Cybersecurity, Privacy, and Artificial Intelligence in Health Data: Advancements and Challenges

May 5th 2023, Ottawa City Centre Delta Hotel. Organized by the University of Waterloo Cybersecurity and Privacy Institute (CPI), and the Waterloo Artificial Intelligence (Waterloo.AI) Institute. With funding and organizational assistance from Health Canada, Statistics Canada, Canadian Institute for Health Information (CIHI), WatSPEED (at the University of Waterloo), and the Master of Public Service Program (University of Waterloo)

Foreword

Effective knowledge dissemination between universities and governments is critically important for Canada, given the need for consensus on how to best collect, protect, and leverage data to ensure high-quality and impactful public healthcare. Such knowledge dissemination helps develop and implement strategies that facilitate the use of advanced AI methods aimed at better health care delivery, while protecting individual privacy and maintaining public trust. These principles are not mutually exclusive.

In response, the University of Waterloo Cybersecurity and Privacy Institute (CPI) and Waterloo Artificial Intelligence (Waterloo.AI) Institute, co-funded and organized a conference that was held at the Delta Hotel in Ottawa on May 5th, 2023, which was attended by representatives from multiple federal government organizations. Its objective was to bring contemporary developments in cybersecurity, privacy, and Artificial Intelligence (AI) research to the attention of national policymakers. The conference also received funding and support from Health Canada, Statistics Canada, the Canadian Institute for Health Information (CIHI), the Master of Public Service (MPS) program at the University of Waterloo, and WatSPEED (University of Waterloo).

The conference builds on some of the recommendations of the Expert Advisory Group of the Pan Canadian Health Data Strategy (PCHDS), established by the Public Health Agency of Canada. The task of the PCHDS was to provide guidance on modernizing national health data collection, sharing and interoperability, and offer guidance on updating standards on privacy and access to health data, as well as overall data governance. This e-book is a collection of research papers based on the presentations that were made by faculty members across different academic units of the University of Waterloo.

With respect to cybersecurity and privacy: (1) Professor Gautam Kamath (Cheriton School of Computer Science) discussed some of the latest research in the field of differential privacy; (2) Professor Guang Gong (Department of Electrical and Computer Engineering) introduced Blockchain concepts along with a survey of healthcare applications; and (3) Professor Helen Chen (School of Public Health Sciences) proposed the use of synthetic data as a means to conduct statistical research while significantly protecting individual privacy. In the session on Artificial Intelligence applications: (1) Professor Sirisha Rambhatla (Department of Management Sciences) outlined various ways in which AI-powered solutions can revolutionize primary healthcare based on her research in liver transplantation and burn surgical candidacy forecasting, and also discussed the potential role of AI in surgical training and the need to address misinformation from large-language models; (2) Professor Plinio Morita (School of Public Health Sciences) delved into the need to better incorporate Big Data through novel data sources and recommended how to incorporate them within public health contexts; and (3) Professor John Hirdes (School of Public Health Sciences) discussed the catastrophic impacts of COVID-19 on Long Term Care (LTC) facilities and how inconsistent data standards, limitations of data access because of privacy concerns, and inertia on knowledge mobilization, all impede the use of AI methods in healthcare.

The conference concluded with presentations from Professor Anindya Sen (Department of Economics) and Professor Samantha Meyer (School of Public Health Sciences). Professor Sen discussed the consequences of limited public understanding of data and statistical methods, which ultimately impedes the use of AI methods that can lead to better healthcare delivery and outcomes. He illustrated this through the recent controversy on alleged mass surveillance of Canadians by the Public Health Agency of Canada (PHAC), which he posited was both unfounded and a result of poor data literacy. Professor Meyer noted how AI use is heavily dependent on the willingness to share data, which is a function of public trust in government. Her paper introduced the concept of institutional trust and focused on the relationship between societal trust health behaviours, such as data sharing. Our hope is that these publicly available studies will stimulate more impactful work and knowledge sharing that is needed for our public health care systems. We gratefully acknowledge the encouragement and help given by Deputy Minister Stephen Lucas of Health Canada and Chief Statistician of Canada Anil Arora, along with other members of their teams, which made this conference possible.

Anindya Sen
Professor of Economics and Associate Director, Waterloo CPI
Member, Waterloo.AI

Preface from Dr. Vivek Goel

During the COVID-19 pandemic Canada, along with many jurisdictions globally, faced many challenges in ensuring that accurate data were available in a timely manner to support the management of the disease at the clinical level and in health care institutions, as well as at the population level to support policy development, research, and innovation. Challenges in Canada were further magnified because of the constitutional distribution of responsibility for management and delivery of most health services to the provinces and territories.

I had the privilege of serving as Chair of the Expert Advisory Group to advise on the development of a Pan-Canadian Health Data Strategy. Through a series of three reports, we identified the barriers standing in the way of better collection and sharing of health data and laid out a series of recommendations for action. It has been pleasing to see the attention paid to these recommendations and the momentum building towards action, although a great deal of work remains to be done.

This e-book summarizes papers presented by colleagues from the University of Waterloo at a conference on Cybersecurity, Privacy, and Artificial Intelligence in Health Data. The papers highlight advances in a variety of areas, from technical tools that could be used to address privacy challenges in sharing of data to novel analytic approaches for large datasets. However, as we noted in our reports, many of the challenges are not technical, but social. Colleagues therefore also addressed topics such as the importance of data literacy and trust.

Every part of our health systems faces significant challenges – all require better data which is essential to building a high quality, accessible, efficient, and sustainable health system. We do hope that the ideas raised in this conference can be useful as policy work continues on improving health data across Canada.

As part of its 'Waterloo at 100' vision, the University of Waterloo has committed to being a national leader in a series of Global Futures; including health futures, particularly the intersection of society, health, and technology, and increasing partnerships with the public sector to catalyze important policy development. This conference was an important step in this direction, and I look forward to future partnerships that will allow us to bring more of the seminal research being conducted at the university to the attention of policymakers at all levels of government.

I am grateful to the conference presenters, organizers, and supporters for making this happen.

Vivek Goel, CM

President and Vice-Chancellor

University of Waterloo

Foreword to Conference Proceedings: Cybersecurity, Privacy, and Artificial Intelligence in Health Data

The "Cybersecurity, Privacy, and Artificial Intelligence in Health Data: Advancements and Challenges" conference, held on May 5th, 2023, brought together an array of distinguished experts, scholars, and professionals, all united by a shared commitment to addressing the evolving landscape of healthcare in the digital age.

The importance of the topics explored throughout this conference in addressing the opportunities and challenges that come with the evolving digital health care landscape cannot be understated. One of the Government of Canada's major health priorities is to harness the full potential of health data and digital systems. When data is effectively used and shared, it can save lives by ensuring health care providers have the information they need to make sound clinical decisions across health settings and across jurisdictions. It can also empower Canadians in their health journeys with access to their own comprehensive health record. De-personalized health data is critical in supporting improvements in health systems, in population health, in public health responses and in innovation.

The federal government has made a clear commitment to enabling a modern and connected health system that fully harnesses the potential of data and digital systems through the *Working Together to Improve Health Care for Canadians Plan*. In support of this plan, Budget 2023 provided close to $200 billion over 10 years, including $46.2 billion in new funding for provinces and territories on four shared health priorities, one of which is modernizing the health system with standardized health data and digital tools.

As part of the *Working Together Plan*, the federal government is leading the work with provinces and territories, with industry and other key stakeholders to realize this vision for a more connected health system by focusing on the adoption of common standards and modernized approaches for safely and securely sharing health data. These efforts must be grounded in principles that put people and their care at the centre of the approaches taken, while balancing the need to share data with privacy protection.

Utilizing innovations, like artificial intelligence, also presents great opportunities for improvements to health care delivery. Whether it's exploring how automation can help improve workflows, using AI to support diagnostics, or employing machine learning to support health workers in their jobs, the goal is to enhance the quality and efficiency of health care in Canada while ensuring appropriate regulation and safeguards to minimize safety risks and build and maintain public trust in how these tools are used.

The conference helped to inform how to advance the changes needed for Canada to fully benefit from a modernized and better connected health system.

Thank you for your participation and engagement in these exciting conversations.

Dr. Stephen Lucas

Health Canada

University of Waterloo AI/Cybersecurity Conference May 5, 2023

Chief Statistician - Foreword to eBook of conference proceedings (Sept 20, 2023)

I’m honored to provide a forward to the summary of such an important event. New data sources, big data, and artificial intelligence present tremendous opportunities to improve population health as does data integration of different sources such as administrative data, biobank data, and surveys. Using other forms of data (such as data from mobile apps) can enhance our understanding of health and health outcomes.

This conference brought together researchers from academia, policy makers, and data experts to discuss responsible ways of leveraging available data and modern technology to gain a deeper understanding of the key factors that lead to better health outcomes. Statistics Canada has a long history and solid track record of using various data sources and statistical frameworks and enabling technical infrastructure to bring data and insights to Canadians. This expertise and capability needs to be better leveraged to drive better health outcomes for Canadians.

