Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

CIHR IRSC
Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

Guidelines for Protecting Privacy and Confidentiality in the Design, Conduct and Evaluation of Health Research:

BEST PRACTICES

CONSULTATION DRAFT

April 2004

Canadian Institutes of Health Research
Privacy Advisory Committee

Canada

For further information, please contact:
Canadian Institutes of Health Research
410 Laurier Avenue West, 9th Floor
Address Locator 4209A
Ottawa, Ontario K1A 0W9
Telephone: (613) 941-2672
Fax: (613) 954-1800
E-mail: **info@cihr-irsc.gc.ca**
Web site: **www.cihr-irsc.gc.ca**

Cat. No.: MR21-48/2004-1E-PDF
ISBN: 0-662-37062-7

Table of Contents

General Comments:

Q. 1. On the draft guidelines.

Q. 2. On ideas for new sections that could be developed and added to the guidelines.

CIHR developed these draft guidelines with the advice of its Privacy Advisory Committee, bringing the perspectives of the following groups:

Privacy Commissioners
David Loukidelis
Information and Privacy Commissioner of British Columbia

Debra Grant (privacy-enhancing technologies)
Research Officer
Information and Privacy Commissioner/Ontario

Research Ethics Boards (REBs)
Sharon Buehler
Co-Chair, Research Ethics Board, Memorial University

Don Willison (Principal Investigator on REB study)
Scientist, Centre for Evaluation of Medicines,
McMaster University

Health Researchers
Charlyn Black
Director, BC Centre for Health Services and Policy Research

Colin Soskolne
Professor, Department of Public Health Sciences, University of Alberta

Voluntary Health Organizations
Roy West
Co-Chair, Science and Research Committee, Health Charities Council of Canada

Patients/Consumers
Mary Vachon
Consultant in Psycho-Social Oncology and Palliative Care

Phil Upshall
Chair, Canadian Alliance on Mental Illness and Mental Health;
President- The Mood Disorders Society of Canada

Policy-makers
Heather McLaren
Director, Legislative Unit
Manitoba Health

Data Producers / Custodians
Joan Roch
Chief Privacy Officer
Canadian Institute for Health Information

Michael Wolfson
Assistant Chief Statistician
Statistics Canada

Aboriginal Interests
Bronwyn Shoush
CIHR Institute Advisory Board Member- Institute of Aboriginal People's Health,
Director, Aboriginal Justice Initiatives Unit
Alberta Solicitor General

Health Service Providers
Denis Cournoyer
Associate Physician, Associate Professor
Dept of Oncology, McGill University

Ethics/Law
Brent Windwick
Health Law Practitioner, Field Atkinson Perraton

Bartha Knoppers
Professor, Public Law Research Centre Law Faculty,
University of Montreal

Ex Officio Members
Interagency Advisory Panel on Research Ethics (PRE):
Pierre Deschamps, PRE member
Member of the Canadian Human Rights Tribunal

Social Sciences and Humanities Research Council of Canada (SSHRC)
Christian Sylvain
Director, Corporate Policy and Planning,

National Council on Ethics in Human Research (NCEHR)
Fern Brunger, Council Member
Assistant Professor, Memorial University

Health Canada
Brian Foran
Director, Office of Health and the Information Highway

International Advisor to the Group
William W Lowrance
International Consultant in Health Policy and Ethics, Geneva, Switzerland

Canadian Institutes of Health Research
Patricia Kosseim- PAC Chair
A/Director, Ethics Office

Sheila Chapman
Senior Policy Advisor, Ethics Office

Mylène Deschênes
Ethics Policy Advisor, Ethics Office

Sylvie Burion
Project Officer, Ethics Office

Introduction

GOALS

These best practice guidelines are intended to:

1. Provide guidance for health researchers across Canada to address privacy, confidentiality and security concerns in the design and conduct of health research involving the use of personal information.

2. Provide a resource for Research Ethics Boards and institutions across Canada to use in reviewing and evaluating the privacy, confidentiality and security aspects of health research.

3. Promote the uptake and application of these best practice guidelines in the development of privacy laws or policy, toward the objective of supporting a more coherent and harmonized policy framework for protecting privacy in health research across Canada.

QUESTIONS

Q. 3. Do you agree with the goals?

Q. 4. Any ideas for improvement?

STATEMENT OF VALUES

This is a time of great change in health research, particularly with respect to privacy issues. For example, technological advances in information technology and the advance of genetic research are challenging existing standards and mechanisms for privacy protection. Also, the sheer number, diversity and complexity of new privacy laws and policies within and beyond Canada's borders are increasing the practical challenges facing researchers, particularly for those conducting studies across jurisdictions. And, while there are increasing demands for privacy protection in health research, there is also clear recognition that health research plays a critical role in improving the health of Canadians and supporting an evidence-based health care system.

At the heart of these best practice guidelines are two core values: (i) respect for the privacy of individuals and (ii) recognition of the social value of health research. The challenge is to balance these values not one against the other, but with a view to maximizing the benefits of both values, while minimizing potential harms from the neglect of either one.

These best practice guidelines are firmly embedded in the ongoing commitment to uphold the broad ethical framework articulated in the *Tri-Council Policy Statement: Ethical Conduct for Research involving Humans* ("TCPS"), of which *respect for privacy and confidentiality* is one principle among the following fundamental and interrelated principles:

- *Respect for human dignity*
- *Respect for justice and inclusiveness*
- *Respect for free and informed consent*
- *Balancing harms and benefits*
- *Respect for vulnerable persons*
- *Respect for privacy and confidentiality*
- *Minimizing harm*
- *Maximizing benefit*

These best practice guidelines are also grounded in an internationally recognized set of core principles regarding privacy, confidentiality and security, known as fair information principles. These principles are at the heart of Canadian privacy legislation and form the basis of the Canadian Standards Association (CSA) *Model Code for the Protection of Personal Information.*

These core principles are:
- *Accountability*
- *Identifying Purposes*
- *Consent*
- *Limiting Collection*
- *Limiting Use, Disclosure, and Retention*
- *Accuracy*
- *Safeguards*
- *Openness*
- *Individual Access*
- *Challenging Compliance*

In conjunction with the broader TCPS ethical framework, these privacy principles form the foundation of these best practice guidelines.

QUESTIONS

Q. 5. Do you agree with the statement of values?

Q. 6. Any ideas for improvement?

SCOPE OF APPLICATION

Health Research

Research is generally defined as a systematic investigation designed to develop or establish principles, facts or generalizable knowledge. The goals of health research, in the context of CIHR's mandate, are to create new knowledge and to enable its translation into improved health for Canadians, more effective health services and products and a strengthened health care system.

This document has been developed as part of CIHR's mandate to foster the discussion of ethical issues and the application of ethical principles to health research. While recognizing the interdependence of a range of activities related to the improvement of health and health services, these best practices are intended as a resource primarily for the health research community. Coverage of related activities, such as public health surveillance, health service management, and program quality assurance and improvement, are beyond the manageable scope of the present document. However, these best practice guidelines could potentially be adapted or customized by others to fit contexts that fall outside what is generally perceived to be the boundaries of health research.

Personal Information

Personal information is generally defined as information that:

- identifies an individual (e.g. name and street address), or
- could potentially identify a person by reasonably foreseeable means if information contained in the record is combined, or information in the records is combined with other available information.

These guidelines deal with all kinds of personal information (and data) about individuals used in the health research context, ranging from clinical data to information relating to the use of health care services to information relating to broad determinants of health, such as education, employment, and income level.

The scope of personal information covered in these best practice guidelines includes personal information derived from human biological specimens and tissues. Although human biological specimens and tissues raise particular issues related to banking and storage, which are beyond the scope of this document, it is their status as records of personal information, like any other recorded information, which is of interest here.

Canadian Context

These best practice guidelines address the current Canadian context for health research. They are intended to complement – not substitute - the relevant privacy and confidentiality principles of TCPS or applicable privacy legislation.

These guidelines have been designed as a practical user manual that synthesizes, clarifies and provides further detail about core privacy principles that are well accepted in Canada today. They do *not* replace existing laws, policies and professional codes of conduct that may apply to certain types of personal information, designated organizations and/or specific kinds of activity.

Researchers, research ethics board (REBs) and institutions should be aware of, and continue to comply with, the relevant laws, policies and codes, including the TCPS, that govern research activities in their respective jurisdictions. In the case of multi-centre research crossing provincial, territorial or even national borders, several privacy laws and policies may have to be considered and complied with.

To help researchers navigate the sea of privacy laws and policies, CIHR has prepared reference documents outlining relevant Canadian and international legal norms.[1]

QUESTIONS

Q. 7. Do you agree with the scope of application?

Q. 8. Any ideas for improvement?

COMMITMENT TO CONTINUOUS LEARNING AND REVIEW

These best practice guidelines are expected to evolve over time, in response to changes in conditions for research and as new best practices emerge. One of the valuable ways in which researchers, research ethics boards and institutions can assist the future evolution of these guidelines is by bringing to the attention of the CIHR Ethics Office lessons learned through the application of these best practice guidelines and offering suggested areas for further development. A feedback mechanism will be set up on the CIHR website to facilitate this input.

QUESTIONS

Q. 9. Do you agree with the commitment to continuous learning and review?

Q. 10. Any ideas for improvement?

OVERVIEW OF THE CURRENT LANDSCAPE OF RESEARCH

To understand the scope of these guidelines, it is helpful to consider the multi-faceted landscape of health research in this country.

Research projects may cross provincial, territorial or national boundaries, and disciplines.

Research teams may be composed of a multidisciplinary network of investigators drawn from across the country— CIHR's 13 "virtual" institutes are founded on this model, promoting collaboration among investigators working on similar questions from different perspectives, in fields such as biomedical, genetic, clinical, health services, health system, and public and population health research. A single health research study may have multiple sites in several provinces, territories, or even at the national level. Because health is a global issue, researchers collaborate with colleagues in other countries as they have for the multi-year international Human Genome Project to map the human DNA sequence.

[1] CIHR's publications on international and Canadian privacy legislation and policies can be accessed at: http://www.cihr-irsc.gc.ca/e/publications/186.shtml

Health research is conducted in various settings, often supported by a mix of public and private funds.

A great deal of research is based at universities, where investigators may have both public and private funding sources. Research into areas with commercial potential, such as the development of new drugs and medical devices, is also conducted by private companies. Governments and affiliated research or statistical agencies conduct research on such things as emerging public health issues, and the effectiveness of the health care system, and increasingly look for private-public partnerships in sponsorship. Statistical and research agencies with a public mandate conduct research on site and frequently also operate as data stewards, permitting access to their data by external researchers under strict controls.

Potential data sources for health research are also diverse.

Individuals are one essential source of health-related data. Individuals are recruited, for example, for surveys on the health status of the population and clinical trials of new treatments and therapies. For long-term research databases (such as clinical research databases, registries with a research mandate, and human genetic material banks) data are collected from individuals and other sources, for research use over an extended period of time.

Another important source of health-related data are existing databases that were not originally created for research purposes. These databases have the potential to provide data that cannot be obtained directly from individuals, such as physician diagnoses and records of hospital treatment (health administrative databases), official registration of births, deaths and cause of death (population records), and disease trends and geographic "hot spots" in the population over time (health surveillance databases).

