

British Columbia Health Data Asset Inventory: Summary Report

14 December 2016

Project commissioned and managed by:

UBC Vice-President Research Office

Helen M. Burt, Vice-President Research *pro tem*

We would like to acknowledge the support of Deloitte in gathering and compiling the information used to create this inventory.

Project Overview & Approach

British Columbia hosts a large and productive health research ecosystem distributed across regions and sectors. This document presents the first inventory on Health Data Assets¹ in British Columbia, with the hope that such an inventory will be valuable to the entire BC health research ecosystem in its ability to:

- 1) Support planning across the sector related to health data, infrastructure and research priorities;
- 2) Support improved awareness and collaboration related to available assets in the province;
- 3) Support reduced duplication of effort and investment related to health data and related infrastructure; and
- 4) Enable BC to maximize the value (in terms of research and patient outcomes) of existing and planned health data assets.

In addition to conducting a scan of available health data assets in the province, our goal was to organize these assets using an approach that would facilitate discoverability and analysis of assets. First, we created a three-level framework to classify health data assets by their functions and role within the health research ecosystem [see Figure 1]. Then we populated the framework based on an environmental scan and through interviews with representatives from across the sector.

For each data asset, we then categorized its affiliation with BC Health Authorities, its geographic scope [Local, Provincial, National, or International], and its clinical focus [Broad-based, Neurosciences, Cancer, Infectious Diseases, Cardiovascular, Pulmonary, Autoimmune Diseases, Other diseased-based, Child and Youth Health, Women's Health, Ageing and Gerontology, Rural and Aboriginal Health, Other Health Research, Genomics, Health Policy Epidemiology and Health Economics, N/A].

This asset inventory was intended to enable basic reporting by a diverse group of stakeholders in a rapidly-evolving field. Likewise, the data asset inventory has been made available online as an Excel document and will be maintained as a living document through the VP Research Office at UBC [curator: kyle.glenn@ubc.ca].

This document describes the approach used to create the British Columbia Health Data Asset Inventory, its intended use, and a few key findings from the initial version of the Data Asset Inventory [14 December 2016].

Health Data Asset Inventory Structure

Framework to categorize assets

At the broadest level, assets were categorized into Governance, Data User, Data Services, and Data Source. Assets were then further subdivided to differentiate between users/sources of primary vs. secondary data and data services targeted at technical support vs. methodological research support.

¹For the purposes of this project, the term "Asset" is used to broadly include initiatives, governance structures, research centres, technology infrastructure and data related to health.

Level 1	Level 2	Level 3		
Governance	Governance	Strategy and Governance		
Data Users	Primary Users	Health Authorities		
	Secondary Users	Health Authority Institutions/Research Centres	Research Programs	Ministries / Universities
Data Infrastructure & Services	Services	Research Methodological Support		
	Infrastructure	High Performance Computing	Data Platform	
Data Sources	Secondary Data	Administrative Data	Aggregated Data	Registries
	Primary Data	Bioinformatic and Genomic Data	Clinical Data	

Figure 1. Three-level framework used to classify data assets. In many cases data assets served multiple functions and in those cases, the primary role was categorized and a secondary role was noted in the inventory. Definitions for each of these categories are provided in Appendix A.

Governance—Assets that serve strategic or governance role within the health sector, with respect to health data and research

Data User—Assets that utilize data as part of their health research

Primary Users—Users that utilize data for direct patient care purposes

Health Authorities—Assets that directly deliver health services to patients

Secondary Users—Users that utilize data for research, analysis and other purposes not related to patient care

Health Authority Research Institutes/Centres—Assets part of a larger research institution/centre with relationships with a health authority and an academic institution

Research Programs—Assets that use data and are research groups found within health authorities as a group or department

Ministries/Universities—Assets that provide funding and direction and/or academic support

Data Infrastructure & Services—Assets that support health research activities through technical and research methodological support