Statistics Canada is uniquely placed to play a data stewardship role. We started from the foundational recognition that everything is about trust and trust is everything. We need data to improve lives through insights and to get that data, Canadians must be able to trust that their privacy is protected, and their data are safe. Some of the tools we use to maintain the trust of Canadians include our governance models, our frameworks, and our standards. Our Quality Assurance Framework and our Quality Guidelines, for example, help ensure that our outputs are of high quality and can be trusted by our data users. We also use an ethical framework built around the principles of Necessity and Proportionality to guide our data collection activities.

This framework helps us strike a balanced approach. The conference also highlighted the need for collaboration to address critical health issues and the important role of data integration. Bringing multiple sources of data together unlocks new capabilities and insights. New data sources, big data, and artificial intelligence present tremendous opportunities to improve health outcomes.

Together, we have a real opportunity to grow the health data ecosystem. Ultimately, it seems clear that data are a team sport – and one that is reliant on public trust to succeed.

Anil Arora

Chief Statistician of Canada

Contents

A Whirlwind Introduction to Differential Privacy

Gautam Kamath*

University of Waterloo & Vector Institute

July 17, 2023

Abstract

When an organization naively releases too many statistics about a sensitive dataset, privacy leakage is inevitable. Differential privacy has emerged as a powerful framework for protecting individuals' privacy while enabling valuable insights from data analysis. We provide an introduction to differential privacy, aiming to present the fundamental concepts, techniques, and applications of this powerful framework. We begin by defining the core principles of differential privacy and discussing its mathematical foundations; most importantly, focusing on the definition and what it means for an algorithm to be differentially private. Subsequently, we delve into various mechanisms and algorithms employed to achieve differential privacy, including the Laplace and Gaussian Mechanisms. We discuss several useful properties of differential privacy, as well as a popular relaxation.

*g@csail.mit.edu. Supported by a Canada CIFAR AI Chair, and NSERC Discovery Grant, and a University of Waterloo startup grant.

1 Failures to Preserve Privacy

Modern data analysis frequently operates on datasets which are sensitive in nature. This could include things like medical records, personal emails, or location data. Troublingly, if the data analyst is naive or careless, private information could easily be leaked. We illustrate a couple of real-world examples where issues like this have arisen, focusing on the medical domain.

1.1 Genomic Studies

In the context of medical and genomic studies, statistical data analysis plays a pivotal role in scientific advancements. However, the importance of preserving data privacy is paramount. A critical concern arises when considering studies that involve sensitive information, such as individuals diagnosed with HIV. Revealing the identities of participants in such studies would be a severe breach of privacy. Homer et al. [HSR+08] demonstrated that, under certain conditions, it is possible to discern an individual's presence in a DNA sample mixture based on aggregate statistics (though we note that the practical applicability of these conditions has been debated [BRS+09]). Nevertheless, this discovery prompted the National Institutes of Health (NIH) in the US to restrict access to previously open-source summary statistics (including minor allele frequencies, chi-squared statistics, and p-values), thereby hindering open science. Thus, ensuring privacy-preserving analyses, particularly for tasks like statistical hypothesis testing, would greatly facilitate progress in genomic studies while upholding privacy concerns.

1.2 Massachusetts Group Insurance Commission

During the mid-1990s, the Massachusetts Group Insurance Commission initiated a program granting researchers free and open access to hospital visit records of state employees. Given the highly sensitive nature of this data, it naturally underwent anonymization processes. The then-governor of Massachusetts, William Weld, assured the public that patient privacy would be safeguarded. As you might expect, this turned out to be untrue.

The anonymization involved removing identifiers like names and social security numbers from the dataset, while keeping other features that were not directly identifying, such as ZIP code, date of birth, and sex. Latanya Sweeney, a graduate student in Computer Science, acquired the voter rolls from the city of Cambridge for a mere $20. These records contained information such as names, addresses, ZIP codes, dates of birth, and sexes of registered voters. As many of these fields were not redacted from the anonymized dataset, it was easy to match them up. Indeed, by cross-referencing these datasets, large-scale reidentification of individuals in the hospital visit dataset became straightforward. Sweeney demonstrated this by sending Governor Weld his own medical records, highlighting the vulnerability of the anonymization process and the subsequent privacy risks involved.

2 Enter Differential Privacy

As we have seen, without special care, statistical analyses are inclined to violate the individual privacy of those who provided their data. Even worse, as these examples demonstrate, best-effort or heuristic methods to “anonymize” a dataset may fail catastrophically. This is unacceptable when privacy is a serious concern. As such, we require rigorous methods to guarantee that privacy is preserved.

It may be a priori unclear how to provide such a rigorous guarantee. We will first intuitively demonstrate via a simple example. Consider a population where every individual is either HIV positive or negative. We wish to estimate what fraction of the population is indeed HIV positive. One option is to simply ask individuals whether they have HIV or not. If the data analyst is not trusted, the issue is immediately apparent: each individual has to reveal their personal information about whether or not they have HIV, which they may be uncomfortable to do. Indeed, individuals are likely to lie, resulting in an undercounting of the infected fraction. Consider instead another extreme: individuals flip a coin, and if it comes up heads, then they report “HIV positive,” otherwise they report “HIV negative.” This perfectly protects individual privacy: each person’s report is independent of their actual HIV status. However, by the same token, the resulting aggregate of their results is entirely unrelated to the fraction of infected individuals in the dataset. We have perfect privacy, but no utility. The right middle ground between these two deficient solutions balances the two desiderata of utility and privacy. Roughly speaking, each individual independently does the following: with a probability of 2/3 they report the truth, and otherwise, they lie. This gives each individual a type of “plausible deniability”: even if they report that they are HIV positive, there is a 1/3 chance that they are not, and the protocol simply instructed them to say that they are. At the same time, by appropriately transforming the aggregate of their reports, it is possible to get an accurate estimate of the true fraction of individuals who are HIV positive. This type of technique is naturally called “Randomized Response.” It is the simplest example of a protocol that satisfies differential privacy. It was actually introduced in 1965 by Warner [War65], decades before differential privacy was invented. Furthermore, it satisfies a stronger notion of privacy, known as local differential privacy.

2.1 The Definition of Differential Privacy

To be more formal, let us introduce the actual definition of differential privacy. We will work in the trusted curator model, sometimes known as central differential privacy. Suppose there exist n individuals each with their own datapoint, X_1 through X_n. Their data will all be sent in the clear to said, “trusted curator” (i.e., the data analyst) – this individual is trusted with their sensitive data, but no one else. In particular, this curator is trusted to not divulge any information that violates the privacy of the individuals who provided their data. The curator then runs a “mechanism” (a fancy bit of domain-specific jargon for an algorithm, a statistic, or generally, any procedure that takes in a dataset and produces an output) M, and publicly releases the result of this computation. Differential privacy is a property of the mechanism M, which informally says that its distribution over outputs is insensitive to modifications of individual datapoints.

Let’s be a bit more precise. Let $M : X^n \rightarrow Y$ be an algorithm that takes in a dataset of size n over a domain X, and outputs something from the range Y. We will consider any pair of datasets $X, X' \in X^n$ that are *neighbouring*, that is, differ in exactly one datapoint. One could, for example, consider a single individual lying about their datapoint, or otherwise opting out of the data analysis. Such neighbouring datasets are sometimes denoted by $X \sim X'$. We say that M is ε-(*pure*) *differentially private* (ε-(pure) DP) if, for all neighbouring X, X', and all $T \subseteq Y$, we have

$$\Pr[M(X) \in T] \leq e^{\varepsilon} \Pr[M(X') \in T],$$

where the randomness is over the choices made by M.

Differential privacy was introduced in 2006 by Dwork, McSherry, Nissim, and Smith, in their seminal paper [DMNS06]. It is considered to be the gold standard of data privacy, receiving both acceptance and acclaim in theory as well as practice. In terms of practical, real-world deployments, it has been adopted at Apple [Dif17], Google [EPK14], Microsoft [DKY17], the US Census Bureau for the 2020 US Census [AACM+22], and much more. For a more comprehensive look at practical deployments of differential privacy, see [Des21].

2.2 Interpreting Differential Privacy

The definition of differential privacy is quite counterintuitive the first few times you see it. Here are some important points to help you better digest it.

To simply interpret the definition of differential privacy: consider the probability distribution that the mechanism M has over potential outcomes, when given some dataset X. Differential privacy bounds how much the probability of any outcome can increase (multiplicatively) after changing one user's data. In other words, if we modify one datapoint, no outcome becomes dramatically more or less likely.