Thus, these guidelines have a broad scope, encompassing a wide spectrum of health research with a common goal of contributing generalizable knowledge to protect and improve human health. (See Table on page 10)

Examples of databases with research potential, held in diverse settings.

DATABASES	EXAMPLES OF DATA ELEMENTS	EXAMPLES OF RESEARCH POTENTIAL	EXAMPLES OF DATA HOLDERS
Health administrative databases	• Health insurance registration • Physician diagnoses (billing data) • Hospital records	Research at the population level on such things as the epidemiology of disease (e.g. interactions between the environment and health, and trends in disease and wellness over time), and the impact of changes in the health care system.	• Government Ministries of Health • Hospitals • Statistical agencies
Population records	• Records of all births, deaths, cause of death	Research on prenatal and post-natal care and health outcomes, and on long-term outcomes of health conditions (e.g. time and cause of death).	• Provincial and Territorial registrars • Statistical agencies
Clinical research databases	• Detailed data on patient status, care and associated health outcomes	Research on such things as the effectiveness of clinical treatments.	• Disease clinics and institutes (e.g. diabetes, heart disease)
Registries	• A central repository of data (e.g. hospital, clinic and survey data relating to individuals) drawn from various sources; may be population-based or specific to a particular group such as those affected by a particular disease	Epidemiological and health services and policy research.	• Government agencies • Charitable foundations
Human genetic material banks	• Primary materials (blood, bone and tissue culture) • Secondary materials (copies of primary samples such as cellular protein) • Tertiary materials (electronically stored information such as DNA sequences)	Research in such areas as disease diagnosis, the genetic basis of variability in drug efficacy and safety (pharmacogenetics), and the genetic and biochemical basis for disease.	• Government public health and research laboratories • Private companies • Universities • Hospitals
Health Surveillance databases	• Public health data on chronic and communicable disease trends • Reports of adverse health effects from marketed products	Research on public health issues (e.g. causes of disease outbreaks; long-term trends in health status) at the community or population level.	• Government Ministries of Health • World Health Organization • Statistical agencies
Survey databases	• Health, employment, housing, income, workplace conditions, population demographics, education, health services availability • Self-reported personal behaviours, health status, medical conditions, lifestyle, attitudes, values, and experiences	Research on broad determinants of health (individual, biological, social, cultural, and environmental).	• Government departments • Statistical agencies • Universities • Research centres

QUESTIONS

Q. 11. Do you agree with the overview of the current landscape of research?

Q. 12. Any ideas for improvement?

FUTURE CONSIDERATIONS: CHANGING LANDSCAPE OF HEALTH RESEARCH

The research landscape is an evolving one, as our knowledge and technological capacities continue to advance. For example, the advent of a system of electronic health records across the country is expected to present a potentially rich data source for health services and population-based health research. The terms and conditions for access to these electronic health records by researchers in such a way as to ensure appropriate safeguards for privacy and confidentiality are currently under development, in consultation with the various levels of government and with Canadians.

QUESTIONS

Q. 13. Do you agree with the future considerations?

Q. 14. Any ideas for improvement?

Executive Summary

These draft privacy guidelines are the outcome of years of ongoing dialogue with the broad research community. Once finalized, they are intended to provide practical guidance to health researchers and research ethics boards with respect to privacy, confidentiality and security issues in the design, conduct and evaluation of health research, and to inform the development of privacy laws and policies. The foundation of these draft guidelines is the ethical framework in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* and the Canadian Standards Association's (CSA) ten fair information principles in its *Model Code for the Protection of Personal Information*.

These best practice guidelines cover ten interdependent elements that can be briefly summarized as follows:

ELEMENT #1: Determining the research objectives and justifying the data needed.

At the outset of the research design process, health researchers should identify and document research objectives and questions as thoroughly as possible, as a basis for determining what data will be needed. They should anticipate and document research questions that are related to the primary research objective, which might become relevant after the initial data analyses. They should also anticipate and document likely future uses of the data, including possible collaborations with other researchers or possible commercial uses.

ELEMENT #2: Limiting the collection of any personal data.

Researchers should plan to collect personal data at the lowest level of identifiability necessary to achieve the research objectives. Limiting data identifiability means minimizing, as much as possible, the collection of:

- direct identifiers (e.g. name, street address); and
- other data elements that could potentially be used to identify an individual.

ELEMENT #3: Determining if consent from individuals is required.

Voluntary and informed consent from legally competent individuals or authorized third parties is a fundamental principle in research involving humans, and specifically for the use of their personal data or human material. When personal data are collected directly from the individual, consent will almost always be required. When there is a legitimate reason for collecting personal data from other sources, consent will generally be required unless the researcher can demonstrate why the consent requirement should be waived. Under limited circumstances, a waiver or partial waiver of a consent requirement may be permitted by law and approved by an REB.

ELEMENT #4: Recruiting prospective research participants.

A proposed recruitment procedure should normally ensure that initial contact with an individual about a research project is made by someone whom the individual would expect to have relevant information about them, or by other non-invasive means. The recruitment procedure should not inappropriately interfere in the lives of potential research participants, nor exert any undue pressure on eligible individuals to agree to participation in research.

ELEMENT #5: Informing prospective research participants about the research.

Researchers should provide full and frank disclosure of all information relevant to voluntary and informed consent, and researchers must ensure that prospective participants are given adequate opportunities to discuss and contemplate their participation. The consent procedure should begin with the researcher or other appropriate person explaining to prospective research participants things such as the nature of the research, what personal information will be collected and how it will be used, and the risks and benefits of the research to participants, so that participants can make a truly informed decision about whether they wish to participate.

ELEMENT #6: Managing and documenting consent.

Consent or the refusal of consent should be given by legally competent individuals or authorized third parties. This may be done in writing (preferred), orally, or clearly indicated by conduct. Evidence of consent or its refusal should be clearly documented and available for audit and legal purposes.

ELEMENT # 7: Safeguarding data confidentiality.

Institutions or organizations where research data are located have a responsibility to establish appropriate institutional security procedures. Data security measures should comprise organizational, technological and physical measures. Researchers and their employers should take a risk management approach, appropriate to the extent, sensitivity and identifiability of the data.

ELEMENT #8: Limiting access to personal data.

Data sharing for research purposes is an important way of enabling socially valuable research. It avoids unnecessary data collection, which reduces the burden on individual respondents and permits researchers to use their resources more productively. But sharing must be done with appropriate protections for privacy and confidentiality, by controlling levels of data access and having secure procedures for data linking, subject to data sharing agreements.

Element #9: Retaining, destroying and archiving data.

Data should be retained as long as is necessary to fulfill the research purposes. Data may be destroyed or returned to the data provider, if appropriate. However, final research data sets should generally be archived for the use of the scientific community, where resources exist and with appropriate protections for privacy and confidentiality.

ELEMENT # 10: Ensuring accountability and transparency.

Individuals and organizations engaged in health research are accountable for the proper conduct of such research in accordance with applicable funding policies, privacy principles and legislation. Roles and responsibilities should be clearly defined and understood. Recognizing that transparency may enhance public support for, and interest in, socially valuable health research, individuals and organizations should be open to the public about their research objectives and policies and practices for managing personal data used for research, and should promote ongoing dialogue between the research community and privacy oversight agencies.

QUESTIONS

Q. 15. Do you agree with the executive summary?

Q. 16. Any ideas for improvement?

Guidelines

FOREWORD

These guidelines are organized into a series of elements that should be considered in the design, conduct and evaluation of health research to address privacy, confidentiality and security concerns. These elements are not meant to represent a linear, step-by-step process, since many of the elements build on one another and are inter-dependent.

In general, these elements address the basic scenario – for example, a single research study with a specific purpose to be conducted over a clearly defined period of time— and go on to address other considerations that might arise in more complex research designs, such as when a disease registry is being created prospectively for general research purposes.

The elements described in these guidelines, and why each element is essential, are described below.

ELEMENTS	WHY IS THIS ELEMENT ESSENTIAL ?
1. Determining the research objectives and justifying the data needed.	• To anticipate and more clearly articulate to research participants and/or relevant oversight bodies the purposes for which data are being collected, used or disclosed. • To avoid the likelihood of situations where new research objectives or related uses of the data that could have been foreseen prior to collection are identified only afterwards when there may be challenges to re-contacting individuals to obtain new consent or returning to the appropriate oversight body for new authorization. • To help determine the appropriate retention period, since as a matter of principle, data should be retained for the research project only as long as necessary to fulfill the stated purposes.
2. Limiting the collection of personal data.	• To promote a considered approach to the collection of personal data, so that only what are necessary, or could be justified as a future potential necessity, are collected. • To avoid unwarranted invasions of privacy and risk of breaches of confidentiality from excessive collection of personal data.
3. Determining if consent from individuals is required.	• To demonstrate respect for the fundamental right of privacy and to allow individuals to determine for themselves how, and to what extent, personal information about them is used and communicated to others. • To minimize or avoid potential invasions of privacy and other risk of harms to individuals, and where appropriate, to groups. • To avoid excessive precautions that may impose undue burdens on individuals and researchers.

ELEMENTS	WHY IS THIS ELEMENT ESSENTIAL ?
4. Recruiting prospective research participants.	• To minimize potential invasion of privacy, by having the initial contact about the research conducted by someone whom prospective participants can reasonably expect already has personal contact information and any other personal health-related information. • To separate, to the extent possible and appropriate, the research activity from the nature of the existing relationship between data holder and research participant. • To eliminate or reduce any undue pressure on participants and enhance the voluntary nature of their consent. • To avoid situations where prospective participants are not aware, prior to being contacted by researchers, of health information about them.
5. Informing prospective research participants about the research.	• To enhance the conditions for informed consent by ensuring that at the initial contact and throughout the consent process, prospective participants are appropriately informed of all the necessary aspects of the research that make consent truly meaningful.
6. Managing and documenting consent	• To consider the most appropriate form that consent should take in the given circumstances. • To demonstrate compliance and accountability with consent requirements by having documented evidence of consent or the denial of consent.
7. Safeguarding data confidentiality.	• To protect the confidentiality of data entrusted to researchers and guard against inadvertent disclosures of individuals' identities. • To demonstrate accountability to research participants, data subjects, data providers, research institutions and the public.
8. Limiting access to personal data.	• To avoid breaches of confidentiality in the use of data within the research team, or the disclosure of data to third parties or the public. • To ensure that data linkages are conducted by appropriate bodies, at appropriate sites and under appropriate controls to enhance the protection of privacy and confidentiality.
9. Retaining, destroying and archiving data.	• To protect privacy and confidentiality while also fostering the availability of data for the use of the general scientific community committed to advancing socially valuable health research.
10. Ensuring accountability and transparency.	• To recognize the shared responsibilities of key players for ensuring effective protection of privacy and confidentiality in health research. • To provide for ongoing compliance with the terms under which research was approved, taking into account the various roles and responsibilities of all those involved in health research. • To ensure that public is well informed about the research (and its risks and benefits) and the good will and trust of the public, research participants, data providers, institutions, funding bodies and governments towards health research are respected and sustained, thus ensuring continuing support and funding for socially valuable research.

QUESTIONS

Q. 17. Do you agree with why these elements are essential?

Q. 18. Any ideas for improvement?

ELEMENT #1: Determining the research objectives and justifying the data needed to fulfill these objectives.

