



# Alberta Innovates Health Solutions Secondary Use Data Project

---

## II. Companion Document from the Steering Committee

*Date:* October 30, 2015

*Version:* 1.1 FINAL



**I. Document Orientation..... 3**

**II. Overview..... 5**

**III. Our Solution..... 9**

**IV. Our Vision and Mission..... 19**

**V. Go Forward Plan..... 20**

**Appendix A – Project Methodology ..... 27**

**Appendix B – Precise Patient Registries ..... 30**

**Appendix C - Phase 1 Recommendations ..... 32**

**VII. Appendix D - Abbreviation..... 41**



## I. Document Orientation

The document is organized into the following sections:

II. Overview – provides an introduction to secondary data use and describes the background and rationale for the project.

III. Our Solution – describes the health information repository, solution, governance, operations, technology, and analytics components that will make up the solution.

IV. Our Vision and Mission – notes the solution vision and mission.

V. Go Forward Plan – documents the critical recommended actions for successful solution implementation. Describes the high level Phase 2 solution development activities as well as the demonstration projects selected for the solution.

The appendices provide information on working group methodology, precise patient registries, and Phase 1 recommendations.

A separate project charter will be developed to support Phase 2 of the project. It will include information on the following topics:

- Solution Implementation Approach
- Goals and Objectives
- Scope
- Schedule
- Deliverables
- Work Breakdown Structure
- Budget and Resources Estimates
- Risks
- Communication and Change Management

### Terms and Definitions

The term ‘Secondary Use Data Platform’ was originally used to introduce the current project to stakeholders. The term was understood and adopted by the working groups as loosely describing the services and technologies that would support a health information repository. This interpretation was not consistent with other IT and business uses of the term ‘platform’. As a result, the Steering Committee chose to omit this term from its final reports. Instead, the overarching business and technical components that together create a ‘solution’ for improved access and use of secondary health data in Alberta are described in the final reports. The solution includes regulations, technology, privacy and security safeguards, analytics services, governance, and supporting resources.

The working group and steering committee members who contributed to this final report note that data definitions can be difficult to standardize. While much of the data used in Alberta for both primary and secondary



purposes contains the same data elements, it is used for different purposes. It is prudent for the purposes of this report to include the following distinction between primary and secondary use data proposed by the College of Physicians and Surgeons of Alberta<sup>1</sup>:

“Primary use – patient health information collected by their physician (or other healthcare provider) for the purpose of providing health services to that patient. Also includes the provider registration information used to document the provision of care. Use of a patient’s health information by another health care provider to provide a service to that patient is still considered primary use.

Secondary use – using a patient’s health information for any purpose not directly related to the care of the individual patient who is the subject of that information.”

---

<sup>1</sup> College of Physicians and Surgeons of Alberta – Data Stewardship: Secondary Use of Health Information, 2009.



## II. Overview

### Introduction

Data is essential to achieve the ultimate goals of improved health and health care for Albertans. Data provides the vehicle via which we are able to collectively and thoughtfully reflect on our health system performance. It allows us to identify areas requiring improvement as well as analyze what options might work best to make those improvements. It allows us to carry out research that pushes the boundaries of how prevention and treatment are carried out today to create the care options of tomorrow. It provides us a window into our inner system workings so that we can become even more innovative and creative, at the same time conserving and allocating our resources in fiscally responsible ways. It allows us to continually improve upon the health and wellness options available to our citizens. Our health care system needs data to drive improvement.

Secondary use of health data, as defined by the College of Physicians and Surgeons of Alberta, includes ‘using health information for any purpose not directly related to the care of individual patients who are the subject of that information’<sup>2</sup>. Secondary use of health data for research and quality improvement can enhance healthcare experiences for individuals, expand knowledge about disease and appropriate treatments, strengthen understanding about the effectiveness and efficiency of our healthcare system, support public health goals, expedite scientific discovery in medicine, ground clinical research, and aid in the advancement of economic efforts. Data is required to guide the daily decisions faced by clinicians, researchers, funders, delivery organizations, and policy makers.

Visionary work has already begun in Canada to frame what an integrated health data system could look like and how its design and implementation should be undertaken<sup>3</sup>. Following the lead of the Advisory Panel on Healthcare Innovation, Alberta will create a secondary data use solution that supports ‘data-intensive models of care and the rapid-cycle innovations that characterize precision medicine as a field’<sup>4</sup>. Alberta has a unique opportunity to harness these frameworks and experiences by developing a secondary data use solution. Alberta’s population of approximately 4 million, combined with its largely centralized, publicly-funded health system, provides a reasonable population and geographic base for building such a solution.

The current interest and willingness of Alberta’s major health data and information stakeholders to seek an integrated data solution place us in an enviable position; very few jurisdictions in the world can provide these conditions.

---

<sup>2</sup> There are a number of definitions available that encompass primary and secondary use of health data. The project participants noted that primary and secondary data include all the same health data, just used for different purposes at different times. However, the functional definition of secondary data applied for the purposes of this report comes from the College of Physicians and Surgeons of Alberta – Data Stewardship: Secondary Use of Health Information, 2009.

<sup>3</sup> [Canadian Council of Academies, CIHI, Canada Health Infoway, Alberta College of Physicians and Surgeons](#)

<sup>4</sup> [Unleashing Innovation: Excellent Healthcare for Canada. Final Report of the Advisory Panel on Healthcare Innovation, 2015.](#)



Our current shared understanding of the benefits of an integrated data solution and its proposed governance and operations are provided in this report.

## Background and Context

Phase 1 of Alberta's secondary data use solution effort launched on March 12, 2015. Phase 1 of the project was sponsored by the Health Research and Innovation Collaboratory<sup>5</sup>. More than 60 senior leaders, executives and members of the public participated in six representative competency-based working groups; participants represented provincial health, research, and quality organizations. The recommendations and action plan provided in this report reflect the best thinking and expert advice from the working group participants who provided their expertise, time, and guidance over the course of four months. The working groups are noted here:

- Public Engagement
- Privacy and Access
- Data and Analytics
- Infrastructure and Applications
- Governance and Structure
- Enabling Legislation

An overarching Steering Committee comprised of the co-chairs of the six working groups synthesized the recommendations of the working groups and facilitated communication with stakeholder organizations and related data governance groups within the province e.g. Health Information and Governance Committee.

More information about Phase 1 methodology can be found in Appendix A.

The working groups investigated, and where possible, made suggestions and recommendations to address current secondary data use challenges facing Alberta health, quality, and research organizations. Alberta currently lacks the ability to integrate the rich health and health-related data sources at its disposal. Challenges noted by the working groups include the following:

**Data sharing** among organizations does not employ standardized processes or universally accepted technical standards. As a result, access to and sharing of secondary data becomes complicated; no two data sharing agreements are the same and the methods for sharing data vary widely e.g. provision of data files versus use of data within a secure technical environment.

**Data management** practices are not consistent among organizations leading to either a lack of data for secondary use and/or data quality issues. The whole life cycle of data use in Alberta from collection of data and metadata through to retention and purge cycles exhibits large practice variations. Some organizations have well-developed data and analytics policies and well-established practices while others consume data using ad hoc data collection

---

<sup>5</sup> The role of the Collaboratory is to provide long-term strategic direction for health research and innovation in Alberta, to identify research and innovation priorities, and to ensure that provincial programs are integrated and aligned. Membership includes leaders from Alberta Innovates Health Solutions (AIHS), Alberta Health Services (AHS) and the Strategic Clinical Networks (SCNs), Alberta Health, Alberta Innovation and Advanced Education (IAE) and Alberta Innovates – Technology Futures.



and analysis strategies. The resulting fragmentation of information is a serious limiting factor to fully understanding what impacts an individual's health.

**Privacy and access** are often at opposite ends of the data use spectrum. Risk mitigation and privacy concerns are not balanced with the value of the proposed data use, thereby limiting research and innovation. Privacy is protected today on the assumption that all access and disclosure for secondary use carries significant risk, and if the identity of an individual can be ascertained (i.e. re-identified) this is often seen as an unacceptable risk and prohibited by the Health Information Act (HIA). There does not appear to be a risk management lens used by custodians where the value of the intended data use and the risks of disclosure are weighed against each other in determining whether or not access to data should be granted.

**Security risks** related to secondary data use have been described as having a large 'surface area' in that there are many secondary data use situations and opportunities where privacy breaches may occur. Organizations and researchers using secondary data each employ their own security measures and technologies; in addition, most operate without regular review of their security practices. This represents significant risks for involved organizations. In contrast, the report from the Canadian Council of Health Academies notes that there have been no security breaches at the six secondary use best practice entities examined as part of their report:

'While there are clear benefits of research using individual Canadians' personal health and health-related data, there are also risks. These can include accidental release of identifiable data, to the public or unauthorized researchers, when proper security and privacy protocols are not followed (e.g., through loss of computer equipment); illicit access to identifiable data (e.g., through hacking); and inadvertent access to identifiable data by those working inside data organizations. While these types of breaches have occurred during research projects, breaches rarely happen at institutions with databases set up specifically for maintaining large volumes of health and health-related data for research and administrative purposes. They are much more likely to occur when researchers or employees are accessing data directly from health-care centres. Importantly, there are no examples of breaches at the six best practice entities identified by the Panel.'

The working groups found that Alberta lacks an overarching provincial strategy to protect secondary data against threats to the security, integrity or loss of health information, or unauthorized access, use, disclosure or modification of health information. In addition, a secure, central environment that employs privacy and security best practices is not currently available to reduce security risks related to secondary data use.

**Cost** for secondary data analysis, as well as associated storage and security infrastructure, is not wholly known. Working group participants know that tens and hundreds of databases and servers are currently dedicated to computation, analysis, and storage of data for secondary purposes. However, given the distributed nature of these services and technologies within Alberta organizations, it is not possible to generate an estimate of the redundant structures and duplicate costs.



**Governance** related to the secondary use of data is lacking in the province. Coordination of secondary data use efforts as well as the opportunity to learn through the sharing of research and quality improvement results is required. One of the major issues with current secondary data use practices is the lack of attention to the integration and coordination of research efforts; the results include increased costs and negative impacts on the quality and effectiveness of research and quality improvement activities. The working groups note that each stakeholder involved in research and quality improvement has specific business and data needs that must be met in order to use data effectively for secondary purposes. When these specific processes vary between organizations, barriers to effective data sharing arise. This has contributed to a lack of trust in research and quality improvement data sharing initiatives and solutions. There is currently no neutral governing entity to assist organizations to work through these issues and barriers.

Overwhelmingly, the working groups agreed that Alberta needs a solution where health and health-related data can be integrated and used for strategy, planning, policy-making, service delivery, care, and research purposes.