Differential privacy is a quantitative definition: a small value of ε means the mechanism is more private, and a larger value means the mechanism is less private. Specifically, ε = 0 means that the mechanism does not depend on the data at all, and thus is *perfectly* private, while ε = ∞ means that the algorithm does not even attempt to preserve privacy. Typical values of ε one might see range from (roughly) 0.1 to 8. Anything outside these ranges should be subject to additional scrutiny. This is because most algorithms which operate on reasonable-sized datasets would be unable to provide reasonable utility at values of ε smaller than 0.1. Indeed, implicit in the design of any differentially private mechanism is the underlying *privacy-utility* tradeoff; privacy never comes for free, and the stronger the privacy guaranteed, the more that utility will suffer. On the other hand, values of ε larger than 8 result in rather weak privacy guarantees, and using larger values of ε is not generally recommended; it may create a false sense of security against privacy violations. As a disclaimer, these suggestions are broad, and consequently, one should take more care when determining appropriate values of ε for a particular domain. Indeed, the most principled approach would be to compute the social and economic benefit of a particular level of privacy-protection, and balance this against the cost of the resulting loss in data utility. This type of quantification is beyond the scope of this paper but has been approached for certain high-stakes applications like the 2020 US Census [AACM+22].

Differential privacy is a worst-case guarantee: the definition requires one to reason about the worst-case outcome over all pairs of neighbouring datasets X and X'. A common pitfall is to try to quantify which types of datasets are more or less likely and give a type of "average-case" privacy guarantee which takes this into account. This is dangerous for two reasons: first, if the data analyst's assumptions about the population are incorrect, then they risk violating user privacy – in settings where privacy actually matters, this is a non-starter. Second, several very helpful technical properties of differential privacy fail to hold, which are described more in-depth by Steinke and Ullman [SU20a, SU20b].

The definition may look like the two neighbouring datasets X and X' play an asymmetric role.
However, one can just swap their positions, and the definition is seen to be symmetric.

Finally, any "reasonable" differentially private mechanism must be randomized. That is, if it does not involve some randomness, then the algorithm must be completely independent of the dataset.

2.3 What Differential Privacy Does and Does Not Do

It is easy to misinterpret what differential privacy actually guarantees. Repeating the definition: differential privacy guarantees that the probability of any outcome is similar when an individual is or is not included in the dataset. In fact, it guarantees exactly this, and nothing more and nothing less. Many misconceptions, even by experienced researchers, are based around the forgetting of this fact.

Differential privacy protects against the attacks seen at the beginning of this paper. That is because these are attacks which re-identify individuals, reconstruct the database, or identify whether individuals were present in the dataset. All of these are highly sensitive to the addition or removal of a single datapoint, and thus, if a mechanism is differentially private, the output would not reveal any of these things. Conveniently, differential privacy is, in some sense, "future-proof"; even if some data became public in the future, it cannot induce additional privacy leakage from whatever the data analyst made public beforehand. This is not true for some of the other heuristic methods we have seen prior. Differential privacy actually bounds the increase in probability of arbitrary risks, one of the key reasons why it is considered to be such a strong and safe notion of data privacy.

On the other hand, differential privacy does not stop statistics and machine learning. A common example to demonstrate this is by considering a study which is meant to investigate whether smoking causes cancer. An individual who smokes cigarettes is deciding whether or not to participate in a medical study investigating this question. A positive causal link would hurt them; since their status as a smoker is known to the public, this would result in an increase in their insurance premiums. However, since they are told that the study is being conducted with differential privacy, they decide nonetheless to participate, since they know that the experiment will be privacy preserving. Regardless of this, the medical study reveals that indeed, smoking causes cancer. Is this a privacy violation or not? The answer is no, or at least not in the way that differential privacy guarantees against. That is, given a large sample size, it would have revealed that smoking causes cancer whether or not this particular individual participated. To borrow terminology from Frank McSherry [McS16], this was not the smoker's secret, but instead it was a secret about the smoker. The only way to avoid such a disclosure would be either to not conduct the study in the first place, or for the smoker to conceal the fact that they smoke.

Differential privacy is also not an appropriate solution to cases when we are explicitly trying to identify particular individuals. Say, for example, when we are trying to do contact tracing for COVID-19. This is when we trace back the sequence of individuals who were in contact with someone who tested positive, in order to either locate the source or warn others who were exposed. There is much discussion of "privacy-preserving" technical solutions for this problem. However, this is due to the term privacy-preserving being overloaded with several different meanings, and differential privacy is not a relevant notion here. Indeed, contact tracing requires information about the identity of individuals, whereas differential privacy is specifically intended to mask the presence of individuals. If several people all attended an event at the same location, however, there may be some hope through tracking of this large group of people. Specific work at Google using private analysis of location traces since COVID-19 began may be applicable for this type of application [ABC+20].

3 The Laplace Mechanism

To make the preceding discussion more concrete, we introduce perhaps the best-known mechanism in differential privacy, the Laplace mechanism. Simply compute the quantity you wish to release non-privately, but then add noise to it before releasing it. Intuitively speaking, adding noise masks the contribution of any individual, thus giving plausible deniability akin to what we discussed in the Randomized Response example. However, formalizing and analyzing this mechanism requires some care.

As a simple running example through this section, one can consider the simple function which averages *n* bits, $f = \frac{1}{n}\sum_{i=1}^{n} X_i$ where $X_i \in \{0,1\}$. Say, to return to a previous example, X_i represents whether the *i*th individual is HIV positive. We wish to release its value on some dataset $X \in \{0,1\}^n$ subject to differential privacy. As the name of the mechanism suggests, the *Laplace distribution* will be a key component of the Laplace mechanism.

Definition 3.1. *The* Laplace distribution *with location and scale parameters* 0 *and* b*, respectively, has the following density:*

$$\rho(x) = \frac{1}{2b}\exp\left(-\frac{|x|}{b}\right)$$

This may be an unfamiliar distribution, but it can be seen as a symmetrization of the exponential distribution, by just reflecting the distribution's PDF across the y-axis. The exponential distribution is supported on $x \in [0, \infty)$, with density $\propto \exp(-cx)$, while the Laplace distribution has support $x \in \mathbb{R}$ with density $\propto \exp(-c|x|)$. We can also consider the Gaussian distribution: this is also supported on $\mathbb{R}$, similar to the Laplace distribution, but it has lighter tails (with density $\propto \exp(-cx^2)$).

With these preliminaries out of the way, we can proceed with describing the mechanism. We first require the concept of sensitivity. This measures how much a function can change if one datapoint is modified. Intuitively, the more sensitive a function is, the harder it is to privatize (i.e., more noise is required), and the less sensitive it is, the easier it is to privatize.

Definition 3.2. *Let $f : X^n \to \mathbb{R}^k$. The ℓ_1-sensitivity of f is*

$$\Delta_1 = \max_{X,X'} \| f(X) - f(X') \|_1,$$

where X and X' are neighbouring databases.

Recalling our function *f* which averages *n* bits, it has sensitivity $1/n$: this is because the most we can change its value is by changing a 0 to 1 or vice versa, which changes the value of the function by $1/n$.

Finally, we are prepared to define the Laplace mechanism. We simply add noise to the function of interest, which is scaled according to the sensitivity of the function.

Definition 3.3. *Let $f : X^n \to \mathbb{R}^k$. The* Laplace mechanism *is defined as*

$$M(X) = f(X) + (Y_1, \ldots, Y_k),$$

where the Y_i are independent Laplace(Δ/ε) random variables.

Note that this is phrased for the more general setting where the function *f* has a *k*-dimensional range, but for our running example, we can instantiate it with $k = 1$. Specifically, it prescribes outputting the value of $\frac{1}{n}\sum X_i$ + Laplace($1/ne$). While we omit the proof here, it is easy to show that the output of the Laplace mechanism enjoys *ε*-differential privacy.

Theorem 3.4. *The Laplace mechanism is ε-differentially private.*

To explain one slightly more complex application, in which we wish to output several statistics about the same population. Say, for example, that for a particular group of individuals, we wish to simultaneously know what fraction of them are HIV positive, what fraction are male, and what fraction are over 25; but as usual, subject to differential privacy.

We can then define $X_i \in \{0, 1\}^3$ to be the vector for individual ***i***, representing the answers to these three queries. Instantiating the Laplace mechanism for this instance gives $\frac{1}{n}\sum X_i$ + Laplace$(3/ne)^{\otimes 3}$, where a simple calculation shows that the sensitivity of $\frac{1}{n}\sum X_i$ is $3/n$ in this case. More generally, we could consider answering *k* such queries for the same dataset, and the Laplace mechanism would tell us to output $\frac{1}{n}\sum X_i$ + Laplace$(k/ne)^{\otimes k}$. This is useful, because it allows us to privately answer several questions about the same sensitive dataset. Indeed, differential privacy would not be useful if such a thing wasn't possible, as a data analyst rarely asks only one query about a dataset.

4 Properties of Differential Privacy

As hinted at, differential privacy is a robust and versatile notion of data privacy. This is one of the reasons for its success and adoption: it has many nice properties which make it very convenient for building up larger and more complex systems in a modular fashion. In this section, we cover some of the most important properties.

