

The Pan-Canadian Real-world Health Data Network (PRHDN)

Overview

November 5, 2015

Context

Canada's world-class provincial health-related administrative data holdings have supported both internationally recognized research and local decision making

But research data infrastructure that supports multi-province and pan-Canadian studies will be essential for future success and relevance:

- Multi-province studies are required if we are to learn from and build on the natural policy experiments that have been occurring; for example, as provinces have taken different approaches to primary health care reform
- It is not practical, and in some cases not possible, for individual researchers or small teams to establish and maintain health research data infrastructure for multi-province studies
- Other jurisdictions are collaborating to create health data holdings that, in some ways, rival or surpass the data assets Canadians have benefited from based on our small number of large single-payer health systems

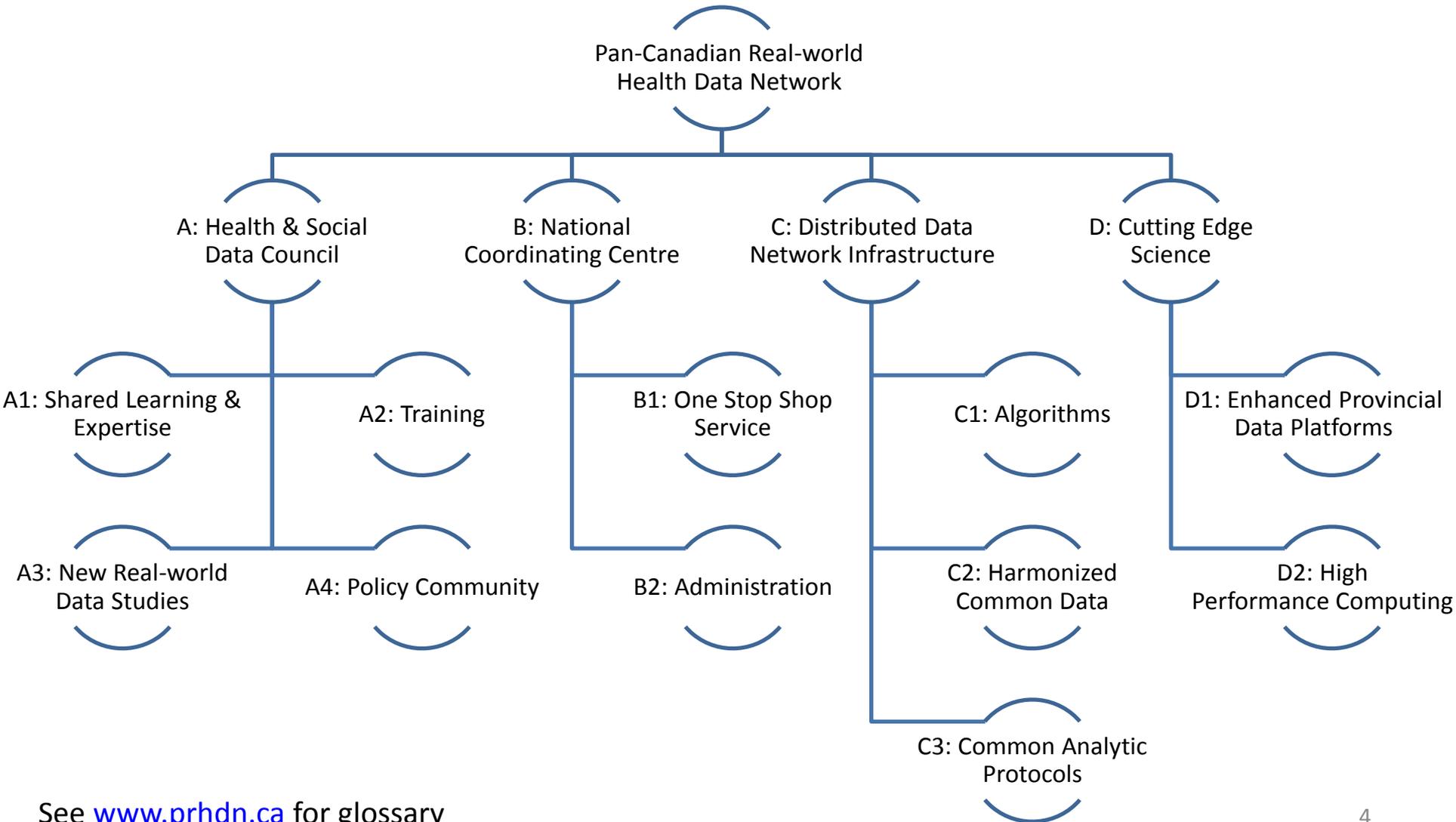
The Opportunity

The PRHDN will be a distributed data network that allows researchers and policy/decision makers from across Canada to make effective use of linked and linkable administrative (real-world) data holdings and expertise in multi-province studies and initiatives without requiring that data leave provincial boundaries