GENERAL PRINCIPLE

At the outset of the research design process, identify and document research objectives and questions as thoroughly as possible, as a basis for determining what data will be needed.

Anticipate and document research questions that are related to the primary research objective, which might become relevant after the initial data analyses.

Anticipate and document likely future uses of the data, including possible collaborations with other researchers or possible commercial uses.

QUESTIONS

Q. 19. Do you agree with the general principle?

Q. 20. Any ideas for improvement?

BEST PRACTICE

1: 1 Research Study

For each research study, identify and document the specific research objectives and related research questions.

Describe and justify data needed to fulfil the research objectives and to answer any related research questions.

Example:

Research Study: *Impact of Ethnic Group Membership and Old Age on Health*

Study Objectives:
To examine and compare the health status, health care, and social involvement of distinct ethnic groups living in [*region X of* p*rovince Y*], to inform policy development by community organizations and governments.

Research questions: (examples)
What is the association between health status, experience of health care and ethnicity?
What is the impact of personal support networks and activity level on health status and perceived well-being?

Data needed and Justification:
Initials: To assist in checking for duplicate records, using a combination of initials and demographic data.

Demographics (date of birth, gender, ethnicity,..): To make between group comparisons on health variables by ethnicity, and between and within group comparisons by other demographic variables.

Physical health and sense of well-being/Use of health services: To investigate and compare health status and perceived health status by health care-related attitudes and use.

Family and friends/Social Activities: To investigate the impact of family structure and interaction and environmental factors on measures of health and well-being.

1: 2 Creation of a database for general research purposes

Define the scope and purpose of the database in a way that will be meaningful for research ethics boards and research participants, even if the boundaries are at a relatively general level.

Even though all of the research studies that will use data from this database cannot be anticipated or explained in detail at the time the database is being created, try to describe the types of studies that will be undertaken.

In addition to the scope and purpose, describe what the database will not be used for. This is an opportunity to reassure research participants and research ethics boards that although the research purposes are not specified, data management and use will occur within a defined framework.

Describe the general types of personal data that are necessary for these general research objectives (e.g. diagnoses, risk factors, outcomes). Include data that are expected to be collected over the lifespan of the database, particularly if there will be multiple data collection periods per participant, or data will be requested from secondary sources. Be as specific as possible.

Example:

Research Registry on Disease X

Research Objectives:

1. Compiling statistics on burden of disease X.
2. Conducting clinical/medical and epidemiological research.

Types of Research Questions (examples):

1. What is the association between disease X and risk factors related to lifestyle such as diet, tobacco use, and physical activity level?
2. What is the risk of developing disease X after exposure to environmental risk factors, such as pollutants in the area of residence?
3. What is the efficacy of screening programs for disease X?

Limits on Data Uses:

Only to be used to compile aggregate statistics and to carry out clinical/medical or epidemiological research on disease X.

Types of Personal Data to be collected over multiple collection periods	**Research Justification**
Name, Address, Telephone number	Contact information to follow-up with participants for further data collection
Demographic data (e.g. age, gender, city of residence, socio-demographic factors)	Assess other variables by demographics of the population
Family history	Disease X is known to have an inherited basis
Diet, Reproductive factors, Physical activity, Anthropometric measures	Assess risk factors for Disease X
Medical conditions, Medication use	Assess co-morbid conditions and medication use and effectiveness
Blood samples	For possible future genetic analysis to assess risk for disease X and genetic inheritance
Tumour tissue samples from pathology departments	For molecular analysis

If appropriate, setting up an advisory committee drawn from the scientific community and those affected by the condition or health event under study, can assist in defining the scope and strategic priorities for a research program in the context of a long-term initiative. (See Element #10 for a model of ongoing monitoring of long-term research by a data stewardship advisory board).

Example:

Multi-year Family-Centered Study on Childhood Condition X

Research Objectives

1. Track and assess the factors that facilitate or hinder the development of family-centered provincial services for children with condition X and their family members, to provide guidance to community organizations and provincial governments.
2. Evaluate family assessment protocols, which include both standardized empirical measures and structured parent interviews, for measurement reliability and predictive validity, to guide the creation of individualized family service plans.

Setting scope of research

- Initial partnership between the research team and Provincial Ministry of Children's Services resulted in agreement on key objectives.
- Set up a Local Advisory Committee to assist in setting out scope and strategic priorities for research program, review research progress and to facilitate the achievement of study objectives; with representatives from the Ministry, provincial clinic for childhood condition X, two community advocacy groups for persons with condition X, and parent representatives.
- National Project Advisory Committee with representation from provinces actively interested in this initiative meets annually to advance services to young children with condition X, and to plan and disseminate project findings.

QUESTIONS

Q. 21. Is the Best Practice acceptable to you?

Q. 22. Any ideas for improvement?

ELEMENT #2: Limiting the collection of any personal data.

GENERAL PRINCIPLE

Researchers should plan to collect personal data at the lowest level of identifiability necessary to achieve the research objectives. Limiting data identifiability means minimizing as much as possible, the collection of:

- *direct identifiers (e.g. name, street address) and*
- *other data elements that could potentially be used to identify an individual.*

QUESTIONS

Q. 23. Do you agree with the general principle?

Q. 24. Any ideas for improvement?

BEST PRACTICE

Data identifiability can be characterized as a continuum or sliding scale, in which the divisions between degrees of "identifiability" and "anonymity" are not always clear-cut. Even a dataset without direct identifiers may present a risk of indirectly identifying data subjects if the dataset contains sufficient information about the individuals concerned.

For example, the collection of certain data elements may increase the likelihood of an individual's identity being inadvertently identified, such as:

- geographic location (e.g. location of residence, location of health event),
- named facilities and service providers,
- rare characteristics of the individual (e.g. a rare health condition), or
- highly visible characteristics of the individual (e.g. ethnicity in certain locales).

These types of data elements, if needed for the research, should be collected at a minimum level of detail consistent with the research objectives.

Direct Collection

2: 1 Consider first whether personal data are needed, or whether aggregate data would serve the research objectives (e.g. data on individuals grouped by age or some other meaningful variable).

2: 2 If personal data are needed for the research objectives, determine the level of identifiability that will be needed.

Does the researcher need to:

- Contact the participant for follow-up data collection?
- Provide data for clinical monitoring of the participant?
- Enable data to be withdrawn from the dataset if the participant withdraws consent?
- Return individual results to the participant?
- Conduct a data linkage with a high degree of accuracy?

If the answer is yes to any of the above questions, the researcher will likely propose:

- the collection of direct identifiers (if data are being collect directly from individuals) or
- the ability to trace data back to direct identifiers (if the research involves secondary use of existing data).

If the answer is no any of the above, the researcher should not collect direct identifiers. However, other potentially identifying elements may be needed to answer the research questions and for other data management reasons, such as to check for duplicate records. The lowest level of identifiability of these other data elements should be used, consistent with the research objectives.

Examples of Reducing Personal Detail in Specific Data Elements Collected

Subject Name

- Full name
- Partial name
- Initials

Age

- Birth day/month/year
- Birth month/year
- Birth year; Age at time of data collection
- Age range (e.g. 5 or 10-year age groups)

Facilities and service providers

- Names
- Specific type of facility, provider (university hospital, family physician)
- Generic class (hospital, medical doctor)

Location

Postal code

- Street address
- 6-digit postal code (e.g. one side of a city street; average of 15 households)
- 3-digit postal code/Forward Sortation Area (average of 7,000 households)
- 1-digit postal code (province or region)

Census area

- Block (e.g. city neighbourhood)
- Census subdivision (e.g. municipality, village)
- Census Agglomeration (urban core: min. 10,000 pop.)
- Census Metropolitan area (urban core: min. 100,000 pop.)

Secondary use

2: 3 For secondary uses of data:

(a) Consider whether aggregate data would serve the research objectives.

(b) If personal data are required for the research, consider whether direct identifiers could be concealed or avoided in shared data, by means of:

- single-coding (the investigator retains the key to link data back to direct identifiers);
- double-coding (an increased level of confidentiality protection over single coding because the investigator does not have access to the key to re-identify individuals); or
- anonymization of identifiers (an even higher level of confidentiality protection, where the link between data provided to the researcher and direct identifiers is permanently eliminated).

(c) Even if the identifiers in shared data have been coded or anonymized, consider how to minimize the collection or sharing of potentially identifying data elements.

KEY-CODING AND ANONYMIZATION OF IDENTIFERS

Single code: A participant's data (or bodily sample) are assigned a random code. Direct identifiers are removed from the dataset and held separately. The key linking the code back to direct identifiers is available only to investigators.

Double or multiple codes (also called "reversibly anonymized"): Two or more codes are assigned to the same participant's data held in different datasets (e.g. health administrative data, clinical data, genetic samples and data). The key connecting codes and providing the link back to participants' direct identifiers is held by a third party and is not available to investigators. If the data recipient is not permitted to request a re-linking of data to identifiers, double-coded data may be considered by the data recipient to have the status of "anonymized data".

Anonymized (also called "irreversibly anonymized"). The key identifying the link between data and the individual's identity is deleted.

Anonymous. No direct identifiers of individuals were collected, or exist, for the data or human sample.

Summary Guide: If personal data are needed, determining the levels of identifiability needed for research-related purposes

RESEARCH-RELATED PURPOSES	SPECIFIC EXAMPLES	DATA REQUESTED FOR THESE PURPOSES WHEN:	
		Collecting data directly from individuals:	**For secondary use:**
CONTACT INDIVIDUALS	Contact the participant for follow-up data collection	Direct identifiers	Key-coded (however, linking back to individuals becomes increasingly difficult for investigators who receive double or multiple coded data, and therefore do not have the key to the code)
	Provide data for clinical monitoring of the participant		
	Enable data to be withdrawn from the dataset if the participant withdraws consent		
	Return individual results to the participant		
DATA LINKAGE	Conduct a data linkage with a high degree of accuracy	Preferred: Direct identifiers and/or a unique personal number (e.g. personal health number)	Preferred: Data holder conducts linkage and provides to researcher the linked dataset (coded or anonymized) that has been stripped of direct identifiers
	Conduct a data linkage with a measurable degree of accuracy sufficient for the particular research	Direct identifiers or potentially identifying data elements (e.g. date of birth, initials, 3-digit or full postal code, gender, specific health data)	
DATA ACCURACY CHECK	Eliminate duplicate records	Direct identifiers or potentially identifying data elements	Key-coded data so that the data holder (preferred) or researcher can use the key to check direct identifiers of records when the researcher detects data duplication
NO LINK BACK TO INDIVIDUALS NEEDED		Anonymous data (i.e. no direct identifiers are collected) although potentially identifying elements may be needed to answer the research questions.	Anonymous or anonymized data; although potentially identifying elements may be needed to answer the research questions.

QUESTIONS

Q. 25. Is the Best Practice acceptable to you?

Q. 26. Any ideas for improvement?

ELEMENT #3: Determining if consent from individuals is required.

GENERAL PRINCIPLE

Voluntary and informed consent from legally competent individuals or their authorized third parties is a fundamental principle in research involving humans, and specifically for the use of their personal data or human material.

Direct Collection

Personal data should generally be collected directly from individuals unless there is a legitimate reason for collecting personal data from other sources. When personal data are collected directly from the individual, consent will almost always be required.