Post-Processing. The first property says that, once you make a release which is differentially private, no data-independent method can remove this privacy guarantee.

Theorem 4.1. *Let $M : X^n \to Y$ be ε-differentially private and let $F : Y \to Z$ be an arbitrary randomized mapping. Then $F \circ M$ is ε-differentially private.*

Group Privacy. Differential privacy guarantees a type of stability for every pair of neighbouring datasets, which seems restrictive at first glance. For example, it would be inconvenient if changing two datapoints would result in massive leakage of sensitive information. Fortunately, the group privacy property of differential privacy allows us to quantify how the guarantees weaken as more datapoints are changed.

Theorem 4.2. *Let $M : X^n \to Y$ be an ε-differentially private algorithm. Suppose X and X' are two datasets which differ in exactly k positions. Then for all $T \subseteq Y$, we have*

$$\Pr[M(X) \in T] \leq \exp(k\varepsilon) \Pr[M(X') \in T].$$

Basic Composition. Perhaps the most important property of differential privacy is *composition*. As previously mentioned, it is unrealistic that a data analyst would only wish to perform a single private analysis on a dataset. It is not a priori obvious how you would quantify the resulting privacy leakage if *k* private analyses were run instead. The composition property roughly says that the privacy parameters "add up":

Theorem 4.3. *Suppose $M = (M_1, . . . , M_k)$ is a sequence of ε-differentially private algorithms, potentially chosen sequentially and adaptively. Then M is kε-differentially private.*

Note that we already saw a similar phenomenon when asking *k* yes-no questions about a dataset using the Laplace mechanism. This result is much more general, allowing for arbitrary private algorithms and for the specific queries/mechanisms to be chosen adaptively.

5 Approximate Differential Privacy

We briefly mention a relaxation of ε-differential privacy, known as (ε, δ)-differential privacy. ε- differential privacy can be a rather stringent notion of privacy, as it requires us to control the probability of every event, even those with very low probability. This proves to be restrictive on the type of mechanisms that can be used to guarantee differential privacy. Instead, we relax to the following definition.

Definition 5.1 (Approximate Differential Privacy). *An algorithm $M : X^n \to Y$ is (ε, δ)- differentially private (i.e., it satisfies* approximate differential privacy*) if, for all neighbouring databases $X, X' \in X^n$, and all $T \subseteq Y$,*

$$\Pr[M(X) \in T] \leq e^{\varepsilon} \Pr[M(X') \in T] + \delta.$$

Note that this generalizes ε-differential privacy, which corresponds to the case where δ = 0. Very loosely speaking, (ε, δ)-DP guarantees that ε-DP holds except with probability δ.
This more permissive definition enables use of a wider variety of algorithms, including the *Gaussian* mechanism, which depends on the ℓ_2-sensitivity of a function (rather than the ℓ_1-sensitivity, as considered before).

Definition 5.2. *Let $f : X^n \to \mathbb{R}^k$. The ℓ_2-sensitivity of f is*

$$\Delta_2 = \max_{X,X'} \| f(X) - f(X') \|_2 ,$$

where X and X' are neighbouring databases.

With the ℓ_2-sensitivity in hand, we are ready to define the Gaussian mechanism, which amounts to addition of Gaussian noise of scale calibrated to the ℓ_2-sensitivity.

Definition 5.3. *Let $f : X^n \to \mathbb{R}^k$. The* Gaussian mechanism *is defined as*

$$M(X) = f(X) + (Y_1, . . . , Y_k),$$

where the Y_i are independent $N(0, 2 \ln(1.25/\delta)\Delta_2^2/\varepsilon^2)$ random variables.

The Gaussian mechanism enjoys (ε, δ)-differential privacy.

Theorem 5.4. *The Gaussian mechanism is* (ε, δ)*-differentially private.*

While (ε, δ)-differential privacy satisfies slight variations of many of the same properties as considered before (post-processing, group privacy, basic composition), one of the most important properties is *advanced* composition. Advanced composition is a useful property for when we need to perform *many* private analyses on the same dataset. To recall, basic composition says that if we run multiple private algorithms, then the privacy parameters "add up." That is, if we run k ε-differentially private algorithms, then the result will be k ε-differentially private. This can be significant if k is large. Roughly speaking, advanced composition allows us to reduce this cost quadratically, to only $\sqrt{k\epsilon}$ −differential privacy, permitting dramatically more queries at the same overall privacy budget. The precise statement of advanced composition is somewhat more nuanced, which we reproduce here.

Theorem 5.5 (Advanced Composition). *For all* $\varepsilon, \delta, \delta' > 0$, *let* $M = (M_1, \ldots, M_k)$ *be a sequence of* (ε, δ)*-differentially private algorithms, where the* M_i*'s are potentially chosen sequentially and adaptively.*
Then M *is* $(\tilde{\varepsilon}, \tilde{\delta})$*-differentially private, where* $\tilde{\varepsilon} = \varepsilon\sqrt{2k\log(1/\delta')} + k\varepsilon\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1}$ *and* $\tilde{\delta} = k\delta + \delta'$.

6 Conclusion

Now over fifteen years old, differential privacy has withstood the test-of-time and has demonstrated itself to be a rigorous yet user-friendly way to preserve individual privacy. In the time since its invention, it has now seen adoption in a variety of companies and organizations in the United States, the most notable of which is the 2020 US Census. Despite this enthusiasm and widespread deployment, Canada lags far behind the United States – we are not presently aware of any production-level deployments of differential privacy by a Canadian organization. We hope for and are happy to advise in the modernization of privacy-preserving data analytics on the sensitive data of Canadians, ranging from the Canadian Census to processing of medical data. Only through such measures will Canadians maintain their trust in Canadian organizations as the responsible data stewards of tomorrow.

References:

[AACM$^+$22] John Abowd, Robert Ashmead, Ryan Cumings-Menon, Simson Garfinkel, Micah Hei- neck, Christine Heiss, Robert Johns, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, Brett Moran, William Sexton, Matthew Spence, and Pavel Zhuravlev. The 2020 census disclosure avoidance system topdown algorithm. Harvard Data Science Review, (Special Issue 2), 2022.

[ABC$^+$20] Ahmet Aktay, Shailesh Bavadekar, Gwen Cossoul, John Davis, Damien Desfontaines, Alex Fabrikant, Evgeniy Gabrilovich, Krishna Gadepalli, Bryant Gipson, Miguel Guevara, Chaitanya Kamath, Mansi Kansal, Ali Lange, Chinmoy Mandayam, Andrew Oplinger, Christopher Pluntke, Thomas Roessler, Arran Schlosberg, Tomer Shekel, Swapnil Vispute, Mia Vu, Gregory Wellenius, Brian Williams, and Royce J. Wilson. Google covid-19 community mobility reports: Anonymization process description (version 1.0). arXiv preprint arXiv:2004.04145, 2020.

[BRS$^+$09] Rosemary Braun, William Rowe, Carl Schaefer, Jinghui Zhang, and Kenneth Buetow. Needles in the haystack: identifying individuals present in pooled genomic data. PLoS Genetics, 5(10):1–8, 2009.

[Des21] Damien Desfontaines. A list of real-world uses of differential privacy. https:// desfontain.es/privacy/real-world-differential-privacy.html, 10 2021. Ted is writing things (personal blog).

[Dif17] Differential Privacy Team, Apple. Learning with privacy at scale. https://machinelearning.apple.com/docs/learning-with-privacy-at-scale/appledifferentialprivacysystem.pdf, December 2017.

[DKY17] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. In Advances in Neural Information Processing Systems 30, NIPS ’17, pages 3571–3580. Curran Associates, Inc., 2017.

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In Proceedings of the 3rd Conference on Theory of Cryptography, TCC ’06, pages 265–284, Berlin, Heidelberg, 2006. Springer.

[EPK14] U´lfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. RAPPOR: Randomized aggregatable privacy-preserving ordinal response. In Proceedings of the 2014 ACM Conference on Computer and Communications Security, CCS ’14, pages 1054–1067, New York, NY, USA, 2014. ACM.

[HSR+08] Nils Homer, Szabolcs Szelinger, Margot Redman, David Duggan, Waibhav Tembe, Jill Muehling, John V. Pearson, Dietrich A. Stephan, Stanley F. Nelson, and David W. Craig. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. PLoS Genetics, 4(8):1–9, 2008.

[McS16] Frank McSherry. Statistical inference considered harmful. https://github.com/frankmcsherry/blog/blob/master/posts/2016-06-14.md, June 2016.

[SU20a] Thomas Steinke and Jonathan Ullman. The pitfalls of average-case differential privacy. https://differentialprivacy.org/average-case-dp/, July 2020.

[SU20b] Thomas Steinke and Jonathan Ullman. Why privacy needs composition. https://differentialprivacy.org/privacy-composition/, August 2020.

[War65] Stanley L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60(309):63–69, 1965.