Research Methodological Support—Assets that provide guidance to researchers regarding research and analytical techniques, tools and standards

Technical Support—Assets that support researchers through technology platforms and computing systems

Data Platform—Assets that help centralize, collect, integrate and manage large sets of data

High Performance Computing—Assets that provide high performance computing capabilities for advanced and/or high-volume analyses

Data Source—Assets that generate primary and secondary data for use by researchers

Primary Data—Assets that generate data for a specific clinical purpose and are used for patient care

Registries—Assets with a consolidated data set focused on a specific clinical area and used for research and/or clinical purposes

Administrative Data—Assets that generate administrative (i.e., non-clinical) data from the health care system

Aggregated Data—Assets that collect and combine data from multiple sources for further analyses, often providing only summary statistics for research and analysis purposes (as opposed to row-level data)

Secondary Data—Assets that generate data that are used for research and/or analysis purposes (the data is originally collected for other purposes)

Bioinformatics and Genomic Data—Assets that generate and store biological and genomic specimens and data

Clinical Data—Assets that generate data at the point of care at a medical facility, hospital, clinic or practice

Initial Findings

As over 14 December 2016, the British Columbia Health Research Data Asset Inventory contained 259 assets. Users [of both primary and secondary data] were the most abundant type of data asset, representing 61% of the inventory followed by data sources at 30%. Data infrastructure and service providers were 8% of the assets and only 1% of the assets were focused on research strategy and governance.

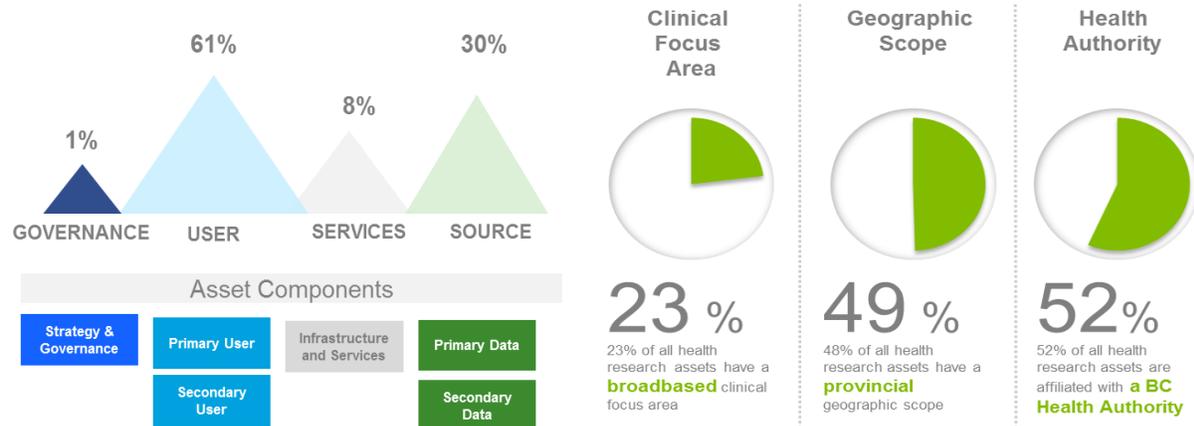


259

Health research assets have been identified as part of the health research network in BC

56%

- of all assets are
- Research Institutes
 - Research Centres
 - Research Programs



The figures above provide a snapshot of the structure and composition of the asset inventory on December 14th 2016 and an example of the types of analyses that can be run using the inventory. The live version of the health data asset inventory file can be accessed here: <https://research.ubc.ca/bc-health-data-asset-inventory>. Please forward any corrections or new additions to the UBC VP Research Office [kyle.glenn@ubc.ca].