Secondary Use

When there is a legitimate reason for collecting personal data from other sources, consent will generally be required unless the researcher can demonstrate why the consent requirement should be waived.

Under limited circumstances, an REB may approve the waiver of a consent requirement, or a partial waiver of some elements of a consent requirement (see TCPS Article 2.1 (c)). In addition to REB approval, access to personal data for research without consent will be subject to specific legal requirements in relevant jurisdictions.

.

QUESTIONS

Q. 27. Do you agree with the general principle?

Q. 28. Any ideas for improvement?

BEST PRACTICE

3: 1 Direct Collection.

When direct contact with individuals is required for the research procedure, their consent will almost always be required.[2] This requirement applies to research involving:

- Collection of personal data from persons (e.g. in face-to-face meetings, by mail or by telephone).
- Therapeutic interventions.
- Medical examinations.
- Collection of genetic data and human material from persons.

[2] TCPS provides exceptions with regard to naturalistic observation (Article 2.3) and in emergency health situations (Article 2.8).

3: 2 Direct Collection and Secondary Use (Hybrid model).

When a research objective requires the collection of personal data directly from individuals to whom the data belong as well as from other sources with the intention of linking these data into a combined file**,** consent should be sought for both types of data collection at the time of direct contact with the prospective research participants.

If the secondary use involves identifying eligible individuals for a study, the procedures under *Element #4* would be applicable.

3: 3 Secondary Use.

3: 3.1 Consent to be obtained under normal conditions.

When personal data are to be collected for research from sources other than the individuals to whom the data belong, consent should normally be obtained from those individuals.

If the data or human material to be collected are highly sensitive, REBs may require that consent be obtained. [3] For example, consent may be required from individuals (e.g. from donors) for secondary use of human materials such as:

- human gametes.
- foetus and foetal tissue.
- human embryos.
- identifiable human tissue**.**

3: 3.2 Waiver of a consent requirement under certain circumstances.

An REB should consider the following factors in determining whether a consent requirement may be waived for secondary use of data in research:

3: 3.2.1 Harm-Benefit Analysis.

1) Potential harm to individuals is minimized.

In assessing potential harm to individuals, REBs should consider:

- the probability of harm (related to the identifiability of data and the adequacy of security measures), and
- the magnitude of potential harm (related to the sensitivity of data), including potential:
  - physical injury,
  - emotional or psychological harm,
  - social harm (e.g. stigmatization, insurability, employability),
  - financial harm,
  - intrusion on privacy,
  - loss of trust, or
  - negative impact of the research results.

[3] TCPS Section 9-10.

2) Potential benefits of the research to the public and individuals outweigh potential harms to research participants or data subjects.

3: 3.2.2 Impracticability/Inappropriateness of a Consent Requirement

a) Seeking consent from individuals may be considered inappropriate because of:

(i) potential harm to individuals from direct contact, where there is:
- A risk of inflicting psychological, social or other harm by contacting individuals or families with particular conditions or in certain circumstances; or
- A risk of creating additional threats to privacy by having to link otherwise coded data with nominal identifiers in order to contact individuals to seek their consent; or

(ii) contact with individuals is not permitted under a previous agreement, law or policy.

b) Seeking consent from individuals for the use of their personal data may be considered impracticable when there are difficulties in contacting or notifying individuals due, for example, to:

- The size of the population being researched; or
- The proportion of prospective participants likely to have relocated or died since the time the personal information was originally collected; or
- Lack of an existing or continuing relationship between prospective participants and the data holder who would need to contact them (e.g. a patient registry that does not have a regular follow-up program to maintain a complete and accurate record of changes in registrants' contact information over time);

such that:

(i) there is a risk of introducing bias into the research due to loss of data from segments of the population who cannot be contacted to seek their consent, thereby affecting the validity of results and defeating the purpose of the study; or

(ii) the additional financial, material, human, organizational and other resources needed to obtain consent, could impose a hardship on the researchers or organization so burdensome that the research could not be done.

3: 3.2.3 Necessity of the Personal Data.

Personal data, at the proposed level of identifiability and sensitivity, are necessary to fulfill the research objectives. (See Element #2)

3: 3.2.4 Consideration of Individuals' Expectations.

- Individuals have not previously objected to the secondary use of their data for research or to the use of their contact information.

- Expectations of a reasonable person in the circumstances have been considered (taking into account the nature of the research, the type of data to be collected, the context in which the data were originally collected, etc).

- Efforts will be made to consult with focus groups, Aboriginal peoples, community representatives, and/or special consumer associations, as appropriate, to:
  - address possible concerns of affected individuals and communities in the design and scope of the research, and
  - discuss how the results of research will be analysed and disseminated to maximize the public benefits of the research, while minimizing the risk of harm to individuals or communities.

  This consultation process will be a high priority when dealing with vulnerable populations and/or controversial issues.

- In the spirit of openness, the researcher should have an appropriate strategy for informing the general public about the research. (See Element #10- *Transparency*)

3: 3.2.5 Legal Requirements.

In addition to REB approval, access to personal data for research without consent will be subject to specific legal requirements in relevant jurisdictions. For example, some jurisdictions require some or all of the following:

- A data-sharing agreement between the data holder and the researcher (see Element #8);
- Personal data will not be used to contact individuals; and/or
- Any other relevant oversight body will be notified and/or give its approval.

QUESTIONS

Q. 29. Is the Best Practice acceptable to you?

Q. 30. Any ideas for improvement?

ELEMENT #4: Recruiting prospective research participants.

GENERAL PRINCIPLE

A proposed recruitment procedure should normally have the following characteristics:

- *Initial contact with an individual about a research project should be made by someone whom the individual would expect to have this information about them, or by other non-invasive means that do not inappropriately interfere in the lives of potential research participants; and*
- *The proposed recruitment procedure and any materials used should not exert any undue pressure on eligible individuals to agree to participation in research; but, rather, they should foster the conditions for voluntary consent.*

QUESTIONS

Q. 31. Do you agree with the general principle?

Q. 32. Any ideas for improvement?

BEST PRACTICE

In order to recruit research participants, the researcher will typically need to complete the following steps, each of which involves the researcher or another more appropriate person having access to personal data:

Step A: Determine eligibility criteria for the research and assemble a list of eligible individuals.
Step B: Establish contact with these eligible individuals.
Step C: Seek consent for participation in research.

Recruitment raises complex issues around who is the appropriate person to seek consent from individuals for participation in research. On the one hand, individuals may feel more comfortable if approached by the data holder whom they already know has their personal information. On the other hand, individuals may be unduly influenced to agree to participate in research if asked to do so by someone on whom they are dependent, for example, their employer or health provider.

Situations to be avoided include where eligible individuals are not aware, prior to being contacted by the researcher, of information about themselves that makes them eligible for participation in the research. For example, the health care provider may not yet have informed the patient of a diagnosis that is in the patient's health records. The researcher would need to confirm with the data holder that individuals have been informed of relevant health-related information, before the researcher initiates contact.

Summary of Scenarios and Preferred Recruitment Processes

Does the researcher have access to eligibility and contact information?	Does the researcher have potential undue influence?	STEPS A & B: Preferred person to determine eligibility and make initial contact	STEP C: Preferred person to seek informed consent for participation in research
1. Yes. It is publicly available.	No.	Researcher	Researcher
2. (a) Yes: Researcher is the data holder.	No (e.g. researcher has access to prior study data).	Researcher	Researcher
2. (b) Yes: Researcher is the data holder.	Yes (e.g. data holder is treating physician).	Researcher (i.e. data holder) may determine eligibility, but initial contact should be by a "neutral" person (e.g. research nurse or receptionist) or by self-selection (e.g., responding to posters).	Someone other than the researcher (data holder), unless under exceptional circumstances as determined by an REB (e.g. where a researcher /clinician is the preferred person to inform his/her patients of the potential risks or benefits of the proposed research).
3. (a) No: Researcher is not the data holder.	No (e.g. external researcher making data access request).	Data holder to: (a) inform eligible individuals about the opportunity to contact the researchers about the project, or (b) seek individuals' consent for release of nominal information to researcher. Under exceptional circumstances, an REB may permit minimal information to be released to the researcher to determine eligibility and/or make initial contact.	Researcher
3. (b) No: Researcher is not the data holder	Yes (e.g. professor in same faculty).	Someone other than the researcher (e.g. the data holder or research assistant) or via self-selection (e.g. responding to posters).	Someone other than the researcher (e.g. the data holder or a research assistant)

SCENARIOS

#1. Eligibility information and the means of notifying individuals about the research are publicly available. For example, eligible participants are in a city telephone directory, and participants may be contacted using random digit dialling.

- In these circumstances, the researcher would be able to make the initial contact and conduct recruitment without seeking the assistance of an intermediary (e.g. a data holder).

#2. The researcher is the data holder or is employed by the data holder. For example, the researcher already has access to personal data from prior research studies and has no current relationship with the participants.

2. (a) The researcher is not in a position of undue influence over prospective participants with regard to the research.

- The researcher should be able to make the initial contact and conduct the recruitment process without needing an intermediary.

2. (b) The researcher is considered to be potentially in a position of undue influence over eligible individuals with regard to the research or there is a potential conflict of interest. For example, the researcher is a provider of health care to eligible individuals and there is potential confusion between the research objectives and the care and treatment being provided to these individuals because of real or apparent conflict of interest (as determined by an REB).

(i) Preferably, the initial contact with prospective participants should be made by neutral means, such that individuals will not be unduly influenced to participate in the research.

Neutral means of recruiting eligible individuals could include:

- using notices in neutral locations or newspapers and not associating the research directly with the non-research context (e.g. the health care or treatment situation); and/or
- having a person conduct the recruitment who is not in a position of authority over the individual, nor someone on whom the individual is dependent for care or benefits of some kind (e.g. a research nurse or a receptionist); or
- a physician could inform patients that there is a study they may be eligible for and then the research nurse provides details of the research and manages the informed consent process.

(ii) Where the clinician/researcher is the preferred person to inform his or her eligible patients of the risks and benefits of the research (e.g. because of special expertise or in-depth knowledge of the patients' cases) and to seek consent for participation, it is critical that the clinician/researcher is not perceived as being in a position of undue influence. For example, patients must be sure that pre-existing entitlements will not be affected by whether or not they agree to participate in the research. Patients must also be clear about how their reasonable expectations for personalized care and treatment will be met within the research context.

#3. The researcher is not the data holder.

3. (a) The researcher is not in a position of undue influence over prospective participants. For example, the researcher is external to the data-holding organization, and is submitting a proposal to conduct research on patients, employees or students of the organization, or to make secondary use of personal data relating to these individuals.

(i) Preferably, the data holder should determine eligibility of individuals for the research and make the initial contact to:

- inform individuals about the research so that they can contact the researcher, if interested, or
- seek consent from individuals to release their nominal information to the researcher who will re-contact them to seek consent for participation.

Examples:

A health professional society mails out a letter (drafted by the researcher) to its members, which explains how to contact the researcher to hear more about the research.

Pharmacists are automatically notified by a computer flag, at the time of filling a prescription, of any patient eligible for the research (e.g. receiving a certain number of concurrent medications). This automatic flag of eligible individuals for the study is visible only to pharmacists in participating pharmacies. Once the eligible study population is identified, pharmacists seek consent from these individuals to release contact information to the researcher.