Alberta has the opportunity to build on the experiences of national work outlined in the recently published Council of Canadian Academies report<sup>6</sup>. Other organizations across Canada are also moving forward to create secondary use solutions – Institute for Clinical Evaluative Sciences<sup>7</sup> (ICES), Population Data (Pop Data) BC<sup>8</sup>, BC Centre for Data Innovation<sup>9</sup>, Manitoba Centre for Health Policy (MCHP)<sup>10</sup>, and Better Outcomes Registry & Network (BORN)<sup>11</sup>; opportunities for collaboration have already been initiated with a number of these organizations and will continue as Phase 2 of the project is initiated. In addition, Alberta has several local examples where innovation and excellence in health have been possible through accessing and analyzing health data. The Alberta Bone and Joint Health Institute<sup>12</sup> and the Child and Youth Data Lab<sup>13</sup> can be harnessed as role models and/or incubators for initial solution development activities.

It is important to note, however, that Alberta has a transformative goal in mind for the solution. While many of the above noted organizations and initiatives have demonstrated benefits related to the integration of and access to health data, Alberta is endeavoring to create a much more inclusive and comprehensive solution. One that includes health and health-related clinical, socio-economic, multi-jurisdictional government, administrative and self-reported data that has been aggregated, analyzed, privacy-protected, and made accessible to inform decision making, strategic planning, policy development, service delivery, care, and research to benefit Alberta and Albertans.

---

<sup>6</sup> [Accessing Health and Health-Related Data in Canada](#), Canadian Council of Health Academies, 2015

<sup>7</sup> [ICES](#)

<sup>8</sup> [Pop Data BC](#)

<sup>9</sup> [BC Centre for Data Innovation](#)

<sup>10</sup> [MCHP](#)

<sup>11</sup> [BORN](#)

<sup>12</sup> [Alberta Bone and Joint Health Institute](#)

<sup>13</sup> [Child and Youth Data Lab](#)

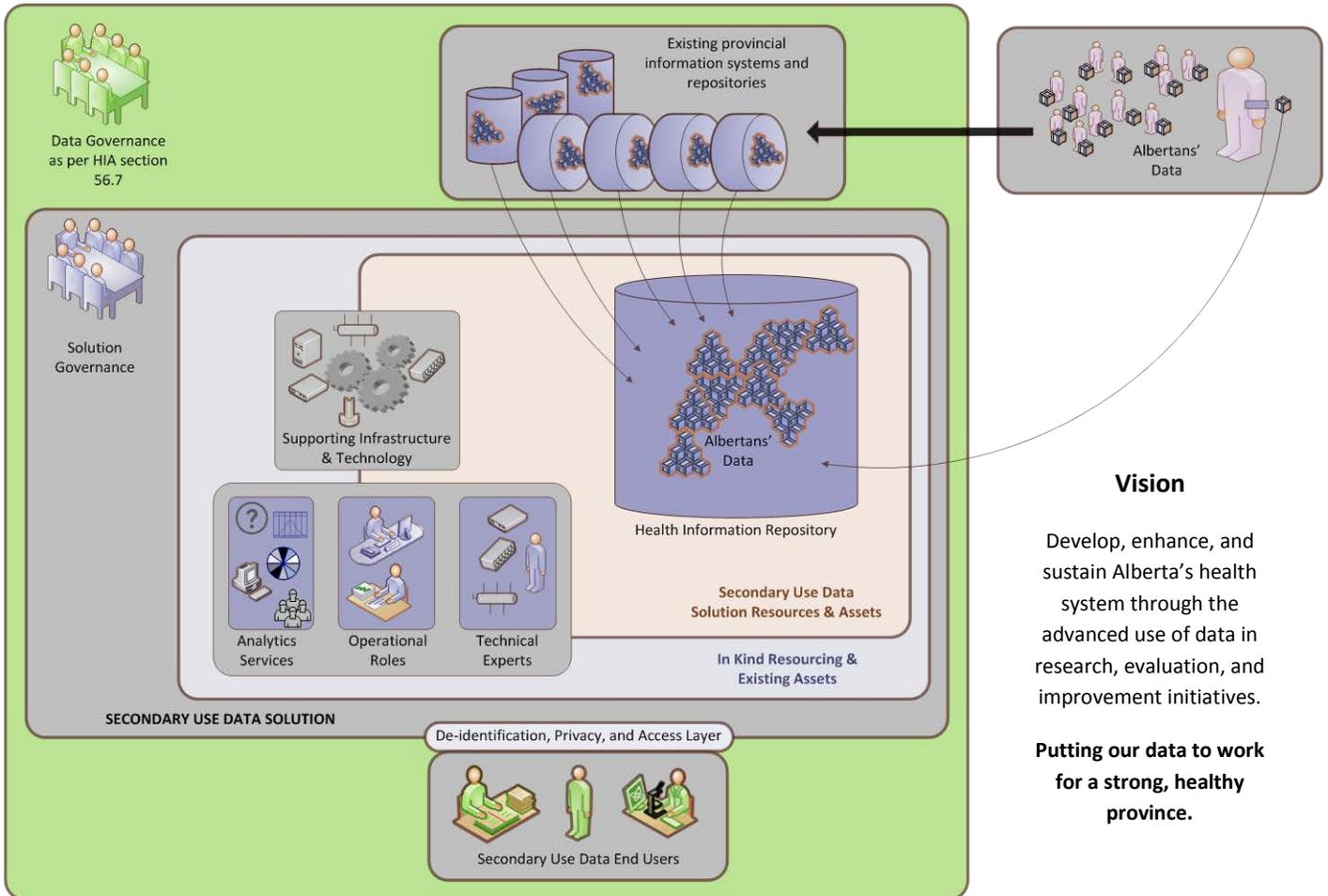


The structure of Alberta’s health system and the available provincial legislation provide for the development of a world-class secondary use solution; the magnitude of health and health-related data slated for inclusion has very few parallels world-wide. As such, Alberta will need to not only collaborate on global work completed to date but also be prepared to lead the way.

### III. Our Solution

The working groups propose that Alberta develop a comprehensive solution that integrates health and health-related data for use in strategic planning, policy-making, service delivery, quality improvement, and research. The solution they recommend involves regulations, technology, privacy and security safeguards, analytics services, shared and equitable governance, and expert supporting resources.

As depicted below, the solution includes a comprehensive health information repository; the repository will integrate health and health-related data on an unprecedented scale, bringing together data sources that are currently used in isolation or only partially integrated. In addition to the health information repository (HIR), the solution will include supporting infrastructure and technology, analytics services, operational roles, and technical experts; these resources will be a combination of net new and existing assets and resources. The solution will be governed by a Solution Steering Committee (SSC) that will report to the data governance entity named under HIA section 56.7. Access to data users and privacy for Albertans will be ensured through the effective use of security and de-identification best practices and technologies.



**Vision**

Develop, enhance, and sustain Alberta’s health system through the advanced use of data in research, evaluation, and improvement initiatives.

**Putting our data to work for a strong, healthy province.**

**Health Information Repository**

The development of a solution relies on a central component – a health information repository (HIR). A comprehensive secondary use HIR will be developed in the province of Alberta. Once a HIR is established with the powers, duties, and functions specified in the regulations, a custodian may, in accordance with the regulations, disclose individually identifying health information to the HIR. Supporting regulations under the Health Information Act (HIA) Part 6.1 will facilitate collection, use, integration, linkage, and maintenance of repository data from disparate health and health-related data sources that are currently used in isolation or only partially integrated. As a result, we will be able to not only access a greater volume of health and health-related data than in the past but also combine this data in new ways to produce new insights.

A HIR provides considerable value-add in terms of researcher access to data; a system that responds quickly and has the capacity to deal with researcher requests is important for continued health system improvements. A central repository (the Act allows for more than one) would formalize and streamline the approach to how



secondary use health data is accessed and used. To quote then minister, Fred Horne, from the Hansard of May 26, 2009, p1285:

“...Health information repositories are intended to improve access to health information for research purposes only. These amendments simply enable health information repositories to deal with research requests in the same manner as custodians currently deal with health information today. Health information repositories cannot authorize the use of health information for any other purpose. In addition, health information repositories do not expand access to health information beyond what is currently permitted under the Health Information Act. By creating health information repositories, greater protection will be provided to health information used for research purposes. Rather than requesting information from multiple custodians, a researcher will be able to request information from a single health information repository. Since the repository is the single source of entry to that information, Mr. Chair, it mitigates threats to privacy. For example, a health information repository will be able to conduct data matching and provide the final data at the highest level of anonymity. Currently researchers conduct their own data matching based on information received from multiple sources. ...[T]he Minister of Health and Wellness will consult with the [privacy] commissioner when preparing regulations. ....”

## Governance

To support the new way forward, the solution will have collaborative and inclusive governance that is aligned with existing governance structures and analytics resources. However, it will fulfill a new mission. Current barriers to the effective use of data are largely cultural and organizational. The solution’s mission will be to remove barriers and provide timely access to data as well as facilitate appropriate and innovative use of Albertans’ data.

Shared equitable governance is required; while existing governance structures need to be considered and leveraged as appropriate, a specific secondary use data governance structure is required to build trust, facilitate collaboration, ensure inclusiveness, and support implementation of the solution. While existing governance committees may offer appropriate representation, knowledge, and competencies, the scope of their current governance responsibilities would not coordinate well with the perspective and scope required to effectively govern the solution’s secondary use functions. In addition, the committees may not include representation from all stakeholder groups that will be involved in the solution.

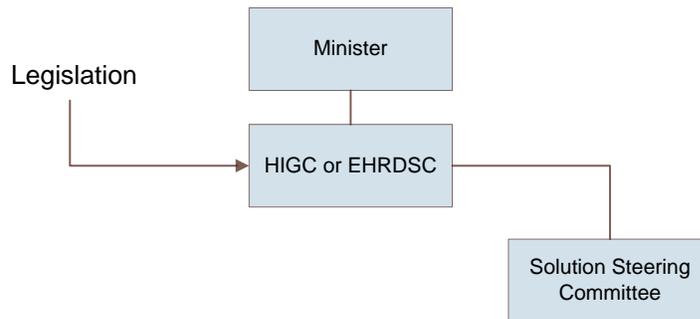
It is important to note that governance will be required at two levels:

- Provincial data governance will be required to ensure alignment of secondary use data initiatives and solution operations with other provincial data initiatives and operations and to provide strategic guidance regarding legislation and regulations that impact secondary use data
- Solution governance will provide the high level policy that guides implementation and operations of the solution. It will ensure solution operational priorities and business practices reflect the needs and interests of Albertans and participating solution stakeholders

In order to quickly provide value to Albertans, it would be prudent to establish the initial solution and secondary use data governance model within existing HIA legislation. However, provincial data governance structures that would normally provide guidance related to secondary use data are currently in a state of change. The working



group and steering committee members note that previously functioning provincial data governance entities such as the Health Information Governance Committee (HIGC) or the EHR Data Stewardship Committee (EHRDSC) would have been able to provide appropriate solution oversight and guidance. However, they also note that discussions are underway regarding the roles, functions, and responsibilities of provincial data governance entities such as HIGC and EHRDSC and that communication regarding clear governance roles, functions and responsibilities is yet to be forthcoming. Despite the changing state of provincial data governance, the Governance and Structure Working group (GSWG) were able to provide solution governance recommendations. The GSWG noted, however, that the recommended Solution Steering Committee would need to report to a provincial data governance entity once the provincial data governance structures are confirmed and communicated with all involved stakeholders.