Establishing FAIR and CARE Synthetic Health Data Ecosystem for Canadian Learning Health Systems and Innovation

Helen Chen[1], PhD, Shu-Feng Tsao, MSc

School of Public Health Sciences

University of Waterloo

[1]corresponding author: helen.chen@uwaterloo.ca

Abstract:

Obtaining access to real-world health data continues to be a significant challenge, mainly due to privacy and security implications. Consequently, researchers and technology innovators - particularly those operating in the health data science and AI technology development space – increasingly resort to synthetic health data to bridge the data gap. High-quality synthetic data has the potential to expedite research and development of novel technologies. However, synthetic health datasets in Canada are scarce, and no existing synthetic health datasets conform to the Findable, Accessible, Interoperable, and Reusable (FAIR) standards. Moreover, while federated machine learning offers the advantage of protecting patient privacy by not requiring the exchange of source data across nodes, it has yet to be optimized in Canada's health research environment, and there is very limited use of federated learning with synthetic health data. This paper explores the ethical considerations and value proposition of generating and sharing synthetic health data. Our goal is to facilitate the development of a reliable and sustainable synthetic data infrastructure that supports the ethical and responsible use of synthetic health data. We believe that this framework will pave the way for a more robust and secure synthetic data ecosystem, enabling the generation of valuable insights that can drive positive health outcomes for Canadians.

Introduction

The effective management and sharing of health data, especially electronic medical records (EMRs), is often impeded by diverse information systems and data formats. This fragmentation poses significant challenges in achieving seamless integration, standardized data representation, and efficient sharing across healthcare and technology innovation ecosystems (Kokosi et al., 2022; Kokosi & Harron, 2022). The handling of health data has been subjected to stringent regulations by various laws, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States (US), the Personal Information Protection and Electronic Documents Act (PIPEDA) in Canada, and the General Data Protection Regulations (GDPR) in the European Union (EU) (Shapiro, 2022). In their quest to test theories, models, algorithms, or prototypes, researchers and developers frequently rely on de-identified or anonymous, aggregated health data. However, it takes substantial time and resources to retrieve, aggregate, and de-identify data before it becomes accessible (Kokosi et al., 2022; Kokosi & Harron, 2022). Legal and ethical sharing of real health data remains challenging, despite initiatives like the Open Data Charter, which promotes Open Science and accessibility through data sharing (Huston et al., 2019). Nonetheless, successful examples like the Medical Information Mart for Intensive Care (MIMIC), a large open healthcare database, have demonstrated the value of sharing real-world healthcare data for research and innovations (Johnson et al., 2016).

One potential solution to this challenge is the creation of realistic, high-quality synthetic health datasets that mimic the complexities of the original data but do not contain any real patient information (Kokosi et al., 2022). The Clinical Practice Research Datalink (CPRD) in the United Kingdom (UK) and the Agency for Healthcare Research and Quality in the US have made synthetic datasets available for research (Synthetic data, n.d.; SyH-DR, n.d.). Synthetic health data can be valuable for health data science education, ML/AI algorithm development, and health technology innovation while safeguarding patient privacy, diversifying datasets, and enhancing health and innovative research (Gonzales et al., 2023).

In Canada, however, there is a scarcity of high-quality, sharable synthetic health datasets that adhere to Findable, Accessible, Interoperable and Reusable (FAIR) standards, despite its involvement in the Common Infrastructure for National Cohort in Europe, Canada, and Africa (CINECA) projects (CINECA, n.d.). Furthermore, there are Collective benefits, Authority to control, Responsibility, Ethics (CARE), and the First Nations principles of Ownership, Control, Access, and Possession (OCAP) pertaining to Indigenous data (Carroll et al., 2021; Mecredy et al., 2018;_Wilkinson et al., 2016), but applying or implementing these principles in generating and sharing synthetic health data remains very limited.

With the fast advent of Machine Learning (ML) and Artificial Intelligence (AI) techniques, researchers and industry have been exploring various deep learning models to generate high-quality synthetic data (Gonzales et al., 2023; Hernandez et al., 2022). Among these, generative adversarial networks (GANs) and their variants have emerged as promising synthetic data generation approaches (Goodfellow, et al, 2014, Xu, et al, 2019, Gonzales et al., 2023; Hernandez et al., 2022; Murtaza et al., 2023). One advantage of AI techniques for generating synthetic data over the conventional de-identification methods is the potentially high-fidelity data quality with a very low risk of one-to-one reverse-engineering back to the original health data at the population level (El Emam et al., 2020; Hernandez et al., 2022; Rajotte et al., 2022).

Additionally, federated learning shows promise as another technique to safeguard data privacy and security by training AI models without centralizing datasets across multiple network nodes, therefore further reducing critical data compromises (Antunes et al., 2022; Brisimi et al., 2018). However, its optimization and broader implementation remain ongoing challenges.

Combining generative AI models and federated learning to generate synthetic health data following FAIR principles, with additional consideration of the CARE principles for Indigenous data, can create a robust and optimal health data network. Such an approach would protect sensitive patient data and accelerate health research and innovations (Antunes et al., 2022; Brisimi et al., 2018).

While AI researchers are consistently seeking new ways to improve the performance of synthetic data generation, other important aspects, such as trust, transparency, and governance of synthetic health data, still require comprehensive study and consensus building. This paper presents key considerations for establishing a pan-Canadian synthetic health data ecosystem which aims to enable learning health systems and foster health technology innovation, providing the foundation for future research and implementation in synthetic health data.

Synthetic Health Data Landscape

A scoping rapid review was conducted to review literature exploring the use of synthetic health data related to health research in general by following the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) and the World Health Organization's (WHO) rapid review approach (Tricco et al., 2012, 2017). Searches were completed on PubMed, Scopus, and Google Scholar using keywords and their variations listed in Table 1. We included peer-reviewed articles and grey literature written in English and published between January 01, 2012, to March 31, 2023. Meta-analyses were excluded because they gave unrelated articles simply due to the term "synthetic" in the search queries. Methodological papers related to improvements of models or AI algorithms were also excluded due to the topic having been reviewed extensively in existing literature (Gonzales et al., 2023; Hernandez et al., 2022; Murtaza et al., 2023).

synthetic health data, synthetic data quality, synthetic data utility, synthetic data governance, synthetic data privacy, synthetic data sharing, data consents, indigenous health data, indigenous health data governance

Table 1 keywords used for literature search.

Methodologies and results of this review can be found in this paper (Tsao, et. al, 2023).

Key Finding I: Data Sources Used to Generate Synthetic Health Data

Data sources used for generating synthetic health data can be broadly categorized as follows: (1) EMRs, (2) health insurance claims, (3) other administrative or health surveys, (4) bioinformatics, (5) medical images, and (6) sensor data. Depending on how synthetic health data is generated, data in these categories can be treated as longitudinal or cross-sectional in corresponding studies.

Despite significant advancements in Natural Language Processing (NLP) techniques, particularly the recent emergence of large language models (LLMs) like ChatGPT (Brekke et al., 2021; Tang et al., 2023), unstructured health data, such as doctors' notes in the EMRs, have been underutilized in synthetic health data generation.

Researchers often rely on open health data, such as the MIMIC-III, in synthetic health data research. This open dataset is attractive not only because it is freely available but also because it enables the reproducibility of research.

While it is widely acknowledged that synthetic health data offers benefits such as preserving privacy, accelerating research, and driving innovations (Chen et al., 2021; Rajotte et al., 2022), some caution against its usage, as well as current barriers to the adoption of synthetic health data are discussed below.

Key Finding II: Concerns of Research Ethics and Legal Regulations

Typically known as the Institutional Review Board (IRB) in academic health centres or Research Ethics Board (REB) in Canadian universities, research ethical governance bodies have universally agreed that studies involving any original or real health data require ethical approvals according to legal regulations (Bassan & Harel, 2018; Nass et al., 2009). However, there is a lack of clear consensus within Canada's REBs when it comes to studies using synthetic health data exclusively. In contrast to de-identified or anonymous real patient data, the regulatory landscape governing synthetic health data and safeguarding patient privacy has remained ambiguous (Bill C-27, n.d.). This ambiguity has impacted the generation and sharing of synthetic data in Canada, with REBs granting ethics approval on a case-by-case basis for such research. Consequently, many legal and ethical questions surrounding synthetic health data in Canada remain unanswered (Arora & Arora, 2022).

In the context of EU's GDPR, synthetic data can be categorized as pseudonymous data, anonymous data, or both, depending on the specific context of its use (López & Elbi, 2022). In the US, if synthetic data is appropriately created, it is exempt from the HIPPA regulations (Varma, 2022). Conversely, in Canada, neither the PIPEDA nor the Consumer Privacy Protection Act (CPPA) explicitly addresses synthetic data, as Canada is still undergoing legal reforms in this area (Bill C-27, n.d.).