(ii) However, this option may be:

- inappropriate if the data holder is considered to be in a position of undue influence over prospective participants, or
- impracticable if:
  - the data holder does not have the resources to carry out the recruitment and therefore the research could not proceed unless an alternative recruitment procedure is utilized, or
  - the data holder does not have an ongoing relationship with the individuals in order to make contact (for example, a registrar of a population records database, or an information processing agency holding health registration and billing information).

If the preferred option is impracticable or inappropriate, the researcher may be given access to minimal personal data only for the purposes of determining eligibility for the research and contacting individuals to seek consent for participation [4]. Personal information should only be released with appropriate confidentiality protection such as a signed confidentiality agreement, and access restricted to the data holder's site.

Minimal personal data provided to the researcher should normally contain only nominal data. However, if health-related data are inherent in the eligibility criteria used to assemble the list of individuals to be contacted, any health-related data provided to the researcher should be minimal or "camouflaged".[5]

Examples:

Study on Hospital Injury: In this case, hospital administrators do not have the resources necessary to search through personnel files in order to identify potentially eligible research participants according to selection criteria specified in the research protocol and to establish prior contact on behalf of the researcher. Therefore, with the approval of the ethics committee, and a signed undertaking of confidentiality by the researcher, hospital administrators provide the researcher with the names of staff, their work location and full or part-time status, in the form of a computer file. The researcher then uses the computer file to exclude staff that do not fit the eligibility criteria and to select a random sample of eligible staff. Senior hospital staff explain the study in general terms to their staff members and inform them that the researcher will be writing in the near future to individuals eligible to be included in the study. Senior staff emphasize that participation is on a purely voluntary basis. Accordingly, the researcher sends letters of invitation to participate in the research only to eligible staff members.

[4] Note: As a condition for allowing data custodians to disclose personal information to researchers without consent, some privacy laws require that researchers undertake not to contact individuals directly.

[5] See definition in Glossary—*Camouflaged disclosure*.

Randomized Health Care Policy Trial: Ministry of Health staff produces a "camouflaged" list of patient names for the researchers, containing scrambled personal health numbers of patients potentially affected by the health care policy with scrambled numbers of a random sample of patients who are not affected by the policy. When the combined list of scrambled numbers is unscrambled and converted to names, addresses and telephone numbers by the Ministry of Health's Client Registry, the health status of each patient remains unknown to the researchers and to the Ministry of Health staff. In order to be most effective, camouflaging should aim to protect the privacy of targeted patients, while limiting the number of patients who need to be contacted overall in order to mask the identity of the target population. In this study, the sample was 80% targeted and 20% camouflaged. The study was approved by the university ethics committee and the privacy branch of the Ministry of Health.

3. (b) The researcher is not the data holder, but does potentially have undue influence over prospective participants with regard to the research. For example, a clinician/researcher at a health care facility wants to do research on patients being treated by another physician in the same facility; or an academic wants to do research on students in his or her university department or program, but not in a class that he or she is currently teaching.

Preferably, the initial contact should be by a person who is not in a position of undue influence or via self-selection (e.g. responding to posters). This neutral person may be the data holder (preferred) or, for example, a research assistant who may be given access to minimal personal data for recruitment purposes only.

QUESTIONS

Q. 33. Is the Best Practice acceptable to you?

Q. 34. Any ideas for improvement?

ELEMENT #5: Informing prospective research participants about the research.

GENERAL PRINCIPLE

Researchers should provide to prospective participants or authorized third parties full and frank disclosure of all information relevant to voluntary and informed consent.

As part of the consent process, the researcher or other appropriate person (depending on the approved recruitment procedure) should explain to prospective research participants such things as the nature of the research, what information will be collected and how it will be used, as well as the risks and benefits of the research to participants, so that they can make a truly informed decision about whether they wish to participate.

Researchers must ensure that prospective participants are given adequate opportunities to ask questions, discuss their concerns and contemplate their participation. (See TCPS Article 2.4)

QUESTIONS

Q. 35. Do you agree with the general principle?

Q. 36. Any ideas for improvement?

BEST PRACTICE

5: 1 Basic information to be provided to prospective research participants

The following categories of information related to privacy, confidentiality and security issues should be included in the information provided to prospective research participants, as they may apply:

Research Objectives & Procedure

- Research objectives and questions.
- What is expected of the participants in the research procedure, including the required time commitment.

Voluntary Basis for Participation

- Voluntary basis for participation, ongoing meaningful opportunities to decide whether to continue.
- Withdrawal without prejudice is possible at any time (but be clear that data which have already been made anonymous cannot be retrieved and destroyed).
- Option of contacting other family members to participate or inform them of the study (e.g. in genetic research, participants should do any contacting of related family members).
- Circumstances under which the researcher may terminate the participants' involvement in the research.

Risks, Benefits, Compensation

- Possible risks or discomforts to the research participant (including physical, emotional and psychological impacts and privacy implications).
- Risks of non-action (e.g. in research related to treatment).
- Expected benefits of the research to the individual, to groups and to society.
- Whether any commercial benefits will be shared with participants.
- Compensation for expenses or inconvenience incurred as a result of the research.

Data types and uses

- Types of data to be collected and why.
- Any planned or foreseeable commercial uses of the data.
- If appropriate, statement indicating whether test results are for research purposes only or can serve other non-research purposes (e.g. clinical care).

Data access and legal disclosure requirements

- Who will have access to the data and for what purposes (include any legal requirements, such as mandatory public health reporting of certain diseases or obligation to produce evidence on court order; access required for scientific integrity such as auditing or verification of data; and any plans to archive or destroy the data).

Confidentiality and Safeguards

- Protection of data confidentiality beyond disclosures and uses stated above (e.g. affirmation that genetic data will not be given to third parties)
- General description of security measures.

Data Retention

- Time period that data will be retained (e.g. a specified time period; or in the case of immortalized cell lines, an indefinite period).

Reporting of Results

- Explanation of the conditions, if any, in which personal results are to be reported back (e.g. results of genetic testing should normally be reported back to the participant through a physician and with provision of genetic counselling; conditions for informing implicated family members of research results should be clearly stated).
- Explanation of the impossibility for researchers to trace results from anonymized data back to individuals.

Inquiries and Complaints:

- Who will be available to answer questions about the research.
- Who to complain to about the research.
- Who to contact if the participant decides to withdraw consent.

5: 2 Direct Collection and Secondary Use (Hybrid model)

For a hybrid project involving the direct collection of data from individuals and secondary use, the prospective research participant should also be informed of:

- All types and sources of personal data to be used.

- Any intended linkages and for what purposes (e.g. provincial health records to be collected and linked to health survey data, to investigate risk factors, health status and health care use in the population).

5: 3 Creation of a Database for General Research Purposes

When personal data are to be entered into database for multiple research uses over an extended period, research participants should also be informed of:

- Type of studies and associated research objectives that might be conducted, with possible examples (e.g. research on cardio-vascular disease).

- Types of data to be collected from all sources including data linkages, and for what research purposes.

- Any anticipated commercial uses.

- How long data will be retained (if for an extended/indefinite period give a specified time for REB review).

- The process being implemented to ensure proper data stewardship and data security, including:
  - The main rules for governing the future uses of the database;
  - The process by which requests for data access will be reviewed and monitored; and
  - The organization or persons to whom the researcher is accountable for the proper management of the data.

- Options for the participant to control future uses of personal data in the database, such as:
  - Re-contacting the participant to seek consent for new uses if desired and practicable; or
  - No re-contacting, but data can be used:
    - Only in a specified form (e.g. coded or anonymized).
    - Only for research related to the stated objectives.
    - Only for research related to specified purposes (researcher or participant to identify specified purposes).
    - Only for research purposes (e.g. no non-research uses such as employment or insurance purposes; and not for any administrative decisions directly involving the individual).

Example: Informing participants and presenting options for control of new uses of registry data

The invitation to participate in the registry is made by a dedicated nurse coordinator employed by, and accountable to, the participating hospital. The nurse coordinator arranges, at an appropriate time for the patient (and his/her family), to explain the registry and seek the patient's consent to participate in the registry. Patients can refuse or can agree to any or all of the following:

- Access to their current hospitalization records by the nurse coordinator to collect information relevant to their condition for entry into the Registry for future research uses;
- A follow-up telephone call by the nurse coordinator 6 months after the onset of their health event to determine longer-term changes in their functional ability—this survey information is also intended for inclusion in the registry for future research purposes;

- Linkage of their data in the registry, with administrative files from the provincial Ministry of health, and other sources, in order to collect information about physician services, laboratory services, subsequent hospitalizations, and deaths for research on resource utilization and health outcomes in condition X patients; and
- Use of their records in future analyses performed at the independent not-for-profit research organization based in City Y. The results of these analyses are released in aggregate form to third-party private companies seeking to improve services and products related to condition X.

5: 4 Using Understandable Language

Information should be communicated to participants so that it is easily understood, in oral and/or written form, as appropriate. Plain language should be used. Recommended comprehension levels range from grade 6 to grade 9. The amount of time taken to communicate information to prospective participants should not be excessive. (See TCPS pg. 2.2 on translation of research information).

5: 5 Temporary Deception/Partial disclosure

Temporary deception or partial disclosure of the research objectives prior to participation is necessary for certain types of research designs (e.g. in some kinds of social science research). The TCPS permits an exception to full disclosure if an REB determines that the following criteria are met:

(i) there is minimal risk to research participants;
(ii) the research is unlikely to adversely affect the participant's rights/welfare;
(iii) it is a practical requirement for the research;
(iv) if possible and appropriate, the participant will be debriefed afterwards; and
(v) the research does not involve a therapeutic intervention.

(TCPS – Article 2.3)

QUESTIONS

Q. 37. Is the Best Practice acceptable to you?

Q. 38. Any ideas for improvement?

ELEMENT #6: Managing and documenting consent.

GENERAL PRINCIPLE

Consent or the refusal of consent should be given by legally competent individuals or authorized third parties. This may be done in writing (preferred), orally, or clearly indicated by conduct. Evidence of consent or its refusal should be clearly documented and available for audit and legal purposes.

QUESTIONS

Q. 39. Do you agree with the general principle?

Q. 40. Any ideas for improvement?

BEST PRACTICE

6: 1 Legal Competence

The TCPS addresses the issue of (a) legal competence to give consent and (b) special procedures for research involving individuals with partial or diminished decision-making capacity.

Competence

"Competence refers to the ability of prospective subjects to give informed consent in accord with their own fundamental values. It involves the ability to understand the information presented, to appreciate the potential consequences of a decision, and to provide free and informed consent…

It does not require prospective subjects to have the capacity to make every kind of decision. It requires that they be competent to make an informed decision about participation in particular research…

The law on competence varies between jurisdictions. Researchers must comply with all applicable legislative requirements." (TCPS, Section E pg. 2.9)

Diminished Competence

"Subject to applicable legal requirement, individuals who are not legally competent shall only be asked to become research subjects when:

(a) the research question can only be addressed using individuals within the identified group(s); and
(b) free and informed consent will be sought from their authorized representative(s); and
(c) the research does not expose them to more than minimal risks without the potential for direct benefits for them." (TCPS, Article 2.5)

6: 2 Managing Express and Implied Consent

6: 2.1 Required Conditions

Express Consent
The majority of research studies use an express (opt-in) consent as the preferred form of consent, when appropriate and practicable (see Element #3).