### Solution Governance Recommendations

The solution is a multi-stakeholder, multi-organizational collaborative that will require shared governance to achieve a collective impact; the SSC will provide an operational governance structure that facilitates achievement of solution goals and objectives. As such, guidelines, rules, and policies will be supported by meaningful discussion and negotiation among SSC members; the SSC will function and display characteristics as noted below:

- Operate under Robert’s Rules<sup>14</sup>.
- Make use of formal votes.
- Operate at arm’s length from any one stakeholder or partner.
- Consider all Solution Steering Committee members equal.
- Be composed of individuals who have the necessary competence and skill sets to support the solution.
- Include representatives of data custodians, data contributors, and data users.
- Include public representation.
- Be limited to 9 individuals.

The SSC will be accountable to regularly report back to stakeholders regarding their stewardship activities as well as areas of success and improvement achieved via solution operations.

<sup>14</sup> <http://www.robertsrules.com/>



In order to support effective governance of secondary use data, governance will consider not only stakeholder representation but also the competencies required to implement and support the solution. SSC membership representation is recommended to include the following six organizations and 1 public member for a total of 9 members.

- Alberta Health
- Alberta Health Services
- Alberta Pharmacists' Association
- Alberta Innovates Health Solutions
- Alberta Medical Association
- Public (2 members)<sup>15</sup>
- Alberta Academic Health Network
- Appointee from Research Community<sup>16</sup>

Not all health professions were included in SSC membership. Membership was prioritized for organizations and professions currently seen to be significant data contributors and/or whose data was currently inaccessible or poorly integrated with other health and health-related data sets e.g. physicians, pharmacists. In contrast, health professions who did not currently have the volume of data available or whose data could already be retrieved by other means were not prioritized for inclusion (e.g. data from nurses and allied health professions is available, for the most part, through AHS clinical information systems; community-based allied health professions may have smaller data sets - these will be prioritized for inclusion later in solution implementation). The rationale for limiting membership was based on governance best practices; creating a small SSC with focused objectives will assist with early solution development and data acquisition/liberation efforts.

The Health Quality Council of Alberta (HQCA) was not included on the membership list as they are primarily data users. However, the HQCA was suggested as a strategic data user whose input may be appropriate for inclusion in provincial data governance entities.

Individuals selected to represent the organizations or groups above will need to demonstrate competencies required to effectively govern the solution. While it is not likely that any one individual would be able to represent all of the required competencies, the seven members selected for the SSC should demonstrate collective competence in the following areas:

- Governance
- Technology startup leadership
- Marketing and communications

---

<sup>15</sup> The Public Engagement Working Group co-chairs noted that this is standard practice to include 2 members of the public on governance groups to ensure full and timely participation in governance activities.

<sup>16</sup> Robyn Blackadar has been put forward as a nominee; this facilitates the competency-based governance approach in addition to the representation provided by other Alberta organizations and public members. Robyn's appointment would build on the success of the Child and Youth Data Lab and set the foundation for cross-ministry data sharing.



- Government affairs
- Privacy Legislation and other Privacy and Security Requirements
- Financial management and fund raising
- Healthcare analytics and data management
- Population and community health
- Clinical research
- Clinical quality improvement
- Healthcare economics
- Analytic technology management
- Governance and regulation of health professionals

The governance, legal, and operational structures associated with the solution will need to evolve as the solution moves from initial start-up phases to ongoing mature operations. The solution will move from a focus on health information repository formation, policy development, legislative compliance, and solution implementation to a mature solution mandate focused on the following:

- Acquiring and integrating data assets, including development of a roadmap for prioritized acquisition of data assets.
- Promoting innovation and excellence in secondary data use e.g. examining licensed access to solution assets and services.
- Evaluating and monitoring solution performance.

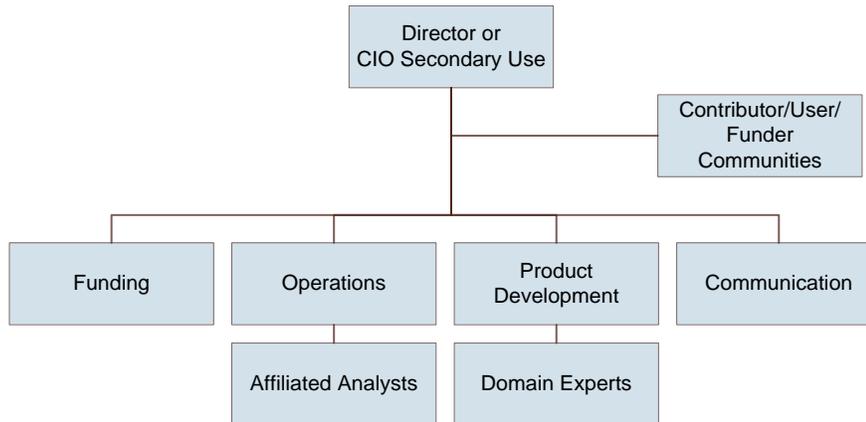
While it is important for the solution to initially leverage existing provincial technical, business, and governance assets to move forward in a fiscally responsible manner, an arms-length entity may need to be established to achieve the mandate noted above. An arms-length entity has the potential to respond to changing end user needs more quickly than traditional government entities. In addition, it would have the added benefit of providing governance independent of data owners. Operating at arm's length would foster excellence in decision making; decisions would be based on end user needs and the needs of Albertans as opposed to being based on the vested interests of SSC member organizations. An arms-length entity would be responsible to assess the needs of its stakeholders and provide strategic solution direction and guidance in the long term; this may eventually include evolving from a SSC to a governing entity such as a board or board of directors. However, before an alternate operational and governance structure are contemplated, the solution must first demonstrate value for its stakeholders and a clear need for independence.

## Operations

A solution operational structure will be required to execute Phase 1 recommendations and facilitate development of policy and regulations to be drafted under Part 6.1 of the HIA. The SSC will need to develop mechanisms for communicating with and providing guidance and direction to the operational team, including support for formal decision making processes. The following initial operational structure is suggested for implementation of the solution. The roles may be a combination of net new solution roles and individuals seconded in part or in full from their current operational roles.



## Secondary Use Data Platform Initial Operations



**DRAFT**

Leadership duties and functions for the solution would include those related to data collection and development, infrastructure (owned or contributed), work flow, general reporting (aggregate data or intelligence), and analytics support.

- The Director or CIO of the solution would be accountable to the SSC.
- Individuals fulfilling solution roles would report to the Director or CIO.
- Solution roles related to operations would include roles related to
  - Operations - procuring and managing an appropriate complement of tools and resources.
  - Funding - working to secure initial and ongoing funding.
  - Product Development - developing an appropriate mix of information and analytics products based on contributor and user needs.
  - Communication - supporting access management, privacy, security, and licensing as well as facilitating adoption and overall system data literacy.
- Solution staff would include a mix of owned and contributory (seconded) resources. Affiliated analysts and domain experts would be linked to the operational structure.

Operational needs of the solution will need to be re-evaluated as solution development activities are carried out in Phase 2; some of the proposed structures and roles may need to be revised or omitted and new structures or roles may need to be added. This includes identifying the knowledge, skills, and abilities that will be required to operate the solution. The roles and individuals required to fulfill these needs may be dedicated to the solution and/or seconded as in kind resources leveraged from existing analytics programs. Example roles may include but are not limited to Director, Chief Analytics Officer, and Chief Privacy Officer.

Solution operations will need to be managed by an existing organization and governed by the SSC. Operational accountabilities would include carrying out the strategies developed by the SSC, managing the solution budget, setting and achieving operational targets, and assuming responsibility for operational roles and priorities. Alberta Health, Alberta Health Services, or Alberta Innovates Health Solutions are three examples of existing organizations



that may 'adopt' the solution as part of their existing organization. Alberta has multiple examples of programs hosted by existing government departments that span multiple organizations; further planning regarding the operational responsibilities will take place in Phase 2 of the project.

## Technology

There are numerous data sources to be considered for secondary use in Alberta. Much of the existing infrastructure and application investment related to secondary use data has been between Alberta Health (AH) and Alberta Health Services (AHS). In contrast, health-related data assets are not yet widely available. In order to better coordinate and access current and future state health and health-related data assets, supporting technology is required. This section includes a preliminary discussion of the technical features, trends, and requirements that will need to be considered as part of solution design and development.

New technologies, based on **open source software frameworks** used to process very large data sets, are competing with traditional relational databases. The solution will need to span the current relational database technologies and build an enabling environment where new non-SQL technology can be assessed and implemented; the aim is to create an environment where the best current and emerging tools available can be continuously leveraged.

Two options exist for configuring the physical infrastructure for secondary use health data within Alberta. The data can be physically co-located within one entity or it can be virtually integrated in a hub and spoke model. When more complicated use cases are introduced to virtual co-location solutions, data binding issues are often encountered. For this reason, a **physical co-location model** is recommended. However, as new technologies evolve, cloud services and increased network infrastructure may reduce the requirement for co-location. It is also important to note that while physical co-location lets you bind data more easily, it is expensive to co-locate and move data. In addition, physical co-location brings with it concerns regarding ownership and access, scalability, and size. The impact of physical co-location as it relates to the governance of health and health-related information in a Health Information Repository (and the reverse impact of information governance on physical colocation) will also require additional analysis and investigation as Phase 2 of the project is undertaken.

Releasing data for secondary use may require the application of **de-identification services** (people, policies, procedures and technology) that work to preserve the privacy of individuals and groups of people according to provincial statutes, while maximizing the analytic usefulness of the data. The solution will leverage plans to proceed with a Provincial Health Analytics Network (PHAN) de-identification RFP on three prioritized data assets: Discharge Abstract Database - DAD, National Ambulatory Care Reporting System (NACRS), and Physician Billing - Claims data. The proof of concept will incorporate the secondary use de-identification levels identified by the Data and Analytics Working Group (DAWG) and the Privacy and Analytics Working Group (PAWG) with respect to risk stratification and risk management.

Work is already underway as part of PHAN to provide an **analytic portal** which will allow analysts to access data securely. This portal is being built from the technology platform and technological assets that currently exist in



AHS. Development of this portal will continue, extending the functionality of the joint PHAN Portal to utilize assets in the solution as well as increase access to researchers and the public in the future.

