# Population Health Digital Research Infrastructure: Building Up and Out from a Learning Collaborative Network

A White Paper submitted by Health Data Research Network Canada (HDRN Canada) to the New Digital Research Infrastructure Organization (NDRIO)

December 9, 2020

Authors: P. Alison Paprica,\* Donna Curtis Maillet, Lindsey Dahl, Brent Diverty, Jean-François Ethier, Frank Gavin, Alan Katz, Lisa M. Lix, Don MacDonald, Ted McDonald, Kimberlyn McGrail, Hude Quan, Robyn Rowe, Michael J. Schull, Catherine Street, J. Charles Victor, Jennifer Walker, Juliana Wu, Nicole Yada

<sup>\*</sup> The designated contact person for this White Paper is Alison Paprica (<u>alison.paprica@ices.on.ca</u>); other authors are listed in alphabetical order by last name

## ABOUT HDRN CANADA

Incorporated as an independent non-profit research institute in January 2020, Health Data Research Network Canada (HDRN Canada) is a pan-Canadian network of organizations that are working together to facilitate and accelerate multi-jurisdiction research in Canada (see Figure 1). The organizations that make up HDRN Canada hold and manage population-wide health and health-related data infrastructure that is used by hundreds of expert analysts and researchers across Canada. Administrative health data, which include longitudinal information about entire populations, are invaluable for high impact health research, particularly when they are linked with social and clinical data. For example, using the population-wide administrative data accessible through HDRN Canada Organizations, researchers can: study diseases of low prevalence, measure the safety and effectiveness of drugs and healthcare interventions, and compare health care system outcomes across provincial and territorial boundaries. However, in Canada, each province and territory is responsible for administering its own publicly funded health care program. This has resulted in multiple sets of administrative health data, often with differences in structure and format of data holdings and/or policies, that pose barriers to multi-jurisdictional research. Under

HDRN Canada, national, provincial and territorial organizations which hold population-wide health and health-related data have come together to make it easier for researchers, analysts and policy makers to use data from multiple provinces and territories for research, benchmarking, health system planning, and innovation. HDRN Canada is a distributed network that distributes its funding in a way that respects each member organization's local context and policy environment (e.g., not requiring jurisdictional or institutional policy changes). HDRN Canada is open to ideas, input, and opportunities, and committed to leveraging and sharing resources and practices whenever possible.



#### CURRENT ISSUES IN NATIONAL DRI AND HOW HDRN CANADA IS ADDRESSING THEM

With few exceptions, including, notably the *SPOR Canadian Data Platform* grant led by HDRN Canada,† health and social data have not had the national support that digital research infrastructure (DRI) requires. There have not been sufficient or stable funds for necessary hardware, software, cybersecurity tools, and personnel to integrate, manage, and provide secure access to data. In the absence of DRI funding and leadership, research datasets and limited project-specific data infrastructure are being established by research teams that are primarily focused on using the data for their own research studies vs. establishing DRI that can be used by others. Project grants do not generally include funding for post-project maintenance of datasets, or the creation of metadata that would facilitate the use of datasets by other research teams. Additionally, there is little funding to support shared infrastructure and partnerships between data holders, even though data assets are known to have more value when they are linked and combined. Overall, the lack of coordination and leadership for DRI encourages:

(i) short-term and narrow planning, (ii) new data collection and duplication of efforts over re-use and /or augmentation of existing data, and (iii) competition between research teams and institutions rather than collaboration and sharing. HDRN Canada was created to address these issues in the population health and health services data space.

A major accomplishment during HDRN Canada's first year of operations was the creation of the online Data Access Support Hub (DASH).<sup>3</sup> The DASH serves as a single point of access and information portal for researchers seeking to perform multi-jurisdictional research using population health data, and includes three interactive inventories:

<sup>†</sup> SPOR is the Strategy for Patient-Oriented Research, a program established by the Canadian Institutes of Health Research (CIHR)

- 1. An inventory of data access processes, which allows researchers to understand and compare processes and requirements across provinces, territories, and pan-Canadian data holding organizations Statistics Canada and the Canadian Institute for Health Information (CIHI).<sup>4</sup>
- 2. A searchable inventory of 380 (and growing) health, social, survey, and registry datasets available at HDRN Canada Organizations across Canada.<sup>5</sup>
- 3. Detailed information about 170 algorithms that can be applied to identify health conditions (e.g., diabetes), health services use (e.g., physician house calls) and determinants of health (e.g., social and material deprivation). The algorithms were identified via a series of systematic reviews, which will be refreshed every six months. Each algorithm has been validated or assessed for feasibility for use in two or more Canadian jurisdictions.