An opt-in mechanism means that prior to the start of the research or data collection, informed individuals give clear indication that they voluntarily agree to participate in the research.

Opt-in consent can be given in writing (e.g. by signing a consent form), orally (e.g. in a face-to-face or telephone encounter with the researcher) or by conduct (e.g. by filling out and returning a questionnaire received by mail). Consent is only voluntary if it can be withdrawn at any time.

Implied consent with opt-out
Implied consent with an opt-out mechanism is to be used only when an REB considers express consent to be impracticable or inappropriate.

A valid opt-out mechanism means that individuals have the opportunity at some time during the research or data collection process to give a clear indication (in writing, orally or by conduct) that they do not want to be participants in the research or to have their data used in the research.

If individuals do not choose to opt-out of the research, their consent is implied as long as they were given reasonable notice of the research and meaningful opportunity to opt-out.

Ranked Forms of Consent and Associated Conditions

Type of Consent	Specific Forms of Consent	Required Conditions for REB consideration
1. Express, opt-in consent *(preferred)*	**Ways of opting in:** 1. Written *(preferred)* 2. Oral 3. Conduct (e.g. returning a questionnaire)	**All of the following:** • Voluntary. • Informed. • Unambiguous. • Obtained before beginning the research. • Consent can be withdrawn at any time, with a clear understanding of what that means, for example: ◉ no further collection of additional data; ◉ no further analyses using these already collected data; or ◉ removal of data from the database to the extent possible (e.g. irreversibly anonymized data will be impossible to isolate and retrieve). The process of consent to be documented by the researcher.
2. Implied consent, with opt-out	Consent is assumed because the person does not opt out **Ways of opting out:** 1. Written *(preferred)* 2. Oral 3. Conduct (e.g. leaving the research site)	For opting out: **All of the following:** • Voluntary. • Informed (e.g. through notices, brochures, letters, media announcements): ◉ of the research ◉ of the opportunity to opt-out. • Accessible means for opting out. • Opt-out may be done at any time before or during the research, with a clear understanding of what opting out means, for example: ◉ no further collection of additional data; ◉ no further analyses using these already collected data; or ◉ removal of data from the database to the extent possible (e.g. irreversibly anonymized data will be impossible to isolate and retrieve). • The process of opting-out to be documented by the researcher.

6: 2.2 Documenting Opt-in and Opt-out Mechanisms

• Written (preferred)
If appropriate and practicable, a written form of opting-in or opting-out of research is preferred.

This should be documented using a consent form or refusal statement signed by the individual.

• Oral
Where written documentation is culturally unacceptable, or where there are good reasons for not recording opt-in or opt-out in writing using a form that the participant signs, an oral procedure should be managed and documented, indicating also that the opt-in or opt-out was conducted orally.

For example, anonymity of data collection or results may be necessary or proposed when the research deals with highly sensitive conditions or activities. Therefore, an REB may determine that oral consent may be documented in a manner that prevents in any way linking the identity of research participants to their data or to results of analyses.

Example: Oral consent

Disease X prevalence study among women undergoing abortion in City Y.

Before undergoing therapeutic abortions, women must necessarily have a blood test. Researchers approached these women in the clinic of a large hospital and sought their consent to participate in the study. Those who agreed to participate were asked to fill out anonymous questionnaires about certain risk factors for disease X.

Also with their permission, leftover blood from the blood test was used to test for disease X. For each participant, a computer generated a specific scrambled code linking the blood sample for the disease test and the answers to the questionnaire. Once the results of the disease tests were linked to the corresponding questionnaire, the computer-generated code was removed.

In this way, it was not possible to identify the research participants, even if one had used the same computer program to try to retrace the scrambled codes. The linked information for each person was thus completely anonymized so that the researchers could look at risk factors and determine the incidence of disease X but could not identify any of the research participants…

• By Conduct
When research participants desire anonymity and the personal data can be collected without the researcher present—such as by using a self-administered questionnaire—individuals could indicate consent by filling out and mailing back an anonymous questionnaire to the researcher. Documentation of the consent should be done separately in order to prevent linking research participants to their data or the results of analyses.

Example: Provision of data indicating unambiguous consent

Study on Workplace Injuries in Nursing and Laboratory Staff

The questionnaire had no name or code number on it and participants were asked not to write their name on it. The cover letter from the researchers asked participants to fill out the questionnaire, put it in the provided envelope and return it through internal [staff] mail. The letter also asked participants to then sign the response card that had their name on it, put it in a separate envelope that was also provided and deposit it into the slotted drop boxes located in each work area. The researcher did not need to know the names of persons who had responded; it was the content of the responses that was of interest. The only identifying information required was on the response card in order to allow the researcher to send targeted reminder letters to those persons who had still not responded. In addition, general reminders to return the questionnaires were also posted in designated work areas in an effort to increase response rates. All of the data collected during the study were destroyed when the study was completed. Questionnaires and response cards were shredded and computer files were deleted.

To minimize the risk of linking questionnaire responses with the names provided on the response cards, the researcher picked up the cards regularly throughout the week and the questionnaires only once every week or two. Furthermore, no data were entered until the end of data collection to reduce the possibility of identifying late respondents. With this method, the researcher could not identify who had filled out each questionnaire, but she would know from the response cards who on the list had or had not returned a questionnaire.

Another risk posed by the study was that information would be revealed about those staff who had suffered an injury at work but who had not reported it, contrary to mandatory hospital reporting policies. Some respondents may not have reported injuries because they did not want to appear careless; others may have wished to avoid the fairly lengthy follow-up procedures required of persons with certain injuries. The researcher had anticipated that this might be the case and understood that this information would be considered quite sensitive. It was for this reasons that the survey was conducted completely anonymously with no ability to identify an individual who might have reported an injury to the researcher but not to staff.

QUESTIONS

Q. 41. Is the Best Practice acceptable to you?

Q. 42. Any ideas for improvement?

ELEMENT # 7: Safeguarding data confidentiality.

GENERAL PRINCIPLE

Institutions or organizations where research data are held have a responsibility to establish appropriate institutional security safeguards. Data security safeguards should comprise organizational, technological and physical measures.

Researchers should take a risk management approach, appropriate for the extent, sensitivity and identifiability of the data.

QUESTIONS

Q. 43. Do you agree with the general principle?

Q. 44. Any ideas for improvement?

BEST PRACTICE

The following safeguards are particularly relevant to research conducted within large institutions or other organizations. However, smaller scale projects should also demonstrate acceptable ways of protecting the confidentiality of data.

7: 1 Threat-risk vulnerability assessment[6]

A vulnerability assessment assists researchers and institutions in determining an appropriate level of security for research data and the means by which the data should be received, used, stored, etc. The following are the main steps in a vulnerability assessment:

[6] Adapted from RCMP Security Information Publication 5- November 1994

Assessment	Examples
1. Determine **what assets** need to be protected	• Databases and files of personal and other confidential data • Database management software • Computer hardware, fax machines
2. Determine **what to protect against**	• Five main classes of threats are: disclosure, interruption, modification, destruction and removal or loss
3. Assess the **probability** of the threat occurring	• Low, Medium or High
4. Assess the **magnitude of the impact** and consequences of the threat if it occurs	• Loss of public trust • Harms to individuals (loss of privacy or trust; social stigmatization; social discrimination affecting financial, employment or other status; loss of benefits) • Loss of assets
5. Assess **existing safeguards** and assess need for additional safeguards.	• For example, are direct identifiers separated from personal records as soon as reasonably practicable? Are highly identifiable and sensitive data stored at highest level of security, e.g. on stand-alone servers?
6. Recommend the **appropriate security safeguards** to protect the assets from threats.	• See security measures proposed in 7.2 below
7. **Update and regularly review these safeguards** (at least annually)	• Respond to changes: ◉ in the internal technological environment, ◉ in the research project and the institution, ◉ in technologies available to threat agents and ◉ in the profile of potential threats.

7: 2 Security Measures

7: 2.1 Organizational Safeguards:

- Ensure ongoing commitment to privacy and continued emphasis of its importance by all involved in the research and the institutional/organizational management.
- All involved in the research project should be subject to an oath of confidentiality.
- Access to personal information should be strictly limited in terms of numbers of persons, for legitimate purposes, and strictly on a realistic need-to-know basis.
- Data sharing agreements between the researcher/institution and all involved should be signed prior to providing any access to data.

- Consequences for breach, including dismissal and/or loss of institutional privileges, should be clearly stipulated.
- Institutions and organizations should, with ongoing commitment of adequate resources:
  - Develop, monitor and enforce privacy and security policies and procedures;
  - Appoint privacy officers and create data stewardship committees; and
  - Implement internal and external privacy reviews and audits.

7: 2.2 Technological measures

- Encryption, scrambling of data and anonymization methods should be used to eliminate unique profiles of potentially identifying information.
- Direct identifiers should be automatically removed or destroyed at the earliest possible opportunity.
- If direct identifiers must be retained, direct identifiers should be isolated on a separate dedicated server/network without external access.
- Camouflage sampling or other techniques could be used to prevent researchers from viewing nominal information prior to gaining consent or from identifying data subjects.
- Strong authentication measures (such as computer password protection, unique log-on ID's, etc.) should be implemented to ensure access to data by only authorized personnel.
- Special protection for remote electronic access and external communications should be installed.
- Virus-checking programs and disaster recovery safeguards, such as regular back-ups should be implemented.
- Where possible, a detailed audit trail monitoring system should be instituted that documents the person, time, and nature of data access with flags for aberrant use and abort algorithms for questionable or inappropriate access.

7: 2.3 Physical security

- Computers and files that hold personal information should be housed in secure settings such as combination lock doors, smart card door entry, locked storage cabinets.
- The number of locations in which personal information are stored should be minimized.
- Architectural space should be designed to preclude public access to areas where sensitive data are held.
- Routine surveillance should be deployed.
- Other special physical security measures should be used to protect data from hazards such as floods or fires.

QUESTIONS

Q. 45. Is the Best Practice acceptable to you?

Q. 46. Any ideas for improvement?

ELEMENT #8: Limiting access to personal data.

GENERAL PRINCIPLE

Data sharing for research purposes— whether of linked or unlinked data sets— is an important way of enabling socially valuable research. It avoids unnecessary duplication of data collection, which reduces the burden on individual respondents and permits researchers to use their resources more productively.

However, data sharing must be done with appropriate protections for privacy and confidentiality by controlling levels of data access and having secure procedures for data linkage, subject to data sharing agreements.

QUESTIONS

Q. 47. Do you agree with the general principle?

Q. 48. Any ideas for improvement?

BEST PRACTICE

8: 1 Controlled levels of data access

Researchers and institutions should protect against unauthorized disclosure and use of sensitive data and data subjects' identities, by controlling access to data.

Controlling access to data for research purposes means, under most circumstances, that:

- sensitive and/or highly identifiable data are accessible to the minimum number of persons necessary on the research team, with appropriate training and subject to security safeguards;

- access to coded data, or to data where the identifiers are anonymized but potentially identifying elements remain in the dataset, may be permitted for researchers outside the research team only under strictly controlled conditions; and

- fully anonymized and aggregated data are made available to the general scientific community and for public use only after appropriate scrutiny to minimize or avoid risks of inadvertent disclosure of individuals' identities.