The **network model** required to move data from source to end point needs to be examined as the solution is designed and developed. As multiple data sources are added, it will be important to ensure that production systems supporting primary provision of care are not compromised. Network options available to support both primary and secondary data will be employed e.g. network segregation, cloud computing, potential use of private networks.

In addition to health data, the solution will include health-related data e.g. social services data, income data, environmental toxin data. Data will be tied together using sophisticated **inference engines** that feed precise registries and support standardization. More information on **precise patient registries** can be found in Appendix B.

Many valuable research data sets exist across Alberta; the data collected for specific research projects may have value beyond the projects. However, not all research data sets are integrated back into health system channels for secondary use. Research data sets are stored in a variety of places, many of which are not shareable with the larger health system; typically, research data sets end up being destroyed after the research has been completed. Solution infrastructure requirements will take the **needs of clinicians and researchers** into account as the solution is designed and developed:

- Researchers and clinicians involved in research require access to not only health data sets but also health-related data sets outside traditional health environments.
- At present, research teams control who accesses identifiable data within a research database; practices related to access and use need to be understood so they can be addressed by the solution.
- Each research database incurs costs for infrastructure; the solution needs to identify opportunities for collaboration.
- The concept of keeping and potentially sharing research data would require changes in process and requirements of ethics boards, especially with data collected using informed consent.
- Data required for research studies is influenced by the researcher's perspective and the hypothesis used to query the data; the solution will need to take researcher perspectives and research process into account during design and development. Infrastructure opportunities such as a **research data vault** may prove useful; a research data vault would contain enough information about a particular research data set to evaluate what additional secondary uses it may be considered for.

There may be value in sharing clinical trial data for secondary purposes; however, it will be important to investigate the complex challenges associated with sharing clinical trial data. For example, some clinical trials in Alberta have international data sets, containing only limited Alberta participant data.

Solution business and technical requirements were not examined in detail during Phase 1; Phase 2 will involve identifying the specific existing and net new technologies that exhibit the above noted features and fulfill solution stakeholder needs. Rapid scalability, performance, and stability will be key when deciding what technology to use.



Given the current fiscal situation and the investment made in the province to date, solution information will initially be managed by Alberta Health Services, with a recognition that concerns around governance of data assets and solution services will be addressed by the new SSC.

## Analytics

It is not enough to provide access to data. Abilities that transform raw data into information that generates insight are also required. Alberta's capacity in data and analytics is scattered and disconnected, with no overarching provincial strategy. The solution will provide services to develop the required capacity in a coordinated manner; solution analytics services will include:

- Registration and management of data users.
- Acquisition of prioritized data assets.
- Solution service planning to decrease duplication, improve analytics role clarity, and align analytics skill sets. This would include freeing up resources that could be re-invested into the solution and using PHAN analytics expertise, previously completed deliverables, and in flight projects to inform solution planning.
- Provisioning data access, including appropriate de-identification based on the user's role or intended use of the data. Principles related to use and disclosure have been clearly articulated by the Privacy and Access Working Group – these will assist to guide the operation of the solution.
- Consulting/advisory services for analysts and decision makers.
- Provisioning analytical tools e.g. software, hardware, server space.
- Learning and development offerings.
- Optimization of existing data holdings for analytical use.
- A metadata framework/strategy and maintenance of a metadata library.
- A data quality framework, aligning with and leveraging existing structures/groups.

A huge variety of analytics tools and applications will be available to the solution. Options related to 'bring your own' versus supplied analytics applications and tools will need to be further investigated in Phase 2. Duplication of tools and applications can result in significant cost to provincial analytics efforts and require careful evaluation.

Resources considerations for emerging open source analytics technologies also need to be taken into account as solution planning proceeds. Hadoop type solution design can be implemented at significantly less cost than relational databases; however, the Hadoop skillset in the marketplace is not yet mature (accessing in-demand, high-cost Hadoop resources may make solution implementation more costly at first).

### Acquisition and Prioritization of Data Assets

A preliminary, high level data asset inventory was completed in Phase 1. It is currently biased toward healthcare data as most individuals use the health system; the inventory needs to be expanded to include information about well individuals and additional health-related data assets.

Initially, currently existing high value data assets, including alignment with work already underway in PHAN, will be prioritized for inclusion in the solution. While this work is important, the true challenge lies in acquiring other categories of data assets noted in the Phase 1 data asset inventory:



- Health data assets currently available but not accessible
- Scattered clinical, research and quality improvement data sets
- Potential future health data sets
- Health-related data sets

These additional data assets will be prioritized early in Phase 2. Both value of the data asset and effort required for integration will be considered as data assets are prioritized. It is important to note that data users and participating organizations will have a variety of criteria that they use to assess data asset value. Given the complexities in prioritizing data assets for integration in the solution, data assets will be prioritized within three streams:

- Stream 1 - Integrate data assets that are readily available and easy to share.
- Stream 2 - Data assets that are available but challenges or barriers to integration exist e.g. culture, trust.
- Stream 3 - Data assets that represent a significant shift in health analytics - they are a 'game changer' e.g. previously non-existent (genomic) or inaccessible (electronic medical record (EMR)) data assets.

These three streams would run in parallel; data asset integration would proceed for each stream at the same time and assets within each stream would be prioritized for integration. This would ensure that data asset integration plans proceed for not only readily available assets but also for more complex assets in the last two streams. To ensure value to stakeholders, it will be important to outline and track high level measurable objectives of the solution. This will assist stakeholders to evaluate the prioritization and acquisition activities and ensure progress is being made. Example outcomes include, but are not limited to, the following:

- Number and type of data assets added over time
- Number of individuals and organizations accessing the data over time
- Number of analytic products created e.g. dashboards, research studies, task force reports.

## IV. Our Vision and Mission

### Vision

Develop, enhance, and sustain Alberta's health system through the advanced use of data in research, evaluation, and improvement initiatives.

Putting our data to work for a strong, healthy province.

### Mission

Through creation and development of a health information repository and secondary use data solution, we will:

- Respect the value of the data resting at our fingertips and ensure it is used to help others.
- Expand the value of the solution over time through not only use of the data and but also focused efforts to advance the mastery, autonomy, and purpose of all data contributors, users, and funders.
- Use data to continuously improve our research and health delivery efforts:
  - Improve the quality and safety of care provided to every patient, on a personal level.



- Improve the effectiveness and reduce the cost of delivering high quality care.
- Improve the overall health of the community while reducing the overall economic drain of poor health on the community.
- Accelerate the pace of clinical research.
- Enhance knowledge.
- Engage and request participation from individual citizens, health care providers, government, and researchers to continuously improve the solution's services and data assets.
- Create a data and analytics system capable of measuring, understanding, and designing interventions that create a high-performing health system.
- Share our learning, expertise, and resources.
- Ensure transparency regarding data use, security for the data, and privacy for the individual.

## V. Go Forward Plan

### Critical Recommended Actions

The following actions are critical to successful solution development and the successful drafting and approval of HIA Part 6.1 regulations. Additional, detailed recommendations suggested by the working groups are also available to guide Phase 2 project planning; they can be found in Appendix C.

1. Draft and approve without delay, the proposed regulation under Part 6.1 of the HIA. The solution cannot exist unless this regulation is put into force. Without it, privacy and access challenges will continue to plague researchers and quality improvement specialists. Program managers, policy makers, and funders will continue to operate with incomplete and inaccurate data sets. Bringing all necessary data into a single health information repository will vastly improve our efforts to advance health system quality. Inquiries made with comprehensive information will provide insight and wisdom previously inaccessible to decision makers. The project solution team will liaise with Alberta Health as they work to draft regulations under Part 6.1 of the HIA; however, regulation development will be dependent on appropriate prioritization and assignment of resources by the Government of Alberta. In order for this work to move ahead within the 10 month timeframe, work on the policy and regulations will need to begin immediately.
2. Engage the public. Albertans' support and trust is required for development and use of the solution. As secondary use of data is a complex subject, it will also be necessary to promote awareness and understanding of the purpose of the solution, and of the safeguards to be put in place with respect to personal information. Data comes from Albertans; as such, the views of Albertans with respect to the use of their data must underpin and inform integrated secondary use efforts.
3. Establish an effective governance structure, to begin functioning immediately. Comprehensive stakeholder engagement and collaboration will break down barriers, improve collective processes, speed access to data, and support a new way of doing business. The initial SSC must review and develop strategies and plans to address the following challenges within the first month it is appointed:
  - Balance between the privacy of individuals and the value of secondary data to the health of Albertans needs to be achieved. Privacy and access principles need to be consistently interpreted and actions regarding privacy rules need to be standardized. The steering committee will need to promote and support adoption of de-identification levels suggested in Phase 1 of the project. Appropriate implementation of these levels will streamline and speed access to data assets. The steering



committee will need to direct solution and project teams to devote special attention to operationalizing de-identification levels, including partial automation of access procedures accompanied by appropriate controls.

- Stakeholder concerns regarding loss of control and fears that data will be used inappropriately and out of context need to be addressed. With proper governance and data stewardship models, data contributors and users can contribute and integrate data with no loss of control. A communication and change management strategy will need to be developed to engage stakeholders in a trusted solution ecosystem where information exchange becomes a value added experience.
  - The use of secondary data for commercial purposes has raised concerns. Use of health data for commercial purposes is not permitted under Alberta legislation; at present, the only fees that can be collected related to secondary use health data are related to cost recovery i.e. a reasonable fee can be charged by the repository for access to the data. The solution implementation is centered on improving the quality of clinical care, achieving related social benefits such as improved person-centered government services, and increasing the speed with which research can be accomplished in the province of Alberta; the governance model will focus on these benefits. However, given that concerns regarding commercial use have already been encountered, it is likely that they will continue to come forward as the solution is designed, developed, and implemented. For this reason, the governance model and operational structure will need to develop mechanisms and processes that are capable of addressing, evaluating, and responding to data use scenarios such as commercial use requests.
  - A new legal structure may be required under the HIA Part 6.1; this structure must respect all related current provincial legislation.
4. Align primary and secondary data use strategies and initiatives. It is important to note that primary and secondary use data are the same data; mandates, objectives, and goals for data use in the province of Alberta must be aligned and coordinated. As such, appropriate decision-making powers are required to coordinate primary and secondary data governance; the proposed HIA amendments include providing the data governance committee under section 56.7 of the HIA with these decision-making powers. If these amendments are successfully implemented, the SSC will work to establish a reporting relationship and/or association with this overarching data governance committee. In addition, strong linkages must be formed between clinical information systems, electronic medical records, electronic health records, other primary data collection systems, and the solution.
  5. Leverage previous and current secondary use efforts. The rationale for net new solution resources and infrastructure needs to be considered carefully and aligned with in-flight projects and initiatives PHAN, Strategy for Patient Oriented Research (SPOR Support Unit), existing research repositories, primary care analytics, and other efforts must be respected and leveraged. Wherever possible duplication must be avoided and reduced; in particular, AHS and AH need to integrate and organize their analytic functions with clear role clarity aligning responsibility with skill capacity. Within this effort PHAN work can be leveraged but should not be viewed as the end point for secondary use in Alberta. By working together, Alberta can build on existing efforts to deliver not only an improved data and analytics system but also a more feasible and sustainable one.
  6. Establish solution funding, employing gated funding strategies that combine minimal net new funding with in kind contributions from stakeholder organizations. Funding models must be finalized early in Phase 2 and continuously evaluated as solution development proceeds; it will be important to demonstrate system cost savings generated through reduced duplication and waste. On the same note,



existing technology assets, initiatives, and procurements need to be leveraged to minimize expenses associated with demonstration projects and solution development.