Since DASH was launched on February 11, 2020, HDRN Canada has worked with over 35 requestors seeking to perform multi-province data-intensive research studies.

Another notable development is HDRN Canada's nascent Canada HDR Alliance. Modelled on the UK Health Data Research Alliance<sup>7</sup> and the UK-based International COVID-19 Data Alliance (which HDRN Canada belongs to),8 the Canada HDR Alliance will focus on partnerships and data linkage between HDRN Canada Organizations and groups holding large cohort, randomized clinical trial, and clinical data collections. HDRN Canada's first formal partnership is with the Canadian Partnership for Tomorrow's Health (CanPath). CanPath's consented research cohort data have been, or are expected to be, linked to administrative data holdings in all provinces, CanPath data are included in the DASH inventory which increases the utility of the DASH and the impact of CanPath data. Several other research teams and organizations have expressed interest in partnering with HDRN Canada, and HDRN Canada is working collaboratively with many data stakeholders including the following individuals and organizations that are submitting White Papers to NDRIO: Canadian Research Data Centre Network (CRDCN, Martin Taylor), Laura Rosella at the University of Toronto Dalla Lana School of Public Health, Diabetes Action Canada (Gary Lewis), and the Canadian Primary Care Sentinel Surveillance Network (CPCSSN, Sabrina Wong). From the perspective of HDRN Canada, the most important consideration when establishing DRI is not who is designated to be the primary collector or holder of data, but rather it is that sites adopt a Collect Once Use Many Times (COUMT) approach, 10 and develop the necessary policies and infrastructure to maximize the impact of data holdings and minimize unnecessary duplication, while ensuring strong governance so that data are protected.

It is important to note that even with the DASH and HDRN Canada tools, services and resources, there is still complexity in data access processes and variation in data holdings. These are not entirely avoidable due to underlying jurisdiction- or site-specific laws, policies, data sharing agreements, and data collection details. The role of HDRN Canada is to provide expert advice, tools, processes, methodologies and facilitation that directly support researchers in performing multijurisdictional research by reducing the impact of variation. As data centres work together to support researchers through DASH, they also learn from each other's practices, and can converge in aspects of the data access process in the longer term. This leads to the third major development of HDRN Canada – the creation of a learning and collaborative network. HDRN Canada's governance and management comprises 11 working groups, teams, and committees, 11 each with representatives from multiple provinces, territories, and pan-Canadian organizations, who are unified by the goal of unlocking the potential of Canada's unique provincial, territorial, and pan-Canadian population health data assets and expertise.

## CURRENT AND FUTURE STATE OF HDRN CANADA TOOLS, SERVICES & RESOURCES

From the time the idea for HDRN Canada was conceived in 2015, an investment in cutting edge science was envisioned. The 2015 budget plan included \$27.5M for *Enhanced Provincial Data Platforms* (i.e., integration and linkage of new sources of data) and *High Performance Computing* that was not in scope for, nor funded by, the *SPOR Canadian Data Platform* grant. Since 2015, additional new priorities and opportunities have emerged, including: increased capacity to perform distributed analyses on datasets that do not move across jurisdictional boundaries (vs. requiring that datasets be brought together in a single location); the need for widespread public engagement to build and maintain trustworthy practices and public trust in data-intensive research; the opportunity to create the space and resources to support Indigenous data sovereignty; and, as highlighted by Canada's experience with COVID-19, the requirement for coordinated, ethical, respectful, and sensitive approaches to data related to race, ethnicity, and other social determinants of health. For these and other topics,

there are examples of existing tools, services, and resources at individual HDRN Canada Organizations that could, and should, be taken to scale. The learning collaborative network of HDRN Canada can serve as a strong foundation to build tools, services, and resources to achieve the desired future state. Table 1 provides some selected examples of tools, services and resources that are poised to be spread in the near-term.