Another significant concern revolves around patients' informed consent. In the US, HIPAA classifies the creation of de-identified data as the healthcare operations of a covered entity, thus eliminating the requirement for informed consent from patients, even if the de-identified data will be used for research (Nass et al., 2009).

In the EU, synthetic health data can be considered as either de-identified or anonymous data under GDPR, but regardless, informed consent is still mandated (Shapiro, 2022). In contrast, Canada's PIPEDA and CPPA generally require organizations to obtain informed consents when individuals' personal information is collected, used, or disclosed (Shapiro, 2022). The lack of clarity in PIPEDA and CPPA has made synthetic health data a grey area, resulting in potentially lengthy and inconsistent ethical reviews and further impending health research and innovations.

Key Finding III: Evaluations of Synthetic Health Data Utility

Synthetic Data generation models, the utility of the generated health data, and privacy concerns of that synthesized health data are co-related. So far, there have been no universal standards to generate health synthetic data, even though GANs have been the most popular ML technique to generate various health synthetic data (Gonzales et al., 2023; Hernandez et al., 2022; Murtaza et al., 2023). These models produce robust synthetic data (as long as the real-world data is also robust) by preserving the cohort characteristics and trends in the real-world data. However, overfitting in these models could become problematic in privacy preservation as it can cause some synthesized records to be too similar to the real-world data (Bhanot, et al., 2021).

Given no common methods for generating health synthetic data, there are no universal evaluation metrics to assess the quality and utility of health synthetic data. Individual studies generally have their own evaluation metrics to examine how closely the health synthetic data have reflected their data sources and specific use cases. Gordon et al. (2021) have proposed a data utility evaluation matrix as a framework for health data curation, which can be modified to evaluate synthetic data. Gordon's data utility framework consists of five categories: data documentation, technical quality, coverage, access provision and value and interest. Each category has four quality levels, ranging from bronze as the lowest quality to platinum as the highest (Gordon et al., 2021).

Key Finding IV: Generation of Synthetic Health Data in Indigenous Data Governance Context

Currently, there are no standardized data-sharing principles stated for synthetic health data. However, the FAIR and CARE principles are two main guidelines for data sharing, with the latter addressing Indigenous data (Carroll et al., 2020; Kush et al., 2020; Wilkinson et al., 2016). The FAIR principles have been applied frequently in real-world datasets, with FAIRsharing.org as a platform to share data that meets FAIR principles (Sansone et al., 2019). However, implementing FAIR principles has remained limited for health synthetic data sharing.

In the synthetic health data network, ensuring the inclusion of Indigenous people's data and upholding their right to control and access it is paramount (Carrol et al., 2021). Historical injustices and the unethical treatment of Indigenous peoples have strained relationships between Indigenous peoples and the government, leading to data-sharing restrictions (Bosacrino et al., 2022). This exclusion of Indigenous datasets poses limitations in fields such as synthetic health data, which could greatly benefit from insights provided by Indigenous datasets to develop FAIR and CARE ML/AI techniques. The scarcity of representative Indigenous datasets impacts the accuracy of generative models, inadvertently introducing biases and influencing data-driven decisions regarding Indigenous health.

To promote healthcare access and equality for Indigenous peoples, accurate Indigenous data representation is essential (Walker et al., 2017). Synthetic health data offers a promising solution to bridge this knowledge gap. However, this goal requires partnerships between the Indigenous community, AI technology developers, and governments to address this limitation. Presently, there is no clear governance framework for generating and utilizing Indigenous synthetic health data. However, the FAIR and CARE principles can be the foundation for Indigenous synthetic data governance. The FAIR principles would act as guidelines for data producers and repositories, and the CARE principles would extend to the 'people' or 'purpose' of why that data is being used

(Carrol et al., 2021). While data stakeholders inform data reuse and research reproducibility, the CARE principles can address historical inequities and provide Indigenous peoples with a platform for preserving data sovereignty.

Generative AI holds great promise for advancing health research and accelerating technology innovations. Despite its potential, generative AI's widespread adoption and utilization in the health sector are yet to be fully realized. To ensure its responsible and effective application, there is a pressing need for further research and establishing regulatory guidelines focusing on data governance and quality evaluation standards. In light of this, we present the following recommendations.

Recommendations

Recommendation I: Establish Synthetic Health Data Governance Guidelines

Conducting a comprehensive policy analysis concerning synthetic data governance is imperative to compare the EU's GDPR, the US's HIPAA, and Canada's PIPEDA, CPPA, and Artificial Intelligence and Data Act (AIDA) (Bill C-27, n.d.). This policy analysis aims to understand how the EU and US address synthetic health data compared to Canada's regulations, identify potential gaps in policies related to synthetic health data, and formulate recommendations for policy amendment or changes in Canada.

Additionally, this policy analysis needs to examine the methods employed to obtain patient consent regarding the use of their health data to generate synthetic health data. Following the completion of the policy analysis, a systematic three-round electronic Delphi survey (Nasa et al., 2021; Okoli & Pawlowski, 2004) can be conducted, inviting experts, data owners (patients), data custodians, and data users to co-design the ethical guidelines and synthetic health data governance framework. Key players shaping Canada's data strategy include Canadian citizens, the Canadian Institute for Health Information (CIHI), Health Canada, Statistics Canada, the First Nations Information Governance Centre (FNIGC), the ethics review boards within academic institutions, and health organizations; all of whom need to work together to co-design the ethical guidelines and data governance framework. Their trust and collaboration are instrumental in successfully developing and deploying synthetic health data research and data uses.

Throughout the Delphi study, an extensive list of factors crucial to the governance of synthetic health data will be identified, with particular attention to factors aligned with the FAIR and CARE principles. The team will assess the level of agreement among experts and potential users regarding the importance of these factors, aiming to bridge the differences and similarities. Should further clarifications be necessary, stakeholders will be interviewed after the Delphi surveys.

Furthermore, most current consent directives for EMRs do not stipulate the possible use of patients' information for producing synthetic health data to be used by hospitals and third-party entities. Policy updates also need to explicitly address whether synthetic health data will be governed as human subject data. Clear and specific policy guidelines are essential in establishing a standardized framework that governs the generation and ethical use, protection, and sharing of synthetic health data. Such guidelines ensure transparency, accountability, and compliance with regulatory requirements, ultimately fostering trust among stakeholders and facilitating data-driven research and innovation in the health domain.

Recommendation II: Demonstrate the Responsible and Beneficial Use of Synthetic Health Data

The deep learning models for synthesizing data, the usefulness of the generated health data, and privacy concerns over sharing such data are interconnected. While the list of companies specializing in synthetic data generation is expanding, including Replica Analytics®, MDClone®, Mostly AI®, to name a few, it is important to note that there is no one-fits-all standardized algorithm or commercialized tool available for generating high-quality synthetic health data. The inherent challenges arise due to the health data's high-dimensional nature and complex interrelations. Major data custodians, such as hospitals, CIHI and Statistics Canada, are key players in exploring the benefits and limitations of generative AI algorithms. Conducting in-depth assessments of the privacy risks and the utility associated with their specific use cases can offer valuable insights. By gaining deeper insights into the potential and challenges of generating and using synthetic health data, all stakeholders can make responsible and well-informed choices that align with their specific objectives and data privacy considerations. These explorations will undoubtedly pave the way for more informed decisions regarding the incorporation of synthetic health data in data management strategies across the entire spectrum of the Canadian health research and innovation ecosystem.

Recommendation III: Establish Standards for the Evaluation and Publication of Synthetic Health Data

The existing literature on evaluations of synthetic health data reveals knowledge gaps that require further exploration. Given that generative AI algorithms are normally tailored for types of data (i.e., images, tabular data, time series, or genome data) and fine-tuned for specific types of use cases (i.e., software testing, epidemiological study, operation optimization, etc.), it is important to assess the quality and value of synthetic health data to fulfill these diverse needs. However, as there are no universal evaluation metrics to assess the performances of generative AI models and the quality of synthetic health data, the evaluation metrics will be co-designed to assess model performances and the quality of synthetic health data based on FAIR and CARE principles.

The key components of synthetic data assessment can be categorized into the following three dimensions:

Privacy Risk: The calculation of the privacy risk involves assessing the uniqueness or identifiability of individuals in the dataset. Since data imbalance is common in health data, it is imperative to employ best practices to eliminate the risk of re-identification in the original data before using it to generate synthetic data. Using privacy-preserving ML/AI models to generate synthetic health data should be investigated and preferred. When calculating the privacy risk, the synthetic health data can be assessed for the risk of re-identification with a common threshold of 0.09 (El Emam et al., 2020). Additional measurements, such as cosine-similarity (NIST, n.d.) between a record in the synthetic dataset and records in the original dataset, can serve as valuable tools to further mitigate the risk of synthetic data records matching those in the original real dataset.