7. Prioritize health-related data for inclusion in the solution as soon as possible. Many health and healthcare determinants lie outside the traditional boundaries of the healthcare delivery system. In addition to clinical and administrative health data, health related-data inclusive of health behavior, social and economic factors and the physical environment are required to achieve personalized medicine as well as community and population health benefits.
8. Establish a solution development agenda that includes demonstration projects. The solution requires real-world activities to anchor its development and incrementally test its governance and operations. Activities related to solution development (e.g. establishing the initial SSC, completing privacy impact assessments, drafting policy and regulations, establishing an initial operations team) cannot occur in isolation. Solution development will be coordinated with real-world demonstration projects that test and anchor the solution. It is essential that dedicated project and solution resources are available for these activities. Solution development activities cannot be coordinated and resourced by any one organization; as single entities, no one stakeholder organization can provide the required resources or combined perspective required to develop the solution single-handedly. In addition, an overarching assessment of stakeholder needs and requirements needs to be coordinated and actioned by a collaborative solution and project team focused on solution development objectives as opposed to separate and distinct stakeholder objectives.
9. Establish a timetable for migration to open source/Hadoop technology. The health information repository will be able to acquire data assets of a size and nature unlike any Alberta has used in the past. In order to realize the value of big data, migration to open source/Hadoop technology will be necessary. The solution will need to span the current relational database technologies and build an enabling environment where new non-SQL<sup>17</sup> technology can be assessed and implemented.

Continuing status quo is not an option. Our current ineffectiveness related to linking health data and health-related data prevents advancements in our understanding of the complex relationships between healthcare, health, and the social determinants of health. The technology, analytics, and reporting systems for linking this data exist today - technology is not the barrier. A lack of supporting HIR regulation, fractured data governance, and uncoordinated data and analytics practices prevent enhanced use of our own data. All that is required is a new way to collectively manage and share data. If we do not seize our current opportunity, redundant funding of uncoordinated analytics services and data management infrastructure will remain. Elevated costs, stifled research, suboptimal quality, and lost innovation will endure.

## Solution Development

A project team will work with newly established solution governance and operational roles to develop privacy impact assessments, solution policy, analytics services, and solution infrastructure to support several solution demonstration projects. The solution project team will also work with PHAN team members to complete PHAN de-identification and portal projects, leveraging their execution, components, and results whenever possible. Early demonstration project efforts will include exposing and amplifying existing data assets to a wide audience

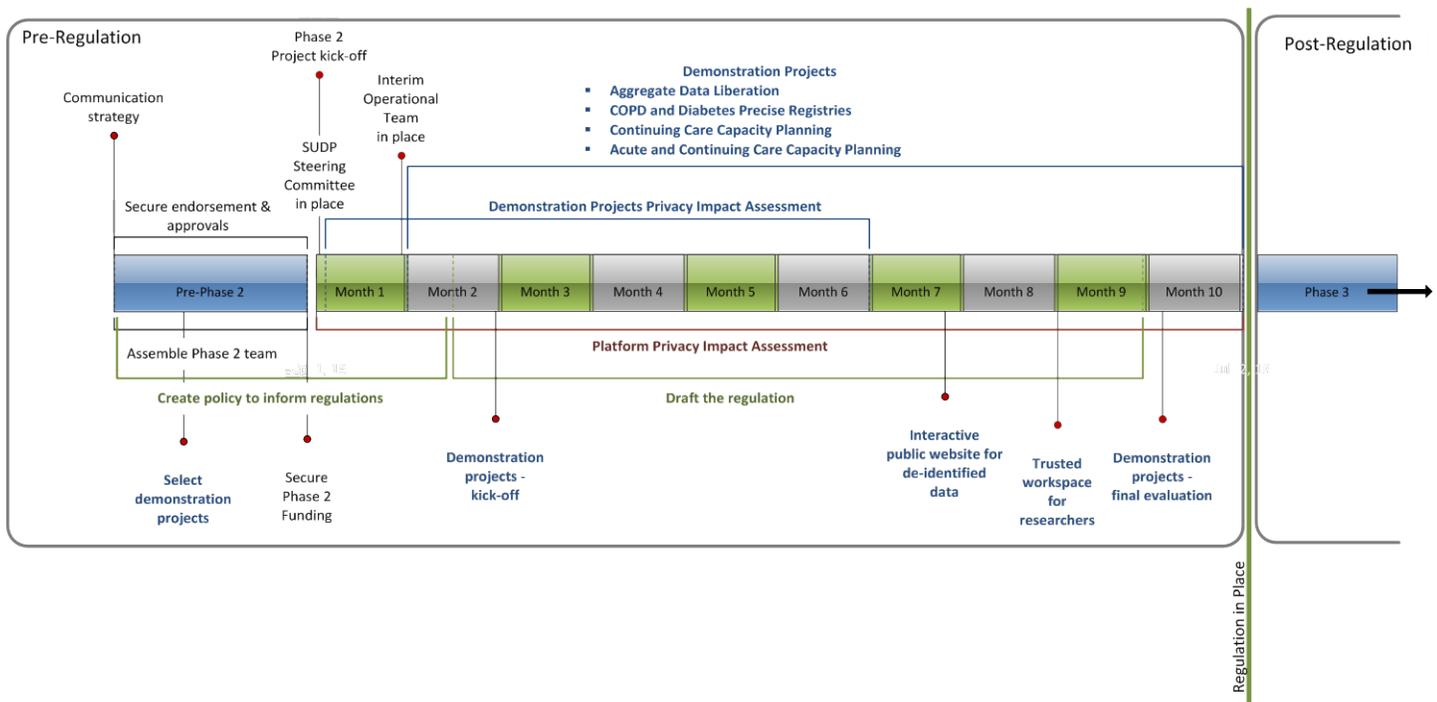
---

<sup>17</sup> Structured Query Language (SQL) – programming language for managing data in relational databases.



using appropriate de-identification strategies and tools. Solution policies, processes, infrastructure, services, and governance will be produced during the demonstration projects and solution development activities. As a result, the solution structure and functions will be available for immediate activation as soon as the regulations are put in force.

Efforts to develop collaborative recommendations have resulted in significant momentum to move to implementation. It is important to continue moving forward with solution development in advance of Part 6.1 regulations being put into force. Pre-regulation activities and deliverables will model, test, and inform the post-regulation phase.



## Demonstration Projects

The need for a secondary use data solution is immediate; there are many Alberta organizations whose current secondary data use needs are unmet. Data access, availability, and integration challenges prevent these organizations from using secondary data effectively and efficiently. Five demonstration projects were chosen to assist with solution development. These demonstration projects will foster collaboration among stakeholders as they work together to address identified secondary use challenges and problems and aim to facilitate data access and integration solutions or strategies within a 6 – 9 month timeline:

- Demonstration Project #1: Release 4 aggregate AHS reports via web interface
  - Total Cost and Prevalence of Disease
  - Community Level Healthcare Utilization Rates



- Surgical Utilization Rates
- Variations in Care Delivery
- Demonstration Project #2: Develop a precise patient registry<sup>18</sup> for Chronic Obstructive Pulmonary Disease (COPD).
- Demonstration Project #3: Develop a precise patient registry for Diabetes, including use of community and primary care data sources.
- Demonstration Project #4: Develop an acute and continuing care capacity planning model that allows us to better understand the capacity required between acute and continuing care to ensure timely transitions from hospitals and community to continuing care.
  - e.g. what health needs do Albertans have now and in the future? How can our policies, services, and programs adapt to meet these needs as they change over time?

The relationships between the many factors will be examined and data that is required to support decisions regarding changes to policies, programs, and services will be identified. The results will be used to create a model to assist decision makers and policy makers respond quickly and appropriately to the needs of Albertans

- e.g. how many acute care and long term care beds are required in Alberta over the next 5 years? 10 years? 20 years? What kinds of community services are required to support individuals living with chronic disease in the community?
- Demonstration Project #5: Expand data assets and related data integration for researchers participating in the Child and Youth Data Lab.

The demonstration projects chosen represent a balanced portfolio of projects that will support solution development. The suitability of the demonstration projects in meeting the needs of the developing solution were examined using demonstration project criteria noted below.

### Demonstration Project Principles

- All demonstration projects will engage researchers as key participants to ensure project data outputs meet end user data access and data use requirements.
- Existing provincial data and analytics assets, resources, in flight projects, and in flight procurements will be leveraged to the greatest extent possible.
- To prevent redundancy and waste, the demonstration projects will make every effort to align their activities with existing data and analytics projects.

### Demonstration Project Criteria

The following criteria were used to determine the suitability of a demonstration project for inclusion in Phase 2 of the project:

- **Prove Value through Data Influence, Data Liberation, and Data Integration** – will new data integration occur as a result of the project? Will new data assets be available for integration and analysis? Will data be liberated in new ways or to new end users?

---

<sup>18</sup> See Appendix B for a description of precise patient registries



- **Process Improvement** – are there opportunities to improve processes related to data access, acquisition, and de-identification?
- **Governance** – will the project assist the solution to further refine the governance roles and functions being put in place?
- **Operations Implementation** - will the project assist the solution to establish roles and functions required for solution operations?
- **Policy and Regulation Development** – will the project objectives and activities provide valuable input to the policies and regulations being developed to support the solution?
- **Privacy Impact Assessment** – will the project facilitate privacy impact assessment activities required for solution start-up and operations?
- **Leverage Existing Resources** – can proposed or existing investment be harnessed to benefit demonstration projects? Can investment dollars be saved by improving access to data and/or improving secondary data analysis processes?
- **Perception of Value** – does the project improve or address a problem that Albertans feel is worth the time, resources, and expertise invested in solving it?

## Funding

Phase 2 solution development funding will include net new and in kind contributions from Alberta stakeholder organizations. The first year solution development will require approximately \$3M net new funding; this will include demonstration projects and project resources dedicated to solution development. The net new and in kind FTE contributions are noted in the table below:

Resources	Annualized FTE
Net New Solution Development Resources	6.0 <sup>19</sup>
In Kind Solution Development Resources	[The project team is still gathering the data required for this calculation]
Net New Demonstration Projects Resources	5.8 <sup>20</sup>
In Kind Demonstration Project Resources	[The project team is still gathering the data required for this calculation]

The long term solution funding model will be finalized early in Phase 2. Preliminary work has provided the following synopsis of other national and international secondary data use initiatives.