Table 1: Examples of Existing DRI at HDRN Canada Organizations that Could be Spread and Scaled

Tool/Service/Resource   Examples of Current Assets and Activities that could be   Desired Future			
and its Importance	Spread and Scaled	State State	
Social Data including Race and Ethnicity Data Data related to education, environmental, justice, social services, income, race, ethnicity, and other social determinants of health are important because health status depends more on these factors than it does on healthcare services utilization	<ul> <li>MB's data platform has grown from 12 health service delivery datasets in 1991 to more than 90 datasets including education, immigration, child protection, justice, social services, and early childhood development data</li> <li>NB has progressive legislation to share non-health provincial government data directly with the NB Institute for Research, Data and Training (NB-IRDT)</li> <li>BC's Data Innovation Program is bringing justice, education, early childhood, social development, income/work, and health data together, which are currently available for COVID research through PopData BC</li> <li>NB and other provinces are transferring health data to Statistic Canada where it is linked with comprehensive social data, anonymized, and made accessible via CRDCN Research Data Centres</li> <li>CIHI and Statistics Canada routinely link pan-Canadian administrative health services data with population surveys, Census, income data, mortality, and other data sources</li> <li>MB and ON have initiated work on frameworks for inclusion of race and ethnicity data in COVID-19 research studies</li> <li>CIHI has released proposed standards for race-based and Indigenous identity data</li> <li>Several provinces have linked data from Immigration, Refugees and Citizenship Canada (IRCC), and HDRN Canada is working with the IRCC to facilitate data sharing agreements in additional jurisdictions and enable studies related to refugees and other vulnerable immigrant groups</li> </ul>	Routine and regularly updated linkage of administrative data on health and social services across Canada with ongoing work to harmonize data across jurisdictions Common approaches to race, ethnicity, and other sensitive data are co-developed, refined and spread with deep involvement of members of affected communities  Anti-racist and positive inclusion frameworks for research are developed, refined, and spread	
Primary Care Electronic Medical Record (EMR) Data  EMR includes important variables that may not be in administrative data, e.g., smoking status, blood pressure, height, weight, coagulation measures and, in some cases, detailed primary health care provider notes	<ul> <li>Newfoundland Centre for Health Information (NLCHI) is working with the NL medical association on a framework for secondary use of EMR data; EMR data covering 60% of the province is expected to be available via DASH in 2021</li> <li>Through the support of the QC SPOR SUPPORT Unit, PARS3<sup>12</sup> has a mandate from the Ministère de la Santé et des Services sociaux to connect all primary care EMR vendors in QC</li> <li>ON has an agreement with TELUS to increase the timeliness of Primary Care EMR data held by ICES</li> <li>Several provinces are able to link data available through CPCSSN to other data sources</li> <li>CIHI maintains a data standard for a primary care EMR minimum dataset, created with clinician and researcher input, that could be more widely adopted</li> </ul>	Near population- wide and timely EMR data is available for research, practice improvement, and innovation	