To fully capitalize on Canada's potential to be a world leader in real-world health data studies, the PRHDN will:

- Inform researchers, trainees, policy and decision makers about data, algorithms and documentation that are common across provinces and already available for multi-province studies/initiatives
- Develop and validate new algorithms that generate new harmonized common data and make harmonized common data available for use for a variety of purposes
- Develop common analytic protocols that can be used in cases where harmonization of data between provinces is not practical or possible
- Establish complementary infrastructure to ensure data have their maximum impact (e.g., a National Coordinating Centre, a Health & Social Data Council, mechanisms so that all provinces can learn from cutting edge science and new developments initiated in a subset of provinces, etc.)
- Leverage data and expertise in provincial/regional SPOR SUPPORT Unit data platforms to provide coordinated and enhanced services to SPOR Networks and large pan-Canadian studies (e.g., Canadian Longitudinal Study on Aging, Canadian Partnership for Tomorrow Project)

Work Breakdown Structure Illustrating Full PRHDN Scope



PRHDN Objectives

- Complement and build on existing expertise, structures and processes; providing connecting, not duplicative infrastructure
- Support knowledge creation that is informed by local connections between researchers and policy/decision makers; building on the strength of provincial resources
- Increase the efficiency and timeliness of research by establishing standing research data infrastructure and processes for multi-province studies
- Increase the robustness of studies, and the capacity of researchers and policy/decision makers to use real-world data, by bringing together experts from across Canada to establish and implement common approaches to multi-province analyses
- Create the multi-province data platform that is essential for transformation and innovation in the Canadian health care system

PRHDN Leads Team Members

Population Data British Columbia (PopData)

[Nancy Meagher](#), MA
Executive Director, PopData

[Kimberlyn McGrail](#), MPH, PhD
Scientific Advisor and Chair of the Advisory Board, PopData

Alberta SPOR SUPPORT Unit

[Finlay McAlister](#), MD, MSc, FRCPC, FACP
Lead, Data Platform, Alberta SPOR SUPPORT Unit

Health Quality Council (Saskatchewan)

[Gary Teare](#), PhD
Chief Executive Officer, Health Quality Council (Saskatchewan) and Co-PI, Saskatchewan SPOR SUPPORT Unit

Manitoba Centre for Health Policy (MCHP)

[Alan Katz](#), MBChB, MSc, CCFP
Director, MCHP

[Mark Smith](#), MSc
Associate Director, Repository, MCHP

[Dan Chateau](#), PhD
Research Scientist, MCHP

Institute for Clinical Evaluative Sciences (ICES)

[Michael Schull](#), MSc, MD, FRCPC
President and CEO, ICES and Lead, Data Platform, Ontario SPOR SUPPORT Unit

[Alison Paprica](#), PhD, PMP
Director, Strategic Partnerships, ICES

[J. Charles Victor](#), MSc
Senior Director, Data Platform, ICES

Québec SPOR SUPPORT Unit

[Alain Vanasse](#), MD, PhD
Director, Data Platform, Québec SPOR SUPPORT Unit

[Jean-François Éthier](#), MD, CM, MSc
Associate Director, Data Platform, Québec SPOR SUPPORT Unit

Maritime SPOR SUPPORT Unit

[Adrian Levy](#), PhD
Nominated Principal Investigator, Maritime SPOR SUPPORT Unit