Secondary Use Initiative/Organization	Cost Estimates*
Institute for Clinical Evaluative Sciences (ICES),	<ul style="list-style-type: none"> <li>• Infrastructure Investment n/a</li> </ul>

<sup>19</sup> 6 FTE for 12 months as the solution development team would need to be in place prior to and following the demonstration project teams

<sup>20</sup> 7 FTE for 10 months



Ontario	<ul style="list-style-type: none"> <li>• \$ 21.9M Annual Operating Cost</li> </ul>
Manitoba Centre for Health Policy (MCHP), Manitoba	<ul style="list-style-type: none"> <li>• \$3.5M (1990) Infrastructure Investment</li> <li>• \$5.7M Annual Operating Costs (AOC)</li> <li>• 40% of AOC funded by Manitoba Government</li> <li>• 60% through awards and grants</li> </ul>
Health and Social Care Information Centre (HSCIC), UK	<ul style="list-style-type: none"> <li>• \$63M for development and maintenance over 5 years</li> <li>• Average AOC \$12.6M</li> </ul>
British Columbia Centre for Data Innovation (BC CDI), British Columbia (Proposed)	<ul style="list-style-type: none"> <li>• Estimates only – Not implemented yet</li> <li>• \$7.0 - \$10.5M One-time cost</li> <li>• \$4.0 – 6.0M AOC</li> </ul>

\* Estimates are based on either publicly available information or preliminary interviews with the organizations. The estimates have not been validated with the source organizations. All figures have been converted to CAD\$ for comparison purposes.



## Appendix A – Project Methodology

### Working Group Objectives and Activities

#### Description

Six working groups were formed to provide recommendations and expert guidance in preparation for development of a secondary use data solution in Alberta. A short description of each working group is provided below:

**Public Engagement Working Group** - represented the interests of Albertans regarding the use of their data. The working group's main objective was to comment on the need for and perceived value of a secondary data solution and to determine what strategies might be required to receive ongoing feedback from Albertans on solution-related topics.

Co-chairs: Sharon Nettleton and Deborah Prowse

**Privacy and Access Working Group** - examined the balance between access to data and the privacy of Albertans' information. This group developed principles related to privacy and access and identified and mitigated risks to privacy. In conjunction with the Data and Analytics Working Group, they also developed de-identification levels to support appropriate access to data.

Co-chairs: Mary Marshall and Lawrence Richer

**Data and Analytics Working Group** - identified a preliminary data asset inventory for secondary use in Alberta as well as a preliminary analysis of the analytics tools and services required by solution end users. The working group suggested a prioritization schema for inclusion of data assets in the solution.

Co-chairs: Kathryn Todd and Stafford Dean

**Infrastructure and Applications Working Group** – commented on the anticipated infrastructure required to house, maintain, and share data in the solution and the impact that business requirements and technology trends may have on solution infrastructure decisions.

Co-chairs: Susan Anderson and Penny Rae

**Enabling Legislation Working Group** – identified and recommended legislative options to support a health information repository (HIR) and outlined the processes required to draft regulations under HIA Part 6.1.

Chair: Roger Palmer

**Governance and Structure Working Group** – developed a governance model designed to represent the interests of all stakeholders and commented on the operational and legal structures that may be required as the solution is designed, developed and implemented.



Chair: Dale Sanders

### Working Group Activities

Each working group met a number of times between March 12 and June 30, 2015. Working group activities included investigating high level requirements for solution development and implementation, discussing issues and concerns related to secondary data use, determining the impact that Alberta-specific factors may have on solution creation and ongoing operations, and recommending strategies and courses of action for Phase 2 of the project.

Each working group produced a final report that is location on the project [SharePoint site](#).

### Participating Groups and Organizations

Over 60 individuals from the following organizations participated in the working groups:

- Albertans
- Academic Health Network
- Alberta Medical Association
- Alberta Innovates Health Solutions
- Alberta Health
- Alberta Health Services
- Health Quality Council of Alberta
- Health Research and Innovation Collaboratory
- University of Alberta
- Office of the Information and Privacy Commissioner
- Cybera
- Strategy for Patient Oriented Research
- Alberta Innovates Technology Futures
- University of Calgary
- College of Physicians and Surgeons of Alberta
- Alberta Centre for Child, Family and Community Research
- Alberta Bone and Joint Health Institute

### Members

Working group members are recorded on the final working group reports posted on the project [SharePoint site](#).

### Literature Search

A preliminary literature search was prioritized and completed for the PEWG, GSWG, and ELWG. The search results and select articles/resources were provided to these working group members; however, the results have not yet been reviewed for the purpose of extracting data using a formal literature review process. The extraction and review processes as well as the literature searches for DAWG, PAWG, and IAWG were deferred to Phase 2 of the project.



## References

A complete [list](#) of the literature search results is available on the project SharePoint site.

## Project Resources

Document Name	Source/Location	Link/URL if applicable
Project Charter v0.9	Secondary Use Data Platform (SUDP) Project SharePoint site	<a href="#">Project Charter v0.9</a>
Secondary Use Data – SUDP Background	SUDP Project SharePoint site	<a href="#">SUDP Background</a>



## Appendix B – Precise Patient Registries

The solution will rely heavily on development of precise patient registries. Precise patient registries differ from traditional registries:

- Content, data elements, metadata, and associated technical and terminology standards are often controlled in a traditional registry by an external organization such as the government or a society. As a result, the information contained in a traditional registry is not easily tailored to local health analytics functions.
- In a traditional registry, contributors have to adhere to standardized definitions that may not provide enough information for detailed local population and cohort analyses. The data is not as precise as if it were maintained by a local health organization with access to specific and comprehensive patient cohort data sets. As a result, metrics and outputs from a traditional registry are often limited to process metrics as opposed to outcome measures or more sophisticated analyses.

In contrast, a precise patient registry is a database designed to store and analyze information about the occurrence and incidence of a particular disease, procedure, event, device, or medication and for which, the inclusion criteria are defined in such a manner that minimizes variability and maximizes precision of inclusion within the cohort. Precise patient registries are necessary to complete comprehensive and accurate analyses for secondary research and quality improvement initiatives. Without precise definitions and registries of patient types, it isn't possible to support precise clinical research, conduct precise comparisons, develop personalized healthcare and facilitate predictable clinical outcomes. The value of precise patient registries also lies in their ability to reduce labor and increase consistency of reporting associated with chronic condition registries; registries don't just pull together data sets, they also include the construction of precise indicators for health system and research measurement functions.

Precise patient registries will be foundational for the solution as all solution activities will be built on standard data definitions and precise content. Alberta has a good registry base to work from:

- Alberta Health Services has created databases with approximately 250 conditions using International Classification of Disease codes; specificity and sensitivity of the information captured for inclusion in these databases is very high.
  - The solution can add value to these databases by de-identifying and exposing additional data and ensuring comprehensive data elements and metadata are represented from as many data sources as possible.
- There is a shared data model program in AHS; cohort and registry designs are identified as opportunities arise e.g. APPROACH<sup>21</sup> database in cardiac research. The researchers are provided an analytics workspace.
  - The solution will be able to simplify processes for access to data and provide an enhanced shared data model environment. At present, researchers come through high end analysts which results in a process bottleneck; the solution will enable researchers to work with de-identified data themselves.

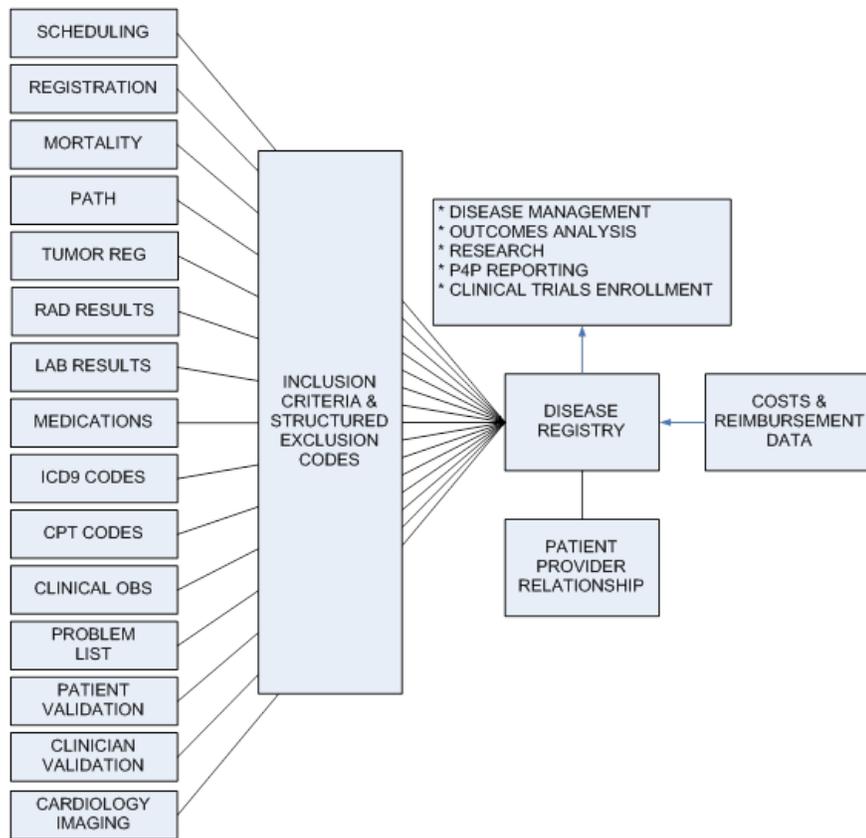
---

<sup>21</sup>Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease <http://www.approach.org/>



- The solution may also be able to improve the precision of the registry.

Precise patient registries are built using inference engines that employ algorithms to automatically search for specific information about a cohort, disease, event, pathway, or other health system grouping that is of interest to researchers, policy makers, and quality improvement experts (as an example, a disease registry is depicted in the following diagram). Inference engines apply specific inclusion and exclusion criteria that increase the comprehensiveness and precision of the registry beyond that available in traditional registries.





