset

Forward sortation area	Gender	Salary ($)
M5M	Female	90,000
M5B	Male	65,000

Direct identifiers and date of birth (indirect identifier) were removed; there is no code to allow for re-identification of the individuals.

By dropping the identifiers, an irreversible record is created. Unless there is only 1 male making $65,000 at that postal forward sortation area, it is not possible to identify the individual.

Best Practice

- If data/samples have been made anonymous, inform registry participants that they are unable to withdraw consent.[7]

What Do I Need to Know About Sharing Aggregate Data?

Why This Matters

Aggregate data are information that has been collected from multiple sources and/or measures, variables, or individuals and compiled into data summaries for the purpose of public reporting or statistical analysis.[18]

> **Why This Matters for Rare Disease Registries**
>
> For rare disease registries, making aggregate data publicly available could potentially stigmatize certain patient populations, depending on the sensitivity of the information.

Best Practice

- Establish policies and procedures for data access and data sharing.

  From the onset of your registry, you should establish policies and procedures around data access and data sharing to facilitate decision-making. You should also consider the level to which the data will be aggregated (i.e., some organizations do not report results based on fewer than 6 individuals) as well as who will be notified about upcoming publications, including presentations, posters, media releases.[7]

Are There Specific Legal Obligations in the Jurisdiction Where Your Registry Operates?

Why This Matters

In Canada, depending on the jurisdiction (province, territory, or Canada-wide) in which data are being collected, used, or disclosed certain legal requirements may apply.[19]

Why This Matters for Rare Disease Registries

Rare disease registries tend to encompass the widest geographic scope because no single institution, and in many cases no single jurisdiction, has enough patients to conduct adequate clinical and translational research. If you are managing a rare disease registry, collaboration and maximum use of limited resources are crucial.[20]

Currently, provincial data centres only allow for a federated data access model, which will likely not fulfill your needs if you are managing a small registry with only a few individuals in each jurisdiction. If you are managing a rare disease registry and wanting to link your data with administrative health services datasets, you should approach 1 of the 2 organizations that hold pan-Canadian administrative health services data (i.e., Canadian Institute for Health Information and Statistics Canada), keeping in mind that there must be clear legislative authority for you to disclose your linkage variables to the organization and for the organization to return linked personal health information to you (if that is required).

Best Practice

- Before sending data to a secondary international registry, find out how they will safeguard the data.

  When you are planning to send data to a secondary international registry to increase the pool of information available about a particular disease, you should request documentation of all the data safeguards of the secondary registry from the time the data are being used to long-term storage.

What Governance Roles Are Needed to Manage a Data Registry?

Why This Matters

It is best practice to establish a governance structure to inform the purpose of the registry and to oversee the registry and how it will operate.[10] Governance is a formalized structure or plan for managing a registry and guiding decision-making related to registry funding, operations, and dissemination of information.[10] Proper governance and oversight should allow registry teams to be transparent with stakeholders in terms of operations, decision-making, and the reporting of results.

These governance roles may be fulfilled in a variety of ways. Many of the roles could be assumed by a single committee (e.g., steering committee). Governance and oversight functions you might consider include:

Executive or steering: Responsible for major financial, administrative, legal/ethical, and scientific decisions that determine the direction of the registry.

Scientific: Responsible for determining the overall direction of the database inquiries and recommending specific analyses to the executive or steering committee. May include experts in database content, general clinical research, epidemiology, and biostatistics.

Data access, use, and publication: Responsible for establishing the process by which registry teams access and perform analyses of registry data. Also responsible for reviewing and responding to requests

for data from external researchers or entities, if the registry's primary stakeholders and participants are agreeable to such use.

Liaison: This function may be required in large registries to focus on the relationships with funders, health care providers and patients.

Adjudication: Responsible for reviewing and confirming cases (relevant in rare disease registries) or outcomes that may be difficult to classify. Adjudicators should be blinded to the exposure (product or process) under study so that the confirmation of the outcomes is made without knowledge of exposure.

External review/advisory board: May be useful for providing independent oversight throughout the course of the registry.[10]

Best Practices

- Document your decision-making process.

  Whatever registry governance model you establish, to ensure transparency with stakeholders and participants, it is good practice to document your process for decision-making, including your process for requesting the use of registry data by internal registry researchers and your process for considering how future, possibly unanticipated, requests for data access by external researchers or entities will be evaluated.

- Develop a plan for stopping or transitioning the registry.

  You should have a plan for stopping or transitioning the registry, including any archiving or transferring of data and informing of participants, as appropriate.