Tool/Service/Resource	Examples of Current Assets and Activities that could be	Desired Future
and its Importance Algorithms Algorithms to measure key attributes of population health, health service use, and the determinants of health are developed, validated, and/or assessed for feasibility of implementation in multiple jurisdictions	<ul> <li>Spread and Scaled</li> <li>HDRN Canada's DASH includes 170 algorithms that have been validated or confirmed to be feasible to apply to administrative data from two or more Canadian jurisdictions</li> <li>HDRN Canada issued a call for proposals for new algorithm development to which 17 multidisciplinary research teams responded</li> <li>Four multi-jurisdiction projects have been selected to move forward</li> </ul>	Validated algorithms widely available and adapted for routine use in multi- jurisdictional analyses
High Performance Computing (HPC) Computational power for modern artificial intelligence and machine learning (AI/ML) methods and for the analyses of "Big Data" such as genomics	<ul> <li>NL Centre for Health Informatics and Analytics (CHIA) is HPC infrastructure at the Faculty of Medicine of Memorial University established in 2014 in partnership with IBM</li> <li>ON's Health AI Data Analysis Platform has supported 40 AI/ML studies (10 related to COVID-19)</li> <li>Statistics Canada is collaborating with the CRDCN to develop an HPC environment based on the Five Safes framework<sup>13</sup></li> </ul>	Accessible HPC for health available following the Five Safes framework and design to meet or exceed privacy and data protection requirements in all provinces and territories
Public Engagement As noted in the key premise of the International Consensus Statement, <sup>14</sup> the public is a key part of the solution to establish socially beneficial data-intensive research	<ul> <li>Following a broad call for applicants, 21 interviews were conducted, and 12 individuals were selected for HDRN Canada's Public Advisory Council has which has met twice</li> <li>HDRN Canada has published academic<sup>15</sup> and plain language articles <sup>16,17</sup> to raise awareness about uses of nonconsented data and the importance of plain language</li> <li>Multiple HDRN Canada Organizations have patient and public advisory councils and are sharing information through HDRN Canada's Public Engagement Working Group and other mechanisms</li> </ul>	A coordinated suite of public engagement activities that leverage and connect with public and patient engagement of and by governments and data institutions
Indigenous Data Sovereignty This refers to the right of Indigenous Peoples to govern the collection, application, and ownership of Indigenous data and information	<ul> <li>HDRN Canada is building understanding and capacity for Indigenous data sovereignty across HDRN Canada organizations to influence the appropriate collection, use, and governance of First Nations, Inuit, and Métis data; for example, ICES has an established Indigenous Portfolio that works closely with First Nations, Inuit, and Métis organizations in ON in partnerships that include data governance and data sharing agreements which has enabled Indigenous-driven research and data analysis using data held at ICES</li> <li>HDRN Canada is supporting Indigenous organizations to conduct Indigenous-led health data research across jurisdictions</li> </ul>	Data related to First Nations, Métis, and Inuit populations has established Indigenous governance and can be used by Indigenous organizations to inform decisions

For clarity, Table 1 presents only a subset of the activities that HDRN Canada and its member organizations are working on currently, with a focus on work that will have near term impacts. HDRN Canada is also engaged in other longer-term and transformational DRI-related work including: (i) leadership on secure research environments and requirements for data trusts, <sup>18</sup> (ii) a comprehensive review of data flows, metadata and client services at all member organization as the first step toward a technology plan that will increase efficiency and

enable future automation, (iii) a review of legal frameworks related to data sharing to distinguish between legislative requirements vs. regulatory practices vs. procedural choices of jurisdictions and lay the foundation for future regulatory reform, (iv) encouraging and supporting pragmatic clinical trials that make use of population health DRI to measure outcomes, (v) linkage to genomic data, and (vi) work related to "FAIRification" of data, i.e., metadata and computational systems to help researchers find, access, interoperate, and reuse data with no or minimal human intervention. Additionally, mechanisms to support "CAREful" consideration of Indigenous-identified data through the application of the CARE principles for Indigenous data governance are also in development. PCARE principles are people- and purpose-oriented and designed to complement the FAIR principles, while reflecting the crucial role of data in advancing Indigenous autonomy and sovereignty.

### ROLES FOR NDRIO IN POPULATION HEALTH DRI

Foremost, NDRIO is encouraged to articulate a commitment to including population health DRI in its work. Admittedly, population health DRI may be more complex than natural sciences DRI or biomedical DRI because of legislation, regulations, and public concerns about privacy and data security.<sup>20</sup> This is because the extremely beneficial impact of health DRI goes hand in hand with an increased need to address risks and stay within social licence.<sup>21</sup> However, as the COVID-19 pandemic has shown, there is no substitute for a jurisdiction having the ability to learn from its own health data. Population health DRI is the foundation for evidence-based decision making, and policy and practice changes, that literally save lives and promote wellbeing at a population level.

Secondly, NDRIO should acknowledge that population health DRI will require funding for personnel in ways that other kinds of DRI may not. Even with increased automation of processes, population health DRI requires people for work such as project- or dataset-specific privacy impact assessments, activities to involve members of the public and Indigenous nations within data governance, data sharing negotiations and contracts oversight, work with regulators and ministries, training and capacity development, cybersecurity, etc.