Stakeholder Interview Insights

Interview questions

In order to validate the data asset inventory and learn about challenges and opportunities in health data research, Deloitte interviewed 22 stakeholders in the BC Health Research Ecosystem. During the interviews, the following questions were asked:

1. *Have we accurately and completely captured the health data research initiatives and/or infrastructure in place in your organization?*
2. *Where additional detail is required (scope, scale, funding, faculty composition, etc.), please provide an overview and/or the location of relevant reference material.*
3. *What are your organization's future plans and goals with respect to health data research? In other words, what initiatives are currently in the development, planning or delivery phases?*
4. *Please describe any key interactions your organization has with other health data research teams in the province.*
5. *In what stage is your initiative/program? (e.g., initial strategy/concept development or well-established organization/entity)?*
6. *Please describe key outputs and/or impact to date (if any) from your initiative/program.*
7. *What opportunities do you see for more coordinated and strategic approach to health data research in BC? What are the key gaps and what are the attributes of a successful solution?*

Two major themes emerged from the 22 interviews conducted with stakeholders in the BC health research field with respect to the data asset network: Data access and awareness/alignment within research the health research network and some opportunities for improvement were identified.

Data access (benefits and challenges)

- BC has some of the richest sets of administrative and clinical health data in the world due to their longitudinal nature; this health data is extremely valuable to researchers.
- The current process to access health datasets is onerous and time consuming for researchers.
- Significant improvements have been made to decrease wait time from “application to access”, however, significant constraints remain.
- Health data in BC is highly fragmented and is not easily accessible in an integrated manner.
- There is a strong interest in sharing research data and data products amongst researchers. Currently legislative, policy and governance constraints inhibit this sharing (although it is recognized that the SPOR initiative and Academic Health Sciences Network are seeking to address these challenges).
- It is perceived that restrictions related to data sharing in BC are the strictest in the country. Some researchers seek datasets from other provinces or the United States due to challenges in accessing and sharing data in BC.
- There is a stigma and general perception of risk associated with allowing data to be used for research purposes. Participants suggest a shift in mindset towards “data as a public good” is required to unlock the full potential of data available for research and healthcare purposes.
- There is an increasing amount of data being generated with technology (e.g. home health monitoring, wearables) but how exactly researchers will gain access to this type of data is not well known.

Degree of awareness and alignment within the health research network

- There was a consistent comment that researchers did not feel they were aware of the breadth of health research assets and data sets available in the province.
- Data standards between the government, external data platforms and health organizations are not consistent and that inconsistency compromises the ability to link and utilize many types of data.
- Opportunities to exchange and integrate data between researchers have been missed due to a lack of coordination and challenges related to data integration and access.
- There is a lack of coordination regarding the data required by data generators and data users (i.e., assets are not created to maximize the impact of the data collected).
- A mindset shift from viewing assets as independent from each other to viewing the entire health research network holistically is needed to encourage collaboration and a strategic approach to data and related infrastructure and capabilities.
- Better alignment and coordination of research organizations with more formalized strategic and governance structures can help facilitate improved collaboration between researchers.
- Thoughtful planning to link infrastructure and data needs of researchers across the sector can maximize the impact of the data collected.
- Further, such coordination can support focused investment in shared provincial assets while reducing the potential for duplication of effort and investment.

Opportunities for improvement

- Increased adoption of electronic health records can increase data standardization and accessibility; this can increase the amount of useable data available to researchers

- Further standardization (e.g. migrating to single EHR format) of electronic health record technology across the province will encourage interoperability of data between health authorities and researchers.
- Data curators / custodians could be established to help direct researchers in gathering data and answer questions regarding the scope of the data and maintain the health research asset inventory.
 - Such a “data navigation” service could facilitate researcher understanding of data sets (definitions, assumptions, quality, etc.) and could also support linking researchers with data sets best suited to their research questions.
- An established data set inventory can help researchers understand the assets available as well as potential areas of collaboration.
- A more coordinated approach from government and other strategic and governance organizations may help to reduce the time required for researchers to gain access to data
 - Additional legislative and/or policy changes and technology enhancements may also be required to reduce the time required for health researchers to gain access to health data.