[Leslie Anne Campbell](#), RN, PhD
Interim Director, Health Data Nova Scotia

Newfoundland and Labrador Centre for Health Information

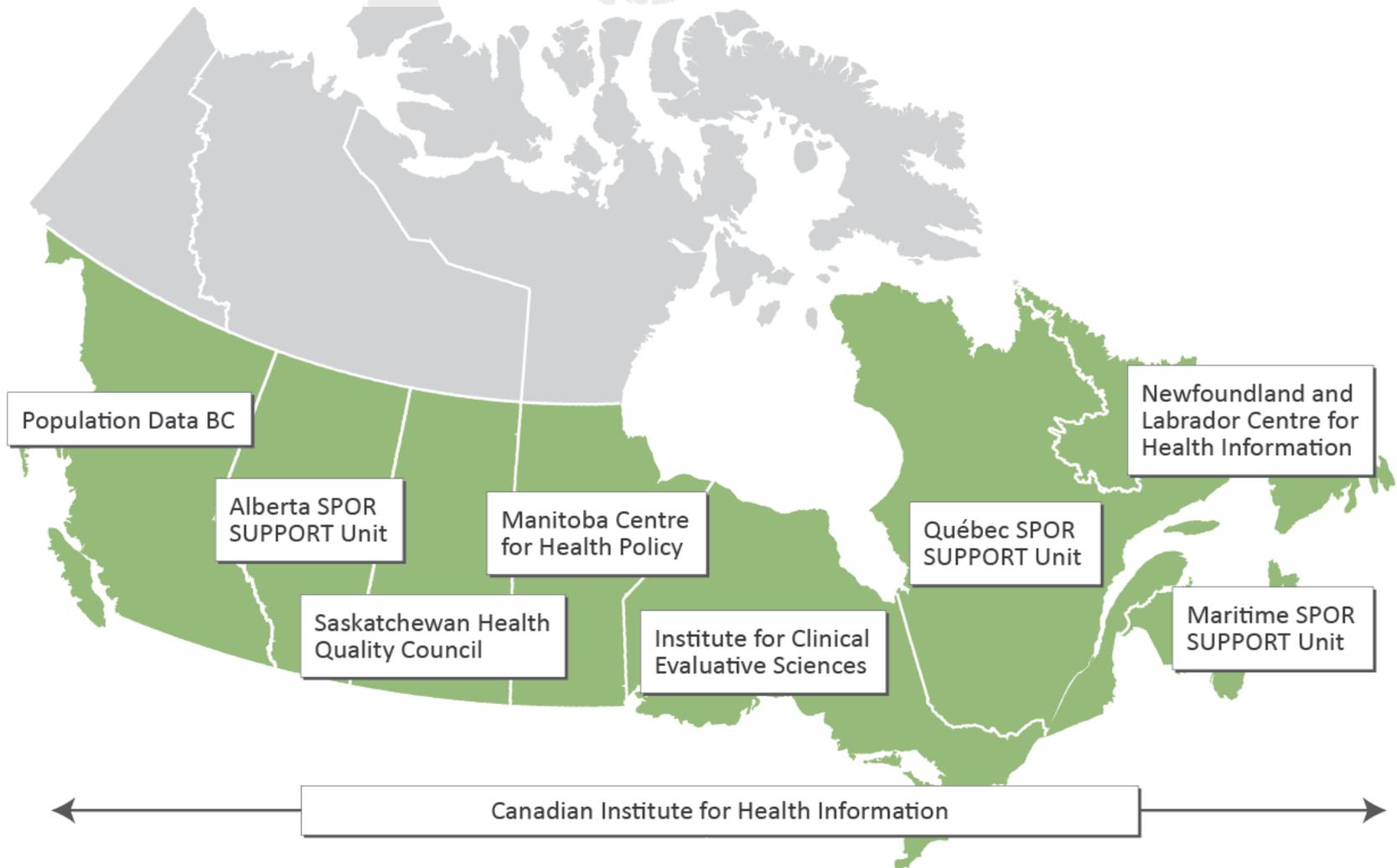
[Don MacDonald](#), MSc, PhD
Vice-President, Research and Evaluation, Newfoundland and Labrador Centre for Health Information

Canadian Institute for Health Information (CIHI)

[Brent Diverty](#), MA
Vice President, Programs, CIHI

[Jean Harvey](#), MHSc
Director, Canadian Population Health Initiative, CIHI

PRHDN Leads Team Organizations



PRHDN Leads, Consortium Members & Potential Collaborators

- The PRHDN Consortium Leads Team includes members from centres with extensive data holdings and expertise: ICES, Manitoba Centre for Health Policy, Population Data BC, Quebec SPOR Support Unit, Alberta SPOR SUPPORT Unit, Saskatchewan Health Quality Council, Maritime SPOR SUPPORT Unit, the Newfoundland Centre for Health Information and CIHI
- These organizations are already nationally and internationally recognized, and have a strong track record of working well together
- To date, 195 researchers and policy/decision makers have joined the PRHDN Consortium (meaning they have an interest in using PRHDN infrastructure once it is established); another 21 individuals have asked to be kept informed of PRHDN developments
- Many centres, networks and initiatives have expressed interest in collaborative work with the PRHDN
 - National: Canadian Network for Observation Drug Studies (CNODES), Canadian Longitudinal Study on Aging (CLSA), SPOR SUPPORT Unit Council, Canadian Partnership for Tomorrow Project (CPTP), Maelstrom Research, Canadian Primary Care Sentinel Surveillance (CPCSSN), Statistics Canada
 - International: FDA-funded Sentinel Network, Australian Centre for Data Linkage, FARR Institute (UK)