For all the potential future state advances noted above, investment is essential, but NDRIO's role can and should go beyond that of funder. NDRIO also has a role to play as a convenor and/or community of practice leader that:

- Identifies and spreads, and in some cases requires, best practices, e.g., related to secure HPC environments, credentialing users, COUMT approaches, cybersecurity, the Five Safes framework, FAIR data principles, use of cloud-based computing resources, technology/software for distributed analytics and/or data federation, and other leading-edge DRI topics
- Advances the priorities of Indigenous nations through the adaptation of the CARE Principles<sup>19</sup> and the
  establishment of Indigenous-led DRIs that could lead to greater community capacity development, advance
  Indigenous innovation, and uphold Indigenous data sovereignty and governance mechanisms
- Encourages researchers to seek out and use data from other disciplines, establishing mechanisms (e.g., data sharing agreements) for data-holding organizations to collaborate, and supports/ pathways for researchers to integrate their data holdings with other DRI (e.g., as a *Canada HDR Alliance* member, or through deposit with an HDRN Canada Organization)
- Hosts national discussions with governments and regulators related to privacy legislation (e.g., discussing the implications of GDPR-like models<sup>22</sup> and recent amendments introduced to federal privacy legislation)
- Brings together private sector organizations and foundations, involving them in the establishment of new and expanded DRI that help Canada become a global leader in collaborative work where private sector science makes use of public sector data in ways that are cutting edge, fully transparent, ethical and aligned with the values of the Canadian public
- Convenes leading DRI organizations, from Canada and around the world, to learn from each other and develop and support opportunities for international research collaborations

In closing, multiple reports and committees have identified the benefits that could be realized with more investment in, and coordination and leadership of, Canada's uniquely comprehensive population health DRI. <sup>23,24, 25,26,27</sup> There are numerous examples of DRI tools, services, and resources that are already operational at a subset of HDRN Canada Organizations, or being developed by HDRN Canada, and that could be spread and taken to national scale if funding is provided. But this White Paper is more than a request for an allocation of funding. The true power of Canada's DRI lies in bringing diverse datasets, expertise, and research communities together. HDRN Canada would welcome NDRIO's leadership as a convenor who helps build a community of practice for DRI that spans multiple sectors.