References

1. Defining patient registries and research networks. In: Workman TA, ed. *Engaging patients in information sharing and data collection: The role of patient-powered registries and research networks.* Rockville (MD): Agency for Healthcare Research and Policy; 2013: https://www.ncbi.nlm.nih.gov/books/NBK164514/. Accessed 2023 Mar 14.
2. Gliklich RE, Dreyer NA, eds. Registries for evaluating patient outcomes: A user's guide, 2nd edition. *(Publication No.: 10-EHC049)*. Rockville (MD): Agency for Healthcare Research and Quality; 2010: https://www.ncbi.nlm.nih.gov/books/NBK49444/. Accessed 2023 Mar 14.
3. Statistics Canada. Statistics: Power from Data! *(Surveys and statistical programs – Documentation: 12-004-X)*: 3.4.5 Record linkage. 2021; https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch3/5214780-eng.htm. Accessed 2023 Mar 14.
4. Optimizing the integration of real-world evidence as part of decision-making for drugs for rare diseases: What we learned. Ottawa (ON): CADTH; 2022: https://www.cadth.ca/sites/default/files/RWE/pdf/optimizing_the_integration_of_real_world_evidence_as_part_of_decision-making_for_drugs_for_rare_diseases.pdf. Accessed 2023 Mar 14.
5. Report of the Drug Safety and Effectiveness Data Access Sub-Working Group. Ottawa (ON): CADTH; 2023: https://www.cadth.ca/sites/default/files/hta-he/HC0029-data_access_sub_working_group-en.pdf. Accessed 2023 Mar 14.
6. Cambridge Academic Content Dictionary. Registry. 2024; https://dictionary.cambridge.org/dictionary/english/registry. Accessed 2023 Mar 14.
7. Gliklich RE, Dreyer NA, Leavy MB, eds. Registries for evaluating patient outcomes: A user's guide, 3rd edition. *(Report No.: 13(14)-EHC111)*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2014: https://www.ncbi.nlm.nih.gov/books/NBK208616/. Accessed 2023 Mar 14.
8. Statistics Canada. Directive on microdata linkage *(updated December 7, 2022)*. 2017; https://www.statcan.gc.ca/en/record/policy4-1. Accessed 2023 Mar 14.
9. Panel on Research Ethics. Introducing the Tri-Council Policy Statement: Ethical conduct for research involving humans – TCPS 2 (2022). 2023; https://ethics.gc.ca/eng/tcps2-eptc2_2022_introducing-presentation.html. Accessed 2023 Mar 14.
10. Gliklich RE, Leavy MB, Dreyer NA, eds. Registries for evaluating patient outcomes: A user's guide, 4th edition. *(Report No.: 19(20)-EHC020)*. Rockville (MD): Agency for Healthcare Research and Quality; 2020: https://www.ncbi.nlm.nih.gov/books/NBK562575/. Accessed 2023 Mar 14.
11. Office of the Privacy Commissioner of Canada. Guidelines for obtaining meaningful consent. 2021; https://www.priv.gc.ca/en/privacy-topics/collecting-personal-information/consent/gl_omc_201805/. Accessed 2023 Mar 14.
12. Chapter 4: Framework for navigating legal and ethical requirements for PCOR. *Legal and ethical architecture for patient-centered outcomes research (PCOR) data ("Architecture")*. Washington (DC): Office of the National Coordinator for Health Information Technology (ONC); 2017: https://www.healthit.gov/sites/default/files/page/2018-02/PCOR%20Architecture_Chapter%204_S_ONC%20FINAL.pdf. Accessed 2023 Mar 14.
13. Administrative data linking consent wording tool *(Version 1.1, October 2021)*. Vancouver (BC): Health Data Research Network Canada; 2021: https://www.hdrn.ca/sites/default/files/2021-12/FINAL%20Administrative%20Data%20Linking%20Consent%20Wording%20Tool%20V1.1_20211027.pdf. Accessed 2023 Mar 14.
14. Office of the Privacy Commissioner of Canada. The Personal Information Protection and Electronic Documents Act (PIPEDA) *(Date modified: 2021-12-08)*. 2021; https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/. Accessed 2023 Mar 14.
15. De-identification guidelines for structured data. Toronto (ON): Information and Privacy Commissioner of Ontario; 2016: https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf. Accessed 2023 Mar 14.
16. UNC Research. De-identified, coded, or anonymous? How do I know? 2020; https://research.unc.edu/2020/05/01/de-identified-coded-or-anonymous-how-do-i-know/. Accessed 2023 Mar 14.
17. Anonymisation: Managing data protection risk code of practice. Wilmslow (UK): Information Commissioner's Office; 2012: https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf. Accessed 2023 Mar 14.

18. The Glossary of Education Reform, Great Schools Partnership. Aggregate data. 2015; https://www.edglossary.org/aggregate-data/. Accessed 2023 Mar 14.
19. Health Data Research Network. Privacy legislation by province or territory. 2024; https://www.hdrn.ca/en/dash/resources/privacy/privacy-legislation/#2. Accessed 2023 Mar 14.
20. European Organisation for Rare Diseases, National Organization for Rare Disorders, Canadian Organization for Rare Disorders. EURORDIS-NORD-CORD joint declaration of 10 key principles for rare disease patient registries. 2012: http://download2.eurordis.org/documents/pdf/EURORDIS_NORD_CORD_JointDec_Registries_FINAL.pdf. Accessed 2023 Mar 14.

Canada's Drug Agency
L'Agence des médicaments du Canada
Drugs. Health Technologies and Systems. Médicaments, technologies de la santé et systèmes.

Canada's Drug Agency (CDA-AMC) is a pan-Canadian health organization. Created and funded by Canada's federal, provincial, and territorial governments, we're responsible for driving better coordination, alignment, and public value within Canada's drug and health technology landscape. We provide Canada's health system leaders with independent evidence and advice so they can make informed drug, health technology, and health system decisions, and we collaborate with national and international partners to enhance our collective impact.

Questions or requests for information about this report can be directed to Requests@CDA-AMC.ca.