Controlled Data Access for Research Purposes

Access to:	Required safeguards to include:	Who should be permitted access: (examples)
Direct identifiers	• Access on need-to-know basis • Appropriate training • Oath of confidentiality by employees or research team • No direct access for external research, except for linkage purposes in exceptional circumstances (see 8.2 below)	• Selected members of the research team • Selected institution employees • "Deemed employees" or trusted third parties, subject to the same oath of confidentiality as institution employees
Single or double coded, or anonymized data (dataset may still contain potentially identifying elements)	• Approved projects • Data sharing agreement (see 8.3 below)	• Research team • Collaborators at local sites of a multi-site study • External researchers
Fully anonymous data (data scrutinized and altered to protect against risks of inadvertent disclosure)[7]	• Public use files may require the user to agree to a basic form of data sharing or "license" agreement.	• General scientific community • General public

8: 2 Conducting Data Linkages

The most secure way of conducting data linkages requested by external researchers is for the data holder to conduct the linkage and provide linked data sets without identifiers to the researcher. If that is not practicable, a trusted third party may conduct the linkage or the researcher could conduct the linkage on the data holder's site. As a last option, a researcher may be permitted to conduct the linkage at a secure site but under strict controls, as specified in a data sharing agreement.

[7] Refer to *Statistics Canada Research Data Centres (RDCs): Guide for Researchers Under Agreement with Statistics Canada*, May 2002. (http://www.statcan.ca/english/rdc/rdc_guides.htm)

Ranked Options for Conducting Data Linkages

WHO SHOULD CONDUCT THE LINKAGE	CONDITIONS for REB CONSIDERATION
Option A: Data holder (preferred)	The data holder performs the linkage(s) and subsequently replaces all direct identifiers with a coded identifier prior to releasing the linked data set to the external researcher.
Option B: A trusted third party (e.g. a statistical agency) or Option C: The researcher conducts the linkage on the data holder's site	When the original data holder does not have the technical capacity or resources to perform linkages in-house: • a trusted third party acting as an information manager may conduct the linkage off site; or • the researcher as a "deemed employee" (e.g. the Statistics Canada model) may conduct the linkage on the data holder's site. The third party and the researchers should be bound by equivalent conditions of confidentiality and security as apply to the data holder and the data holder's employees.
Option D**:** The research conducts the linkage off site	If Options A, B or C are demonstrably impracticable, the researcher may conduct the linkage in compliance with a data-sharing/confidentiality agreement with the data holder, setting out their respective and shared obligations, including restrictions on use and disclosure and appropriate security requirements (see 8.3 below). In this situation, any direct identifiers or other personal data not required to answer the research question should be destroyed or returned to the original data holder as soon as is practicable, and in compliance with the terms of the data sharing agreement.

8: 3 Data-sharing agreements

Data sharing agreements bind data providers and researchers to their respective responsibilities and obligations for protecting personal data.

Data-sharing agreements should set out the terms and conditions under which data providers will allow researchers to access personal data for research purposes.

Data-sharing agreements typically include:

Research Purposes
- A meaningful description of the research objectives and method;

Public Benefits of Research
- The anticipated or desired social value of the project and the benefits to be derived from it;

Permission
- Copies of the explanatory material provided to prospective participants and of any consent form;

Data uses
- A meaningful explanation of why the research objectives cannot reasonably be accomplished without access to personal data;
- Identification of data sources for the project and any linkages to be conducted;
- A statement that the researcher will not use the data for any other purpose or disclose the data to other parties without prior authorization by the data provider;

Data access
- A listing of who will have access to personal data within the research team or the institution, and a requirement that each of these individuals be subject to an oath of confidentiality;

Confidentiality/Security
- A description of the physical, organizational and technological security measures in place to safeguard against risks of unauthorized use or disclosure, corruption and destruction (see Element #7);

Retention/Destruction of Data
- The time period for data retention and conditions for the return or destruction of direct identifiers at the earliest reasonable time consistent with the research objectives;
- The possibility for the data provider to authorize an extended retention period;

Required Approvals/Authorizations
- Requirement to obtain REB approval and other relevant authorizations;
- Setting out the duration of the agreement or a time for which the parties are to review the agreement;

Reporting Results
- A requirement that results and data not be released in a form that identifies individuals to whom the information relates;

Contact
- Statement that the researcher will not attempt to re-identify data subjects without prior authorization by the data provider, as appropriate;

Accountability
- Researchers to allow on-site visits by the data provider to monitor or audit data use or to respond to allegations of breach;
- No further data to be provided to researchers if the conditions of the agreement are breached; and
- Referral of matters to regulatory or judicial bodies.

QUESTIONS

Q. 49. Is the Best Practice acceptable to you?

Q. 50. Any ideas for improvement?

Element #9: Retaining, destroying and archiving data.

GENERAL PRINCIPLE

Data should be retained as long as is necessary to fulfill the research purposes. Data may then be destroyed or returned to the data provider, if appropriate, given the terms of the original collection or data sharing agreement. However, final research data sets should generally be archived for the use of the scientific community, where resources exist and with appropriate protections for privacy and confidentiality.

QUESTIONS

Q. 51. Do you agree with the general principle?

Q. 52. Any ideas for improvement?

BEST PRACTICE

9: 1 Retention of Personal Data

9: 1.1 Specific research project

Where personal data are collected and used in the context of a specific research project, identifying personal data should be retained by the researcher as long as necessary to fulfill the original research objectives (identified in Element #1), including related purposes such as tracing, validating or auditing research results as required by regulators, study sponsors and/or publishers.

9: 1.2 Database for general health research purposes

When personal data are collected in a database to support general health research purposes in the future, personal data may be retained for the general purpose originally consented to, subject to security safeguards that are proportional to the extent, nature, and level of identifiability and sensitivity of the data.

Administrative databases used to support health research may retain personal data over the long-term, provided that this is permitted according to legislation or the mandate of a public body such as a government health department.

Any long-term retention of personal data for general health research purposes should also be subject to periodic audits and effective oversight by independent third parties.

9: 1.3 Identifiers and linked datasets

Following the linkage of datasets, researchers should reduce datasets to the lowest level of identifiability needed to accomplish the research objectives. For example, direct identifiers (e.g. name or personal health

number) or potentially identifying elements (e.g. a full date of birth or full postal code) may be needed for data linkage but may not be needed to answer the research questions. In such cases, these identifiers should be destroyed as soon as is reasonably practicable or returned to the data holder, as per the terms of the data sharing agreement (See Element # 8- *Limiting access to personal data*).

Researchers should either destroy the new linked dataset immediately after use, or use enhanced security measures to store it, as per the terms of the data sharing agreement. Within some research or statistical agencies it may not be practicable to unlink datasets after each use, however these institutions should ensure that the linked datasets are used only for authorized purposes.

9: 2. Data archiving[8].

9: 2.1 Availability and confidentiality.

Final research datasets should be archived and made available to the scientific community to promote new research to the fullest extent possible while safeguarding the privacy of data subjects and protecting confidential and proprietary data.

Data do not necessarily need to be anonymized to be archived, but the conditions under which the data were collected need to be respected and weighed. Where datasets contains identifiable elements, investigators would normally be required to sign data-sharing agreements. (See Element #8)

9: 2.2 Timeliness.

Data should be archived as soon as possible, depending on the type of research and the nature of data involved. For time-limited research, final datasets could be archived after the acceptance for publication of the main findings from the dataset. In long-term studies, where data are collected over extended periods or waves, data from each wave could be archived soon after publication of findings from these data.

Where research is co-funded by the private sector, the need to protect patentable and other proprietary data should be recognized and accommodated. Generally a delay of 30 to 60 days after publication of results is viewed as a reasonable period to suspend archiving of data.

QUESTIONS

Q. 53. Is the Best Practice acceptable to you?

Q. 54. Any ideas for improvement?

[8] Some national-level examples: SSHRC is committed to the principle that the various forms of data collected with public funds belong in the public domain, with appropriate protection for confidential data (see SSHRC Research Data Archiving Policy). Statistics Canada promotes use of its data by external researchers through such means as Regional Data Centres, Remote Data Access and the Data Liberation Initiative. In the United States, the National Institutes of Health provides a Data Sharing Policy and Implementation Guidance (http://grants2.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm).

ELEMENT # 10: Ensuring accountability and transparency.

GENERAL PRINCIPLE

Individuals and organizations engaged in health research involving personal data are accountable for the proper conduct of such research in accordance with applicable funding policies, privacy principles or legislation. Processes and practices must be clearly established and implemented in order to give meaningful effect to these policies, principles or laws.

Roles and responsibilities of all those involved in the conduct and evaluation of research should be clearly defined and understood, including that of researchers, their employing institutions, research ethics boards and institutional research ethics review committees; any data stewardship committees and advisory boards; Privacy Commissioners and other legally-designated privacy oversight agencies. Their concerted efforts should aim to provide a coherent governance structure for effective and efficient data stewardship.

Recognizing that transparency may enhance public support for, and interest in, socially valuable research, individuals and organizations engaged in the conduct and evaluation of health research should:

- *be open to the public with respect to the objectives of the research being conducted;*
- *be open about their policies and practices relating to the management and oversight of personal data used for research; and*
- *promote ongoing dialogue between the research community and privacy oversight agencies.*

QUESTIONS

Q. 55. Do you agree with the general principle?

Q. 56. Any ideas for improvement?

BEST PRACTICE

Key roles and responsibilities with respect to privacy and confidentiality include:

10: 1 Researchers (Principal investigator, Researchers)

- To be aware of all applicable policies and laws in the jurisdictions in which the research is to be conducted and to conduct their research in accordance with such legal and policy requirements.
- To seek REB and institutional approval, and where required or considered appropriate, the review or approval of other relevant legal privacy oversight bodies.
- To provide a mechanism to handle queries and complaints about the privacy and confidentiality aspects of the research from participants.

- To promote openness and accountability through publicly available materials, which describe the purpose and conduct of the research project(s), and how privacy and confidentiality concerns are being managed.

10: 2 Federally-funded Institutions

- To develop and apply institutional privacy policies and procedures for the conduct and review of research that meet, as a minimum, the requirements set out in the TCPS, and other applicable funding policies and laws.
- To designate an individual who is accountable for the institution's compliance with those policies and procedures.
- To provide for the education and training of researchers and REB members on how to manage personal data in health research.
- To provide a mechanism for handling queries and complaints about the privacy and confidentiality aspects of research.
- To demonstrate impartial and accountable procedures to investigate allegations of individual non-compliance, with appropriate sanctions for non-compliance.
- To be open with the public about research supported by the institution; processes and practices for managing personal information; and procedures for receiving and handling complaints.
- To foster coordinated data stewardship and institutional review processes within and between institutions.

10: 3 Research Ethics Boards and Institutional Research Ethics Review Committees

- To review any proposed and ongoing research involving humans in accordance with the TCPS and its principles, as well as other applicable laws and policies, including:
  - the institution's own policies;
  - federal, provincial and territorial legislation; and
  - relevant laws, regulations and/or policies of other countries, when research is to be conducted in those countries.
- To serve as a consultative body to the research community and thus contribute to education in research ethics.
- To foster coordinated and consistent REB review processes, particularly with respect to multi-jurisdictional and multi-site research.
- To undertake regular monitoring of research and coordinate reviews of multi-centre research to ensure equivalencies in standards across jurisdictions, by conducting:
  - an annual review of the research (required under TCPS);
  - an audit of critical aspects of the research protocol including: the consent process, safeguards, and methods of anonymizing data prior to disclosure; and
  - other effective monitoring mechanisms, as appropriate.