#### REFERENCES

1-----

- <sup>4</sup> HDRN Canada. Data Access Processes [Internet]. Vancouver BC: HDRN Canada; 2020 [cited 2020 Dec 7]. Available from: <a href="https://www.hdrn.ca/en/dash/dar">https://www.hdrn.ca/en/dash/dar</a>
- <sup>5</sup> HDRN Canada. Inventory [Internet]. Vancouver BC: HDRN Canada; 2020 [cited 2020 Dec 7]. Available from: <a href="https://www.hdrn.ca/en/inventory/">https://www.hdrn.ca/en/inventory/</a>
- <sup>6</sup> HDRN Canada. Algorithm [Internet]. Vancouver BC: HDRN Canada; 2020 [cited 2020 Dec 7]. Available from: <a href="https://www.hdrn.ca/en/algorithm">https://www.hdrn.ca/en/algorithm</a>
- <sup>7</sup> UK Health Data Research Alliance. UK Health Data Research Alliance [Internet]. London: UK Health Data Research Alliance [cited 2020 Dec 7]. Available from: <a href="https://ukhealthdata.org/">https://ukhealthdata.org/</a>
- <sup>8</sup> UK Health Data Research Alliance. International COVID-19 Data Alliance (ICODA) [Internet]. London: UK Health Data Research Alliance [cited 2020 Dec 7]. Available from: <a href="https://www.hdruk.ac.uk/covid-19/international-covid-19-data-alliance/">https://www.hdruk.ac.uk/covid-19/international-covid-19-data-alliance/</a>
- <sup>9</sup> HDRN Canada. HDRN Canada Announces New Partnership with CanPath to Enable Data Linkage [Internet]. Vancouver BC: HDRN Canada; 2020 [cited 2020 Dec 7]. Available from: <a href="https://www.hdrn.ca/en/news-updates/hdrn-canada-announces-new-partnership-canpath-enable-data-linkage">https://www.hdrn.ca/en/news-updates/hdrn-canada-announces-new-partnership-canpath-enable-data-linkage</a>
- <sup>10</sup> Joukes E, Cornet R, de Keizer N, de Bruijne MC. Collect Once, Use Many Times: End-Users Don't Practice What They Preach. InMIE 2016 Jan 1 (pp. 252-256). https://doi.org/10.3233/978-1-61499-678-1-252
- <sup>11</sup> HDRN Canada governance and management includes: Board of Directors, Executive Committee, Leads Team, Public Advisory Council, SPOR Canadian Data Platform Advisory Committee, Modelling and Informatics Group, DASH Working Group, Algorithms and Harmonized Data Working Group, Indigenous Research Team, Public Engagement Working Group, and Privacy Team
- <sup>12</sup> Research Group in Health Informatics (GRIIS). PARS3: Learning platform in health research and social services [Internet]. Sherbrooke Québec: GRIIS [cited 2020 Dec 7]. Available from: https://griis.ca/en/solutions/pars3/
- <sup>13</sup> Ritchie F. The 'Five Safes': a framework for planning, designing and evaluating data access solutions. Zenodo [Internet]. 2017 Sept [cited 2020 Dec 7]. Available from: https://doi.org/10.5281/zenodo.897820
- <sup>14</sup>Aiken M, Tully MP, Porteous C, Denegri S, Cunningham-Burley S, Banner N et al. Consensus statement on public involvement and engagement with data intensive health research. *Int J Popul Data Sci.* 2019;4:1:06. https://doi.org/10.23889/ijpds.v4i1.586
- <sup>15</sup> Paprica PA, McGrail KM, Schull MJ. Notches on the dial: a call to action to develop plain language communication with the public about users and uses of health data *Int J Popul Data Sci.*. 2019 Aug;4(1). <a href="https://doi.org/10.23889/ijpds.v4i1.1106">https://doi.org/10.23889/ijpds.v4i1.1106</a>
- <sup>16</sup> Paprica PA, McGrail KM, Schull MJ. The public needs to know why health data are used without consent. The Conversation [Internet]. 2019 October 6 [cited 2020 Dec 7]. Available from: <a href="https://theconversation.com/the-public-needs-to-know-why-health-data-are-used-without-consent-123669">https://theconversation.com/the-public-needs-to-know-why-health-data-are-used-without-consent-123669</a>

<sup>&</sup>lt;sup>1</sup> HDRN Canada. Member Organizations [Internet]. Vancouver BC: HDRN Canada; 2020 [cited 2020 Dec 7]. Available from: <a href="https://www.hdrn.ca/en/about/member-organizations">https://www.hdrn.ca/en/about/member-organizations</a>

<sup>&</sup>lt;sup>2</sup> Dahl LT, Walker JD, Schull M, Paprica PA, McDonald JT, Gavin F, et al. The SPOR-Canadian Data Platform: a national initiative to facilitate data rich multi-jurisdictional research. *Int J Popul Data Sci.* 2020 Nov 9;5(1). https://doi.org/10.23889/ijpds.v5i1.1374

<sup>&</sup>lt;sup>3</sup> HDRN Canada. Data Access Support Hub (DASH) [Internet]. Vancouver BC: HDRN Canada; 2020 [cited 2020 Dec 7]. Available from: <a href="https://www.hdrn.ca/en/dash">https://www.hdrn.ca/en/dash</a>