## Appendix C - Phase 1 Recommendations

### Regulation Recommendations

Drafting and enactment of a regulation under HIA Part 6.1 is the foundational activity that will drive establishment of a HIR to support secondary data use in Alberta. As the heart of the solution, the HIR is what will provide the impetus to take secondary data use to the next level in Alberta. The following key recommendations form the basis for policy that will inform drafting of the regulation. No statutory changes will be required to the HIA in order to implement the HIR:

- The interests and priorities of the citizens of Alberta will be central to the purpose and function of the solution.
- The solution will be established as a multi-stakeholder, multi-organizational collaborative.
- While primary and secondary use data are the same data, they are used for different purposes and have distinct data management needs. As such, appropriate decision-making powers are required to coordinate primary and secondary data governance and provincial initiatives. It is recommended that the mandate of the HIA legislated EHR Data Stewardship Committee be expanded in order that its role include providing guidance and input (strategic and otherwise) to the SSC.
- The solution will function as a shared-governance entity between data funders, users, and contributors. A SSC will provide the operational governance structure that facilitates achievement of solution goals and objectives.
- The initial SSC will be an appointed, representative/competency-based steering committee. It will be supported, as needed, by implementation working groups.
- The health information repository should be a new data entity class, 'repository', with rights, duties, powers, and responsibilities separate and distinct from a custodian or information manager.
- The health information repository will collect, use, and disclose health information for the purposes of strategy development, planning, policy making, service delivery, care, research, monitoring, and economic development to benefit Alberta and Albertans.
- The solution will include clinical, socio-economic, multi-jurisdictional government, administrative and self-reported data that is aggregated and integrated.
- Per the HIA, secondary use data will be considered as either non-identifying or individually identifying. Non-identifying means that the identity of the individual who is the subject of the information cannot be readily ascertained.
- Non-identifying data will be categorized into minimum, moderate, and maximum risk of identification or re-identification.
- The solution should be funded to include core data content, business and technical infrastructure, and applications.
- Governance will operate at arms-length from any one of the involved stakeholders. While data contributors and users will have an opportunity to actively contribute to governance of the solution and its operations, no single stakeholder will make decisions or set strategic priorities in isolation from the others.



- As the solution matures, the SSC will evaluate the cost and benefits of establishing an alternate solution entity and governance structure, as appropriate.

### Implementation Recommendations

In addition to regulation drafting and enactment, a number of set-up and operational activities must be undertaken to support implementation of the solution. During Phase 1 of the project, six working groups put forward recommendations to guide implementation of the solution; these recommendations were themed and categorized. They have provided the foundation for the Phase 2 project plan and when implemented, will guide stakeholders through to successful implementation of the solution. Working group recommendations were analyzed and themed in the following table. Each supporting recommendation is followed by the acronym of the working group (or steering committee) that created the recommendation:

- PEWG – Public Engagement Working Group
- GSWG – Governance and Structure Working Group
- IAWG – Infrastructure and Applications Working Group
- DAWG – Data and Analytics Working Group
- PAWG – Privacy and Access Working Group
- ELWG – Enabling Legislation Working Group
- SUDPSC – Secondary Use Data Platform Steering Committee

It is recommended that:	Supporting Working Group Recommendations
<p>A public engagement strategy for the solution be co-developed with members of the public. The strategy will be based on best practices and aligned with other relevant provincial health information initiatives.</p>	<ul style="list-style-type: none"> <li>• That a public engagement strategy be developed as part of Phase 2, based on the considerations and issues outlined in this report. The engagement strategy should be based on best practices identified from a literature review and the experience of other jurisdictions, including the expertise of patient engagement leaders. It should include different levels of engagement from general awareness and communication to full participation of citizens in solution operation and governance. The public engagement strategy should be co-developed with members of the public. PEWG</li> <li>• That the engagement strategy build on and be combined with other current and relevant provincial health data information initiatives such as the electronic health record, personal health portal, and primary use health data systems. PEWG</li> </ul>
<p>The solution ensures adequate privacy protection for individuals at the same time recognizing the benefits for the appropriate secondary use of health data for research,</p>	<ul style="list-style-type: none"> <li>• The working group recommends that all ten of the privacy and access principles be included as part of solution development, governance, and implementation, including representation within the proposed HIA 6.1 Regulations. Specifically, the principles should be considered during the policy development for the HIA 6.1 Regulations, with consideration as to what content that should become part of the HIA 6.1 Regulations. A description of each principle and associated impact/factors is provided in Appendix A. PAWG</li> <li>• That the needs and interests of Albertans related to privacy, security,</li> </ul>



It is recommended that:	Supporting Working Group Recommendations
<p>innovation and quality assurance. Specifically, the solution will develop and support robust de-identification practices to specify the level of risk that is acceptable given the potential benefit of any particular research.</p>	<p>consent, and the protection of personal information are fundamental requirements that must be considered in the development and operation of the solution. PEWG</p> <ul style="list-style-type: none"> <li>• The working group recommends that non-identifying and identifying HIA definitions and practices be operationalized to the extent possible, and that additional non-identifying levels be operationalized by adoption of minimum, moderate and maximum de-identification criteria. HIA definitions and practices may need to be reviewed and modified in the future to fully support the secondary use and disclosure of health information by the repository. PAWG</li> <li>• Once de-identification stratification guidelines are developed for the data solution, it is recommended that               <ul style="list-style-type: none"> <li>○ Software to assist with understanding the risk of re-identification be investigated</li> <li>○ Operational definitions be developed to facilitate maximal automation of access to specific levels of de-identified data. PAWG</li> </ul> </li> <li>• Proceed with a de-identification proof of concept on the three PHAN prioritized data assets: Discharge Abstract Database - DAD, National Ambulatory Care Reporting System (NACRS), and Physician Billing - Claims data. Within the proof of concept, incorporate the needs identified by the Data and Analytics Working Group (DAWG) and the Privacy and Analytics Working Group (PAWG) with respect to privacy and access risk stratification and risk management. IAWG</li> </ul>
<p>All relevant health and health-related data assets be included in the solution (clinical, socio-economic, multi-jurisdictional government, administrative and self-reported data). A multi-stakeholder process will be developed to prioritize data assets for integration.</p>	<p>Data Assets</p> <ul style="list-style-type: none"> <li>• Identify data asset gaps i.e. data that does not exist and that is required; build the strategy and implementation plan to fill in the gaps. DAWG</li> <li>• Implement the Patient Reported Outcomes Measures Strategy (PROMs).<sup>22</sup> DAWG</li> <li>• Include research data sets as part of the solution, including development of the research data vault concept, but in a later phase. DAWG</li> </ul> <p>Data Asset Prioritization</p> <ul style="list-style-type: none"> <li>• Develop data asset prioritization methods               <ul style="list-style-type: none"> <li>○ Make ‘high value, already accessible data’ available as a first priority.</li> <li>○ Ensure alignment with The Provincial Health Analytics Network (PHAN) priorities already underway. DAWG</li> </ul> </li> <li>• Fast track acquisition of clinical data from physician electronic medical</li> </ul>

<sup>22</sup> <http://www.cihr-irsc.gc.ca/e/46848.html>



It is recommended that:	Supporting Working Group Recommendations
	<p>records e.g. primary care data, specialists working in hospitals. DAWG</p> <ul style="list-style-type: none"> <li>• Align the solution’s strategic data acquisition plan to the Information Management/Information Technology roadmap and clinical information system implementation. DAWG</li> <li>• Prioritize data acquisition in three streams               <ul style="list-style-type: none"> <li>○ Stream 1 - Integrate data assets that are readily available and easy to share.</li> <li>○ Stream 2 - Data assets that are available but challenges or barriers to integration exist e.g. culture, trust (such as EMR)</li> <li>○ Stream 3 - Data assets that represent a significant shift in health analytics - they are a ‘game changer’ e.g. previously non-existent (genomic) or inaccessible (immigration and education files) data assets DAWG                   <ul style="list-style-type: none"> <li>▪ Expand the audience for data asset prioritization beyond the Data and Analytics Working Group in Phase 2 of the project. DAWG</li> </ul> </li> </ul> </li> </ul>
<p>A core set of analytics functions, tools, and resources be designed and implemented as part of solution operations.</p>	<p><b>Analytics</b></p> <ul style="list-style-type: none"> <li>• Appoint a solution role accountable to assist with integration activities e.g. steering committee chair. This individual would assist with integrated solution service planning to decrease duplication, improve analytics role clarity, and align analytics skill sets. This would include freeing up resources that could be re-invested into the solution. DAWG</li> <li>• Develop the core services the solution would provide:               <ul style="list-style-type: none"> <li>○ Register/manage data users e.g. analysts, data management staff, decision makers</li> <li>○ Optimize existing data holdings for analytical use e.g. master data elements, data governance, quality, documentation, auditing, privacy</li> <li>○ Provision data access i.e. includes appropriate de-identification based on the user’s role or intended use of the data</li> <li>○ Provision analytical tools e.g. software, hardware, server space</li> <li>○ Consulting/advisory services for analysts and decision makers e.g. question formulation, advising on the most appropriate analytical methods for the problem at hand, determination of analytical rigor required for the question/purpose (when is the data good enough?), recommended outputs to serve the needs/questions</li> <li>○ Acquire existing data from various agencies DAWG</li> </ul> </li> <li>• Develop a metadata framework/strategy               <ul style="list-style-type: none"> <li>○ Leverage existing metadata</li> <li>○ Maintain a metadata catalog with data features, elements and terminologies. DAWG</li> </ul> </li> </ul> <p><b>Data Quality</b></p>



It is recommended that:	Supporting Working Group Recommendations
	<ul style="list-style-type: none"> <li>• Develop a data quality framework, aligning with and leveraging existing structures/groups DAWG               <ul style="list-style-type: none"> <li>○ Develop data quality evaluation tools</li> <li>○ Periodically evaluate data quality for reporting consistency and standard of data collection</li> <li>○ Document level of data quality in each dataset. DAWG</li> </ul> </li> <li>• Link datasets to enrich data quality overall. DAWG</li> </ul>
<p>The solution receive oversight from a committee that has broad health system representation and is focused on data governance.</p>	<ul style="list-style-type: none"> <li>• The solution receive oversight from a committee that has broad health system representation and is focused on data governance; existing committees recommended to serve this role include either the Health Information Governance Committee or the Electronic Health Record Data Stewardship Committee. The solution governance will report to Alberta’s multi-disciplinary data stewardship committee as defined in section 56.7 of the HIA. GSWG</li> <li>• It is recommended that the minister appoint members to the multi-disciplinary data stewardship committee defined in s.56.7 of the Act and give it a mandate to oversee an integrated health information system for Albertans. ELWG</li> <li>• It is recommended the data stewardship committee be given specific authority through the regulation to oversee Health Information Repositories under Part 6.1 of the Act. ELWG</li> <li>• It is recommended that the data stewardship committee, in consultation with the Privacy Commissioner, provide direction to the operating procedures of the HIR so that privacy is protected as required by law but the public good is served in the improvement to patient care and patient safety. ELWG</li> <li>• It is recommended that the data stewardship committee monitor the operation of the HIR and require all requests for data access to be fulfilled within four months of the initial request being received. ELWG</li> <li>• It is recommended that the data stewardship committee take into account the key findings of the Council of Canadian Academies’ report in creating the privacy and ethics requirements of the solution. ELWG</li> </ul>
<p>Stakeholders leverage existing technologies and infrastructure, including in-flight projects and procurements, to design and develop a big data analytics solution capable</p>	<ul style="list-style-type: none"> <li>• Create an environment to support high performance data access regardless of the location of the data or the analyst; ‘compute as close to the data’ as possible. IAWG</li> <li>• Utilize the in-flight Alberta Health Services open source technologies procurement to explore the capability of moving data quickly from source to analytics, including an examination of transition needs for the potential shift from relational database technologies and skills to open source technologies and skills. IAWG</li> </ul>