10: 4 Independent Data Stewardship Advisory Board (See Diagram)

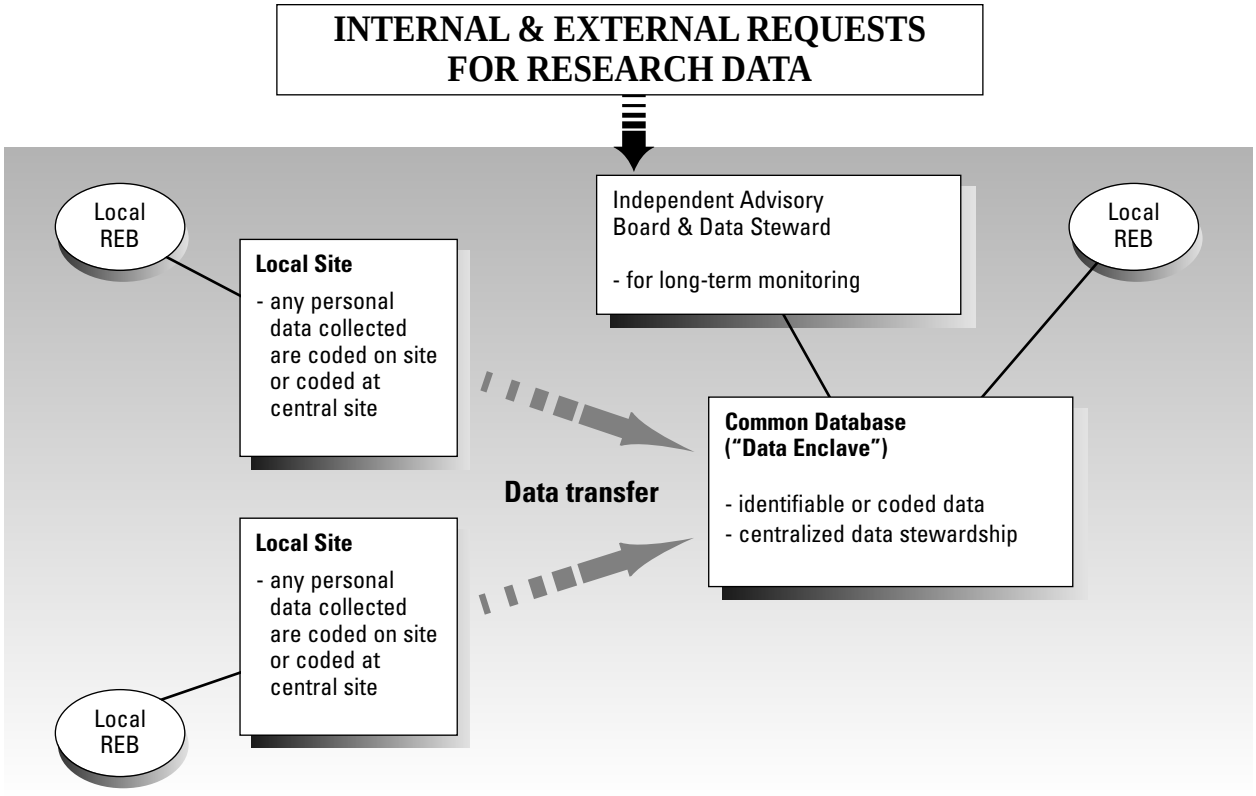
When a database is created for multiple research purposes, or in multiple sites or jurisdictions, researchers and institutional data holders should promote coordinated and streamlined approaches to review of privacy and confidentiality concerns, and to data stewardship over the long-term.

A centralized data stewardship advisory board could be put in place to authorize future uses of the database in accordance with the research objectives and, where applicable, within the parameters set by consent obtained from participants. The responsibilities of this advisory board could include:

- the review of data access requests;
- ongoing management of the database; and
- coordinating reviews by local REBs, for example, by means of agreements between boards, institutions and researchers, as appropriate.

The composition of the advisory board should include scientific experts in the field and representatives from the population being studied. To promote independent oversight of the research database, the principal investigators should not sit on the advisory board.

Diagram: Example of a centralized multi-site research model providing centralized stewardship and approval functions



10: 5 Legally-designated Privacy Oversight Agencies

As specified in legislation, the responsibilities of privacy oversight agencies, such as the Office of the Privacy Commissioner or Ombudsman in each jurisdiction, may include all or any of the following:

- To monitor and investigate compliance with legal requirements.
- To investigate and adjudicate complaints from the public with regard to non-compliance.
- To review and provide comments or approvals on proposed laws or policies.
- To review privacy impact assessments for proposed research.
- To support education with respect to privacy issues.

QUESTIONS

Q. 57. Do you agree with the best practice?

Q. 58. Any ideas for improvement?

APPENDIX (not included)

A-1 Summary and Flow Chart

A-2 Sample Consent Forms

A-3 Sample Data Access and Approval Processes

A-4 Bibliography of relevant literature

A-5 Concordance with TCPS, CSA and Selected Laws

A-6 Privacy Impact Assessment (example)

QUESTION

Q. 59. Any ideas for improvement?

GLOSSARY

The following terms are defined here as used in this document. Readers should be aware, however, that these terms are not yet standardized and may be used somewhat differently in other contexts.

Aggregate data. The data have been averaged or grouped into ranges (e.g. 5 or 10 year age groupings).

Anonymization. Anonymized identifiers (also called "irreversibly anonymized"): the key identifying the link between data and the individual's identity is deleted. **Anonymous:** No direct identifiers of individuals were collected, or existed, for the data or human sample**. Fully Anonymous:** data have been scrutinized for risks of inadvertent disclosure of individuals' identifies, and altered if necessary to minimize or avoid these risks.

Camouflaged contacting. This is an approach to sampling and contacting patients with particular medical conditions in such a way that the individual making the contacting remains unaware of the health status of that individual at the time of contacting. Records of individuals with and without the condition of interest are sampled in some pre-determined proportion from the original source (e.g. administrative or clinical records). Contact information about the combined-sample group is then released without any information about the health status of the individual being disclosed to the person making contact (by telephone or mail). The health status of the individual remains concealed until such time as the individual agrees to participate in the research and to disclose whether or not he or she has the condition of interest.

Coding. Single code: A participant's data (or bodily sample) are assigned a random code. Direct identifiers are removed from the dataset and held separately. The key linking the code back to direct identifiers is available only to investigators. **Double or multiple codes** (also called "reversibly anonymized"): Two or more codes are assigned to the same participant's data held in different datasets (e.g. health administrative data, clinical data, or genetic samples and data). The key connecting codes and providing the link back to participants' direct identifiers is held by a third party and is not available to investigators. If the data recipient is not permitted to request a re-linking of data to identifiers, double-coded data may be considered by the data recipient to have the status of "anonymized data".

Consent. Agreement to participate in research (which may include the collection, use or disclosure of personal data) by a legally competent person or by authorized third parties on behalf of those who lack legal competence. Consent, to be valid, must be voluntary and informed. For consent to be voluntary, the consent must be given without the exertion of undue influence on the person, and with the option of withdrawing from the research at any time without penalty. For consent to be informed, the person must be given information about the research, and must understand this information (See TCPS, Section 3)

Confidentiality. Confidentiality exists when information is communicated in the context of a special relationship (such as doctor-patient, lawyer-client, or researcher-research participant) whereby information is intended to be held in confidence or kept secret. (CSA)

Data. Facts or figures from which conclusions can be drawn. Data can take various forms, but are often numerical, such as daily weight measurements of each person in a group (ref. *Statistics: Power from data*!-

Statistics Canada On-line: http://www.statcan.ca/english/edu/power/toc/contents/htm). See also definitions for *Information*.

Data Custodian. See *Data Holder*.

Data Holder. The Data holder may have custodianship and/or stewardship functions. These functions may be conducted within the same institution/body or may be delegated to distinct but coordinated institutions/bodies. Data custodianship relates primarily to responsibility for data storage and integrity. Data stewardship relates primarily to responsibility for data definition and access authorization, particularly data access and disclosure to third parties.

Data Steward. See *Data Holder*.

Data Subject. The individual who is the subject of personal data/information collected for research purposes. Distinguished from *Research Participant*.

Direct Collection. Collection of data directly from individuals.

Direct Identifiers. These are variables such as name, address or telephone number, etc. that provide an explicit link to a respondent. (Statistics Canada)

Indirect Identifiers. These are variables such as date of birth, sex, marital status, area of residence, occupation, type of business, etc. that, in combination, could be used to identify an individual. (adapted from Statistics Canada)

Impracticable. For the purposes of this document, "impracticable" means a degree of difficulty in doing something under present conditions, where the degree of difficulty is greater than would arise if something is merely inconvenient to do but may be less than if something is impossible. The conditions for assessing "impracticability" of consent are described in Element #3.

Information. Data that have been recorded, classified, organized, related, or interpreted within a framework so that meaning emerges. Information, like data, can take various forms. An example of the type of information that can be derived from data is the number of persons in a group in each weight category or changes in weight over time.(ref. *Statistics: Power from data*!- Statistics Canada On-line: http://www.statcan.ca/english/edu/power/toc/contents/htm). See also definitions for *Data* and *Statistics*.

Personal data/information. Data or information about an identifiable individual, which: (a) identifies a specific individual or (b) can be manipulated or linked with other accessible data or information, by a reasonably foreseeable method, to identify a specific individual.
[Note: TCPS has the following definition of "Identifiable personal information": Information relating to a reasonably identifiable person who has a reasonable expectation of privacy. It includes information about personal characteristics such as culture, age, religion and social status, as well as life experiences and educational, medical or employment histories. It does not include publicly available information, such as documents, records, specimens or materials from public archives or published works, to which the public is granted access. (*pg. 3.2*)]

Privacy. This refers to the right of individuals, groups or institutions to determine for themselves when, how and to what extent information about them is communicated to others.

Research. This refers to a systematic investigation designed to develop or establish principles, facts or generalizable knowledge, or any combination of them, and includes the development, testing and evaluation of research. (Ontario Bill 31).

Research Participant. The individual who consents to participation in research and who is the subject of personal data or information collected for research. See *Data Subject*.

Secondary Use of Data. The data may have been collected originally for (i) a non-research purpose (e.g. for health care administrative purposes or for health care insurance billing purposes), or (ii) a different research purpose (e.g. for a study on a different but related disease).

Sensitivity. The degree of data sensitivity relates to the extent of perceived potential harm to a person or organization—including loss of privacy— from unauthorized collection, use or disclosure of data. Most if not all personal data related to health are considered sensitive. Data are considered particularly sensitive when they carry the risk of social stigmatization or discrimination if disclosed, such as a diagnosis of a sexually-transmitted disease or a genetic test result indicating that an individual has a certain or greater than normal likelihood of future disease or disability.

QUESTIONS

Q. 60. Do you agree with the glossary?

Q. 61. Any ideas for improvement?