- <sup>17</sup>. Paprica PA, McGrail KM, Schull MJ. Plain language about health data is essential for transparency and trust. The Conversation [Internet]. 2019 October 9 [cited 2020 Dec 7]. Available from: <a href="https://theconversation.com/plain-language-about-health-data-is-essential-for-transparency-and-trust-123319">https://theconversation.com/plain-language-about-health-data-is-essential-for-transparency-and-trust-123319</a>
- <sup>18</sup> Paprica, P. A., Sutherland, E., Smith, A., Brudno, M., Cartagena, R. G., Crichlow, M., Courtney, B., Loken, C., McGrail, K. M., Ryan, A., Schull, M. J., Thorogood, A., Virtanen, C. and Yang, K. (2020) "Essential Requirements for Establishing and Operating Data Trusts: Practical Guidance Based on A Working Meeting of Fifteen Canadian Organizations and Initiatives", *Int J Popul Data Sci.*, 5(1). https://doi.org/10.23889/ijpds.v5i1.1353
- <sup>19</sup> Carroll SR, Garba I, Figueroa-Rodríguez OL, Holbrook J, Lovett R, Materechera S, Parsons M, Raseroka K, Rodriguez-Lonebear D, Rowe R, Sara R. The CARE Principles for Indigenous Data Governance. *Data Science Journal*. 2020 Nov 4;19(1). http://doi.org/10.5334/dsj-2020-043
- <sup>20</sup> Teng J, Bentley C, Burgess MM, O'Doherty KC, McGrail KM. Sharing linked data sets for research: results from a deliberative public engagement event in British Columbia, Canada. *Int J Popul Data Sci.* 2019 May 7;4(1). <a href="https://doi.org/10.23889/ijpds.v4i1.1103">https://doi.org/10.23889/ijpds.v4i1.1103</a>
- <sup>21</sup> Paprica PA, de Melo MN, Schull MJ. Social licence and the general public's attitudes toward research based on linked administrative health data: a qualitative study. CMAJ open. 2019 Jan;7(1):E40. <a href="https://doi.org/10.9778/cmajo.20180099">https://doi.org/10.9778/cmajo.20180099</a>
- <sup>22</sup> GDPR is Europe's General Data Protection Act. . European Commission. EU data protection rules. Brussels BE: European Commission; 2019 [cited 2020 Dec 7]. Available from: <a href="https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules">https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules</a> en
- <sup>23</sup> Council of Canadian Academies. Accessing Health and Health-Related Data in Canada. Ottawa ON: The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation, Council of Canadian Academies; 2015 [cited 2020 Dec 7]. Available from: <a href="https://cca-reports.ca/wp-content/uploads/2018/10/healthdatafullreporten.pdf">https://cca-reports.ca/wp-content/uploads/2018/10/healthdatafullreporten.pdf</a>
- <sup>24</sup> Health Canada. Unleashing Innovation: Excellent Healthcare for Canada Executive Summary [Internet]. Ottawa ON: Government of Canada; 2014 [cited 2020 Dec 7]. Available from: <a href="https://www.canada.ca/en/healthcanada/services/publications/health-system-services/unleashing-innovation-excellent-healthcare-canada-executive-summary.html">https://www.canada.ca/en/healthcare-canada-executive-summary.html</a>
- <sup>25</sup> Task Force on Artificial Intelligence for Health (AI4Health). Building a Learning Health System for Canadians [Internet]. Toronto ON: Canadian Institute for Advanced Research; 2020 [cited 2020 Dec 7]. Available from <a href="https://cifar.ca/wp-content/uploads/2020/11/AI4Health-report-ENG-10-F.pdf">https://cifar.ca/wp-content/uploads/2020/11/AI4Health-report-ENG-10-F.pdf</a>
- <sup>26</sup> Government of Canada. Health and biosciences economic strategy table [Internet]. Ottawa ON: Health and biosciences economic strategy table; 2018 [cited 2020 December 7]. Available from: <a href="https://www.ic.gc.ca/eic/site/098.nsf/vwapj/ISEDC\_Table\_HB.pdf">https://www.ic.gc.ca/eic/site/098.nsf/vwapj/ISEDC\_Table\_HB.pdf</a>/
- <sup>27</sup> Forest PG and Martin D. Fit for Purpose: Findings and Recommendations of the External Review of the Pan-Canadian Health Organizations Summary Report [Internet]. Ottawa ON: Health Canada; 2018 [cited 2020 Dec 7]. Available from: <a href="https://www.canada.ca/en/health-canada/services/health-care-system/reports-publications/health-care-system/findings-recommendations-external-review-pan-canadian-health-organization.html">https://www.canada.ca/en/health-canada/services/health-care-system/reports-publications/health-care-system/findings-recommendations-external-review-pan-canadian-health-organization.html</a>