It is recommended that:	Supporting Working Group Recommendations
<p>of integrating data from health and health-related source systems.</p>	<ul style="list-style-type: none"> <li>• Examine network capacity required to support both primary and secondary data, ensuring that production systems supporting primary provision of care are not compromised as the solution is implemented. IAWG</li> <li>• Review technical solutions for agility to ensure secondary use metadata and data element structures are able to support potential shifts in analysis needs over time. IAWG</li> <li>• To the greatest extent possible, automate data sources to feed directly into solution. IAWG</li> <li>• Separation principle or practice is recommended i.e. there are two databases - one with identifying information for integration and linkage and the other with encoded (i.e. unique identifier) non-identifying information. PAWG</li> </ul>
<p>Alberta Health draft and put into force regulation to establish a health information repository. The regulation will be based on policy emerging from SUDP working group recommendations.</p>	<ul style="list-style-type: none"> <li>• It is recommended that Alberta Health be charged with drafting a Health Information Repository (HIR) regulation and taking that regulation through public consultation and government’s approval processes by June 2016. ELWG</li> <li>• It is recommended that the recommendations and findings of the secondary use data platform project working groups are included in solution policy development that will then be used inform HIA Section 6.1 regulations. ELWG</li> <li>• SUDP governance requirements defined by the Governance and Structure Working Group be articulated in the proposed Regulations for Section 6.1 of the Health Information Act (HIA). GSWG</li> </ul>
<p>The initial solution governing entity be composed of 7 or 9 representative stakeholders that employ formal governance practices as equal participants.</p> <p>Solution governance evolve to an entity with formal accountabilities</p>	<ul style="list-style-type: none"> <li>• Plans for evolution of solution governance take into account the potential shift from a SSC to some other type of entity that provides strategic guidance and direction to the solution. The group notes that the entity may be a board, a board of directors, or another entity that functions similarly. While the name and structure of the entity will be determined as the solution evolves, the working group recommends that the entity exhibit board-like characteristics that will enable and guide future state solution functions. GSWG</li> <li>• A representation-based SSC be formed as the solution is implemented, progressing more and more to a competency-based SSC as the solution evolves. GSWG</li> <li>• The SSC be composed of a maximum of 7 or 9 individuals. GSWG</li> <li>• The SSC be ultimately accountable to the Minister and Albertans. GSWG</li> <li>• The proposed solution be linked to governance and operations across participating organizations, where appropriate e.g. PHAN. GSWG</li> </ul>



It is recommended that:	Supporting Working Group Recommendations
<p>such as a board or board of directors<sup>23</sup>.</p>	<ul style="list-style-type: none"> <li>• Initial SSC members demonstrate the following competencies:               <ul style="list-style-type: none"> <li>○ Technology startup leadership</li> <li>○ Marketing and communications</li> <li>○ Government affairs</li> <li>○ Financial management and fund raising</li> <li>○ Healthcare analytics and data management</li> <li>○ Population and community health</li> <li>○ Clinical research</li> <li>○ Clinical quality improvement</li> <li>○ Healthcare economics</li> <li>○ Analytic technology management</li> <li>○ Governance and regulation of health professionals GSWG</li> </ul> </li> <li>• Initial SSC membership include the following organizations:               <ul style="list-style-type: none"> <li>○ Alberta Health</li> <li>○ Alberta Health Services</li> <li>○ Alberta Pharmacists’ Association</li> <li>○ Alberta Innovates Health Solutions</li> <li>○ Alberta Medical Association</li> <li>○ Public (2 members)</li> <li>○ Alberta Academic Health Network</li> <li>○ Appointee from Research Community SUDPSC</li> </ul> </li> <li>• Members of the SSC be equal participants with formal representation via Robert’s Rules or similar governance practices. GSWG</li> <li>• That the solution governance structure include members of the public. PEWG</li> </ul>
<p>Solution design take into account high level requirements noted by the working groups and involve members of the public.</p>	<ul style="list-style-type: none"> <li>• That the solution design and development teams include members of the public. PEWG</li> <li>• Consider multiple perspectives in the design of the solution; the solution will be used by multiple stakeholders for multiple purposes, each with their own privacy and access needs. PAWG</li> <li>• Include mechanisms for data quality improvement and evaluation as part of the solution. DAWG</li> <li>• Define the analytic functions required to support a high performing health system, one that learns through research and quality improvement efforts. DAWG</li> <li>• Develop a process to advance a data-driven culture across all levels, starting with leadership, which assists individuals using health and health-related</li> </ul>

<sup>23</sup> This recommendation was later qualified by the steering committee: in order for the solution to evolve to an independent entity with an alternate governance entity such as a board or board of directors, it must first demonstrate value for its stakeholders and a clear need for independence.



It is recommended that:	Supporting Working Group Recommendations
	<p>data to ask and answer better and deeper questions. DAWG</p> <ul style="list-style-type: none"> <li>• Create a development methods and training hub for analytics practitioners across the province:               <ul style="list-style-type: none"> <li>○ Training curriculum e.g. how to use various datasets, tools, analytical methods</li> <li>○ Analytics resources e.g. software tools, computer hardware</li> <li>○ Metadata library</li> <li>○ Regularly-scheduled data/analytic ‘rounds’ i.e., presentations from analytics practitioners and/or business leaders demonstrating how analytics were used to inform a business/clinical issue</li> <li>○ Self-organized communities of practice and learning</li> <li>○ Job rotation and field replacement assignments/secondment opportunities</li> <li>○ Coaching/mentoring opportunities DAWG</li> </ul> </li> <li>• Access and leverage PHAN deliverables as part of solution planning. While lessons learned are not necessarily documented, PHAN sponsors and business managers working as part of Phase 2 solution development can share lessons learned in the solution’s working groups. DAWG</li> <li>• There is discrepancy among health system stakeholders as to how they deal with consent and masking of personal health information e.g. technology capabilities, business process variability and inconsistent practices. The solution will need an integrated approach to provide guidance to all custodians and contributors; however, development of a comprehensive consent and masking approach was not feasible within the Phase 1 timeline. The working group recommends that further investigation regarding this principle be undertaken in Phase 2 of the project. PAWG</li> </ul>
<p>Solution funding processes assess current analytics investments and capacity, identify opportunities to reduce duplication and waste, and redirect savings to current data and analytics gaps.</p>	<ul style="list-style-type: none"> <li>• Assess current investments and capacity in data management and analytics, tools, resources, and infrastructure across Alberta organizations serving health care functions. DAWG</li> <li>• Identify opportunities to reduce spending and eliminate duplication, redirecting the savings to current data and analytic gaps. DAWG</li> <li>• Build analytical capacity and share analytical technologies to optimize data access and analysis efficiency. DAWG</li> <li>• Utilize existing projects and funding currently in place.               <ul style="list-style-type: none"> <li>○ Continue development of the PHAN portal.</li> <li>○ PHAN de-identification project leverage the de-identification and access levels developed by the PAWG and DAWG. IAWG</li> </ul> </li> </ul>
<p>A Director lead the operations of the</p>	<ul style="list-style-type: none"> <li>• A Director or Chief Information Officer lead an operations group of secondary use data experts, including solution staff and seconded staff from data contributors, users, and funders. GSWG</li> </ul>



It is recommended that:	Supporting Working Group Recommendations
<p>solution.</p> <p>A Chief Privacy Officer be put in place for the solution.</p>	<ul style="list-style-type: none"> <li>It is recommended that a chief privacy officer be in place for the solution. PAWG</li> </ul>
<p>Both a project and a public communication strategy be developed to support subsequent implementation phases.</p>	<ul style="list-style-type: none"> <li>Recommend creating a communication strategy regarding implementation of the solution. It is important to present information to the public regarding privacy of and access to secondary use data. Key messages need to be developed regarding the use of identifiable data via the solution; in addition, standard practices and safeguards need to be explained in lay terms. PAWG</li> <li>That a clear understanding of the purpose, background, benefits, and requirements of the solution be developed and shared as part of the public engagement strategy; this should include a business case composed of a budget, systems design, and timelines for solution testing and implementation. PEWG</li> </ul>
<p>Implementation proceed with several Phase 2 demonstration projects focused on quick delivery of data assets, with concurrent development of solution governance and operations structures.</p>	<ul style="list-style-type: none"> <li>Act quickly to deliver data assets and produce value as soon as possible. The shorter the time period for solution deliverables, the better. DAWG</li> <li>Provide dedicated resources to design and implement the secondary use data solution technical requirements. IAWG</li> <li>Expose existing data using the future state governance model, at a level that allows open access to large de-identified data sets. IAWG</li> <li>It is recommended that provincial stakeholders continue secondary data use in-flight initiatives and, where possible, immediately carry out additional secondary data use initiatives as the solution’s governance is established; the focus of initiatives will be to make data holdings accessible to data users within the current system of law and regulation, even in advance of the regulations being approved. ELWG</li> </ul>



## VII. Appendix D - Abbreviation

AH	Alberta Health
AHS	Alberta Health Services
AOC	Annual Operating Costs
APPROACH	Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease
BC CDI	BC Centre for Data Innovation
BORN	Better Outcomes Registry & Network (BORN)
CIO	Chief Information Officer
COPD	Chronic Obstructive Pulmonary Disease
DAD	Discharge Abstract Database
DAWG	Data and Analytics Working Group
EHRDSC	Electronic Health Record Data Stewardship Committee
ELWG	Enabling Legislation Working Group
EMR	Electronic Medical Record
FTE	Full Time Equivalents
GSWG	Governance and Structure Working Group
HIA	Health Information Act
HIGC	Health Information Governance Committee
HIR	Health Information Repository
HQCA	Health Quality Council of Alberta
HSCIC	Health and Social Care Information Centre
IAWG	Infrastructure and Applications Working Group
ICD 9	International Classification of Diseases version 9
ICES	Institute for Clinical Evaluative Sciences
IT	Information Technology
M	Million
MCHP	Manitoba Centre for Health Policy (MCHP)
NACRS	National Ambulatory Care Reporting System
PAWG	Privacy and Access Working Group
PEWG	Public Engagement Working Group
PHAN	Provincial Health Analytics Network
Pop Data BC	Population Data British Columbia
SPOR	Strategy for Patient Oriented Research
SQL	Structured Query Language
SSC	Solution Steering Committee