Developing the PRHDN

- At this point in time, the PRHDN is a concept that has generated much interest, but does not have funding
- ICES is leading the PRHDN Leads Team work on the first draft of a PRHDN Implementation Plan that will be shared with potential funders by fall 2015
- The PRHDN Implementation Plan will emphasize research data infrastructure that can be established and made available in the short term
 - Beginning with a foundation of common data collected using CIHI standards and algorithms already developed and validated by the Canadian Chronic Disease Surveillance System, within one year the PRHDN Leads Team centres will begin to establish standardized sets of harmonized common data and to make them available to researchers and policy/decision makers
 - Harmonized common data will generally stay within provincial boundaries and be accessed by using existing provincial processes (the exception being data that are already provided to CIHI by provinces)
 - This will prompt studies to be organized with one research lead in each province who works closely with their local provincial policy and practice stakeholders, consistent with the PRHDN objective of building on existing regional strengths and leveraging local knowledge and capacity

Summary

- Though the PRHDN is only in the planning stage, it is already bringing together expertise from across Canada as researchers, policy makers and data holders begin to work together to identify opportunities for enhanced use of health data in Canadian research, policy making and practice
- The PRHDN will build on existing provincial and national strengths, and also has the potential to further develop local data resources
- PRHDN research data infrastructure will be built once using an inclusive process, then used many times over for a variety of purposes including: benchmarking, performance monitoring to identify gaps and opportunities for improvement, evaluation of novel interventions, research and innovation
- Because the PRHDN will be a distributed data network that does not require data to leave provincial boundaries, there are no legislative or privacy barriers to address and the first PRHDN harmonized common data can be available in the short-term
- The PRHDN is open and inclusive; researchers, policy makers and decision makers can find out more, and join the PRHDN Consortium, by visiting www.prhdn.ca

Appendix – PRHDN Glossary

- Algorithm – A reusable process, ideally in the form of documented code, that implements a common approach or definition. Algorithms can be used for a variety of purposes, such as to define cases or to create derived variables.
- Case or Case definition – Definition of characteristics that qualify individuals as having a particular condition (e.g., diabetes, hypertension) or experience of care (e.g., hospitalization).
- Common analytic protocol – A series of technical and process documents detailing the steps for analyses designed to yield results that can be pooled or combined even though the studies are performed using non-harmonized data. Common analytic protocols may include simple statistical methods and/or more complex techniques such as propensity score matching.
- Common data – A collection of variables that are comparable across time and/or jurisdictions through agreed upon content and terminological standards for data collection. Common data variables are a subset of harmonized common data.
- Data – A collection of observations (records) on variables that are available for analysis.
- Derived variable – A variable that does not exist in an original data set but can be created by combining or otherwise manipulating existing information. Examples include the number of chronic conditions, an indication of seeing a primary care provider within a set time period after hospital discharge and total cost of care.
- Emergent properties – Properties that cannot be measured at a service encounter level; they emerge only by considering an individual's experience over time or by aggregating the experience of many individuals. Examples include continuity of care, accessibility and equity.
- Federated data analysis – An analytic approach wherein analyses are run simultaneously on data that remain in separate locations. Federated data analysis is a way to run pooled analyses without requiring data to be in a central location.

Appendix – PRHDN Glossary cont.

- Harmonization: The process through which variables from different time periods, sources, or jurisdictions are made to be inferentially equivalent and thus can be considered common data. Harmonization can be prospective (i.e., by design) or retrospective (i.e., with rigorous methods and validation). Not all variables can be harmonized because there is a requirement for a minimal baseline level of concurrence in the data collection.
- Harmonized common data: A collection of variables that are comparable across jurisdictions, including data that are equivalent from the time that they are collected by the primary source and data that have been harmonized.
- Metadata – Commonly referred to as “data about data”, metadata provide information about the origin and structure of data sets and are intended to inform appropriate use of data. Metadata can include inclusion and exclusion criteria for data sets, definitions for any derived variables, and summary statistics for each field included in a data set.
- Pooled data – Different data sets brought together for the purpose of analysis. This may or may not involve pooling of individual-level data.
- Summary data meta-analysis – An analytic approach wherein individual analyses are performed on common data in their original locations. This is followed by pooling of non-confidential summary results and using statistical techniques to make inferences about the whole (pooled) population. Similar to meta-analysis that is often conducted in systematic reviews of the literature, but in this case it is part of the planned primary analysis of a study.
- Variable – A data element that represents a measurement; for example sex, age, diagnosis. Sometimes referred to as a field or column.