Fidelity and Utilities: In comparative studies of synthetic data quality, the evaluation typically assesses how closely synthetic variable distributions resemble those of the respective real data variables. The correlation matrix of key variables serves as another common measure of gauging the fidelity between synthetic and real datasets. Additionally, the quality of synthetic data is further scrutinized by examining the performance of a machine learning model(s) built with both synthetic and real datasets (Muller et al. 2022, Foraker et al., 2020, El Emam et al. 2021). These initial empirical indicators provide valuable insights into the data quality of a given synthetic dataset.

However, it is important to acknowledge that such evaluations are not exhaustive and highly dependent on specific use cases and the intended goals of the model. Nonetheless, establishing a minimal set of fidelity measurements and benchmarking machine learning models for assessing the utility of synthetic data marks a significant step toward standardizing synthetic data quality measures. Publishing these quality indicators will bolster confidence in synthetic data generation and ensure responsible and informed usage by third parties, equipped with explicit knowledge of the limitation inherent in the synthetic dataset.

Cost-Benefit Assessment: Similar to health technology assessment (HTA), there is a need for a cost-benefit framework that quantifies the societal impacts of generating and using synthetic health data for research and commercial purposes.

The first challenge in developing this cost-benefit analysis framework is accurately capturing the costs and benefits to each stakeholder. The first step would be recommending a measurement to calculate the benefits of creating and sharing synthetic health data. In HTA, the quality-adjusted life years (QALY) is the academic standard used to evaluate how well medical treatments lengthen and/or improve patients' lives. In the evaluation of synthetic health data, the benefit derived from its utilization may not be directly reflected in QALY. The benefit measurement may require considerations of additional factors, including the likelihood of improved patient well-being resulting from the proposed research or commercial products. This encompasses advanced data analytics and the value of innovative technology, which could lead to better patient outcomes and optimized healthcare delivery.

The second challenge is estimating the probability of individual privacy compromise when releasing synthetic health data. Such privacy breaches could arise either through membership inferencing or reverse engineering of the synthetic health data itself, or its combination with other external sources of information. Merely assessing re-identification probability is insufficient to evaluate the overall costs of synthetic health data release fully. It is important to estimate the potential harm inflicted upon individuals to capture the total costs borne by stakeholders.

Furthermore, this framework must take into account the diverse range of stakeholder groups, considering the potential systemic harmful effects experienced by certain communities. Improper handling of synthetic data generation could exacerbate systematic bias in the original data. Therefore, a comprehensive approach is necessary to ensure that equity, diversity, inclusion, and anti-racism (EDI-R) considerations are effectively integrated into the cost-benefit analysis.

As we consider the broader use of synthetic health data, we must be vigilant in avoiding the “garbage in, garbage out” problem, by recognizing the significance of the quality of the original real data. Synthetic data presents a promising venue to mitigate privacy risks while sharing data, thereby promoting the democratization of data for diverse applications. These data usages will help identify potential data gaps and quality issues and incentivize data owners and custodians to prioritize maintaining high-quality original data. This, in turn, fosters a positive cycle of data-driven advancement in Canadian’s health ecosystem.

References:

Antunes, R. S., André da Costa, C., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated learning for healthcare: Systematic review and architecture proposal. ACM Transactions on Intelligent Systems and Technology, 13(4), 1–23. https://doi.org/10.1145/3501813

Arora, A., & Arora, A. (2022). Synthetic patient data in health care: a widening legal loophole. Lancet, 399(10335), 1601–1602. https://doi.org/10.1016/s0140-6736(22)00232-x

Bassan, S., & Harel, O. (2018). The ethics in synthetics: statistics in the service of ethics and law in health-related research in big data from multiple sources. Journal of Law and Health, 31(1), 87-117.

Bhanot, K., Qi, M., Erickson, J. S., Guyon, I., & Bennett, K. P. (2021). The problem of fairness in synthetic healthcare data. Entropy (Basel, Switzerland), 23(9), 1165. https://doi.org/10.3390/e23091165

Bill C-27: An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts. (2022, November 4). Justice.Gc.Ca. https://www.justice.gc.ca/eng/csj-sjc/pl/charter-charte/c27_1.html

Boscarino, N., Cartwright, R. A., Fox, K., & Tsosie, K. S. (2022). Federated learning and Indigenous genomic data sovereignty. Nature Machine Intelligence, 1-3. https://doi.org/10.1038/s42256-022-00551-y

Brekke, P. H., Rama, T., Pilán, I., Nytrø, Ø., & Øvrelid, L. (2021). Synthetic data for annotation and extraction of family history information from clinical text. Journal of Biomedical Semantics, 12(1), 11. https://doi.org/10.1186/s13326-021-00244-2

Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C., & Shi, W. (2018). Federated learning of predictive models from federated Electronic Health Records. International Journal of Medical Informatics, 112, 59–67. https://doi.org/10.1016/j.ijmedinf.2018.01.007

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for indigenous data governance. Data Science Journal, 19. https://doi.org/10.5334/dsj-2020-043

Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F. K., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. Nature Biomedical Engineering, 5(6), 493–497. https://doi.org/10.1038/s41551-021-00751-8

CINECA - common infrastructure for national cohorts in Europe, Canada, and Africa. (n.d.). CINECA. Retrieved December 12, 2022, from https://www.cineca-project.eu/

El Emam, K., Mosquera, L., & Bass, J. (2020). Evaluating identity disclosure risk in fully synthetic health data: Model development and validation. Journal of Medical Internet Research, 22(11), e23139. https://doi.org/10.2196/23139

El Emam K, Mosquera L, Jonker E, Sood H. Evaluating the utility of synthetic COVID-19 case data. JAMIA Open. 2021 Mar 1;4(1):ooab012. doi: 10.1093/jamiaopen/ooab012. PMID: 33709065; PMCID: PMC7936723.

Foraker, RE, and others, Spot the difference: comparing results of analyses from real patient data and synthetic derivatives, JAMIA Open, Volume 3, Issue 4, December 2020, Pages 557–566, https://doi.org/10.1093/jamiaopen/ooaa060

Gonzales, A., Guruswamy, G., & Smith, S. R. (2023). Synthetic data in health care: A narrative review. PLOS Digital Health, 2(1), e0000082. https://doi.org/10.1371/journal.pdig.0000082

Goodfellow Ian, Pouget-Abadie Jean, Mirza Mehdi, Xu Bing, Warde-Farley David, Ozair Sherjil, Courville Aaron, and Bengio Yoshua. 2014. Generative adversarial nets. In International Conference on Advances in Neural Information Processing Systems. 2672–2680.

Gordon, B., Barrett, J., Fennessy, C., Cake, C., Milward, A., Irwin, C., Jones, M., & Sebire, N. (2021). Development of a data utility framework to support effective health data curation. BMJ Health & Care Informatics, 28(1), e100303. https://doi.org/10.1136/bmjhci-2020-100303

Hernandez, M., Epelde, G., Alberdi, A., Cilla, R., & Rankin, D. (2022). Synthetic data generation for tabular health records: A systematic review. Neurocomputing, 493, 28–45. https://doi.org/10.1016/j.neucom.2022.04.053

Huston, P., Edge, V. L., & Bernier, E. (2019). Reaping the benefits of Open Data in public health. Releve Des Maladies Transmissibles Au Canada [Canada Communicable Disease Report], 45(11), 252–256. https://doi.org/10.14745/ccdr.v45i10a01

Johnson, A. E. W., Pollard, T. J., Shen, L., Lehman, L.-W. H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A., & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. Scientific Data, 3(1), 160035. https://doi.org/10.1038/sdata.2016.35

Kokosi, T., De Stavola, B., Mitra, R., Frayling, L., Doherty, A., Dove, I., Sonnenberg, P., & Harron, K. (2022). An overview on synthetic administrative data for research. International Journal for Population Data Science, 7(1). https://doi.org/10.23889/ijpds.v7i1.1727

Kokosi, T., & Harron, K. (2022). Synthetic data in medical research. BMJ Medicine, 1(1), e000167. https://doi.org/10.1136/bmjmed-2022-000167

Kush, R. D., Warzel, D., Kush, M. A., Sherman, A., Navarro, E. A., Fitzmartin, R., Pétavy, F., Galvez, J., Becnel, L. B., Zhou, F. L., Harmon, N., Jauregui, B., Jackson, T., & Hudson, L. (2020). FAIR data sharing: The roles of common data elements and harmonization. Journal of Biomedical Informatics, 107(103421), 103421. https://doi.org/10.1016/j.jbi.2020.103421

López, C. A. F., & Elbi, A. (2022, September 22). On synthetic data: a brief introduction for data protection law dummies. European Law Blog. https://europeanlawblog.eu/2022/09/22/on-synthetic-data-a-brief-introduction-for-data-protection-law-dummies/

Mecredy, G., Sutherland, R., & Jones, C. (2018). First Nations data governance, privacy, and the importance of the OCAPÂ® principles. International Journal for Population Data Science, 3(4). https://doi.org/10.23889/ijpds.v3i4.911

Muller, E., Zheng, X. and Hayes, J. Evaluation of the Synthetic Electronic Health Records. In Proceedings of the 1st Workshop on Scarce Data in Artificial Intelligence for Healthcare (SDAIH 2022), pages 17-22 DOI: 10.5220/0011531300003523

Murtaza, H., Ahmed, M., Khan, N. F., Murtaza, G., Zafar, S., & Bano, A. (2023). Synthetic data generation: State of the art in health care domain. Computer Science Review, 48(100546), 100546. https://doi.org/10.1016/j.cosrev.2023.100546

Nasa, P., Jain, R., & Juneja, D. (2021). Delphi methodology in healthcare research: How to decide its appropriateness. World Journal of Methodology, 11(4), 116–129. https://doi.org/10.5662/wjm.v11.i4.116

Nass, S. J., Levit, L. A., Gostin, L. O., & Institute of Medicine (US) Committee on Health Research and the Privacy of Health Information: The HIPAA Privacy Rule. (2009). HIPAA, the Privacy Rule, and its application to health research. National Academies Press. https://www.ncbi.nlm.nih.gov/books/NBK9573/

National Institute of Standards and Technology (NIST) (n.d.) COSINE DISTANCE, COSINE SIMILARITY, ANGULAR COSINE DISTANCE, ANGULAR COSINE SIMILARITY (nist.gov)

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. Information & Management, 42(1), 15–29. https://doi.org/10.1016/j.im.2003.11.002

Rajotte, J.-F., Bergen, R., Buckeridge, D. L., El Emam, K., Ng, R., & Strome, E. (2022). Synthetic data as an enabler for machine learning applications in medicine. IScience, 25(11), 105331. https://doi.org/10.1016/j.isci.2022.105331

Sansone, S.-A., McQuilton, P., Rocca-Serra, P., Gonzalez-Beltran, A., Izzo, M., Lister, A. L., Thurston, M., & FAIRsharing Community. (2019). FAIRsharing as a community approach to standards, repositories and policies. Nature Biotechnology, 37(4), 358–367. https://doi.org/10.1038/s41587-019-0080-8

Shapiro, J. (2022, May 20). Why digital privacy is so complicated. Progressive Policy Institute. https://www.progressivepolicy.org/publication/why-digital-privacy-is-so-complicated/

Synthetic data. (n.d.). Cprd.com. Retrieved December 12, 2022, from https://cprd.com/synthetic-data

Synthetic healthcare database for research (SyH-DR). (n.d.). Ahrq.gov. Retrieved December 12, 2022, from https://www.ahrq.gov/data/innovations/syh-dr.html

Tang, R., Han, X., Jiang, X., & Hu, X. (2023). Does synthetic data generation of LLMs help clinical text mining? https://doi.org/10.48550/ARXIV.2303.04360

Tricco, A. C., Langlois, E. V., & Straus, S. E. (2017). Rapid reviews to strengthen health policy and systems: A Practical Guide. World Health Organization.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for Scoping Reviews (PRISMA-ScR): Checklist and explanation. Annals of Internal Medicine, 169(7), 467–473. https://doi.org/10.7326/M18-0850

Tsao SF, Sharma K, Noor H, Forster A, Chen H. Health Synthetic Data to Enable Health Learning System and Innovation: A Scoping Review. Stud Health Technol Inform. 2023 May 18;302:53-57. doi: 10.3233/SHTI230063. PMID: 37203608.

Walker, J., Lovett, R., Kukutai, T., Jones, C., & Henry, D. (2017). Indigenous health data and the path to healing. Lancet, 390(10107), 2022-2023. https://doi.org/10.1016/S0140-6736(17)32755-1

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(1), 160018. https://doi.org/10.1038/sdata.2016.18

Xu L., Skoularidou M., Cuesta-Infante A. and Veeramachaneni K.,2019, Modeling tabular data using Conditional GAN. In Advances in Neural Information Processing Systems, Wallach H., Larochelle H., Beygelzimer A., d'Alché-Buc F., Fox E., and Garnett R. (Eds.), Vol. 32. Curran Associates, Inc.

Blockchain Privacy and Its Applications in Healthcare

Guang Gong

Department of Electrical and Computer Engineering
University of Waterloo
Waterloo Ontario N2L 3G1, Canada ggong@uwaterloo.ca, https://uwaterloo.ca/scholar/ggong

Abstract

In this short survey, we first introduce some basic concepts of blockchain technologies and Bitcoin networks, address the problem on privacy in Blockchain systems, and explain a current effective cryptographic tool - zero-knowledge (ZK) proofs employed for preserving blockchain privacy. We then present our new findings for tackling the challenges on blockchain privacy using zkSNARK systems when applying to health care. Finally, we conclude the survey by providing remarks on the adoption of blockchain technologies in security and privacy of health data.

Keywords: Blockchain, privacy, zero knowledge proof, zkSNARK, health data

1 Introduction

With the advancement of science and communication technology, traditional paper-based text file systems are being steadily replaced by digital data systems, such as electronic medical records (EMR). Over the past three years, the COVID-19 pandemic has significantly impacted people's lives, making the demand for digital health systems and smart healthcare increasingly pressing. The healthcare industry and medical research sectors now require greater interoperability, compliance, verifiability, privacy, and traceability of patient records and medical information.

Furthermore, numerous digital health applications and Internet-of-Medical-Things (IoMT) have surfaced, such as the Apple Health personal medical information monitoring system. However, this proliferation of digital health technologies has also heightened the risk of sensitive personal health information being compromised or leaked. Consequently, securely sharing health data with multiple stakeholders and addressing the issue of fragmented and incomplete medical records in heterogeneous environments [4] constitutes significant challenges.

We are interested in exploring the application of decentralized approaches, particularly blockchain technologies, to address security and privacy concerns within health data systems.

1.1 Existing security mechanisms

Our primary inquiry revolves around securing and preserving the privacy of massive health data records. To achieve this, it is essential to effectively implement existing security mechanisms and technologies. Presently, various basic security mechanisms exist to ensure confidentiality, integrity, and authenticity. Additionally, there are single-sign-on (SSO) protocols available for accessing resource servers. We will elaborate on these concepts in the following sessions (for further details, please refer to [5]).

Confidentiality: Confidentiality serves to maintain authorized restrictions on information access and disclosure, encompassing the protection of personal privacy and proprietary information. There are several methods available to achieve confidentiality, which are summarized below:

1. Encryption: Utilizing encryption involves transforming data into an unreadable format through cryptographic algorithms. Only the intended receiver can recover the information using a decryption key, ensuring its confidentiality.

2. Access control: Confidentiality can also be achieved through access control. Access control involves implementing rules and policies that restrict access to confidential information only to individuals or systems with a legitimate "need to know." This determination can be based on various factors, including identity (such as a person's name or a computer's serial number) or role-based access control. However, it is important to note that relying solely on access control provides only a limited level of protection for confidentiality.

3. Entity authentication: The third way to achieve confidentiality is through entity authentication, which is the determination of the identity of role that someone has, for example:
- Something the person has (like a radio token with cryptographic keys)
- Something the person knows (e.g., password)
- Something the person is (like a human with a fingerprint)

However, this also must be used in conjunction with encryption for preventing cyber-attacks on server systems.

4. Physical security: This is the establishment of physical barriers to limit access to protected computational recourses (i.e., locks on cabinets and doors, or a shield room for a server, etc.). Similarly, this cannot solely provide strong confidentiality of data.

Integrity and authenticity: Integrity check and authenticity is to guard against improper information modification or destruction, including ensuring information nonrepudiation and authenticity. Some naive ways include backups, checksums, and error correcting code; i.e., data stored with a small change can be detected and corrected. However, those are very weak protection. The main secure way to achieve integrity and authenticity is through the following two cryptographic methods:

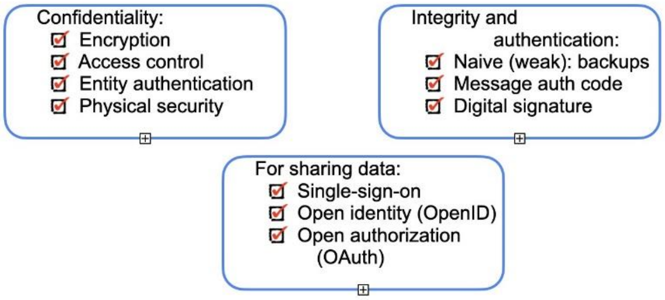
1. a **message authentication codes** (MAC) when a secret, (i.e., an authentication key), is pre-shared between two communication entities or,
2. a **digital signature scheme** if they do not pre-share any secret.

Those mechanisms can provide a strong protection on integrity and authenticity.

Data sharing There are several security mechanisms for issuing single sign-on authentication tokens to use resource servers; i.e., Kerberos and OAuth. Kerberos issues a sign-on authentication token to allow a user to access a resource server (e.g., the data store in the resource server).

OAuth, standardized by IETF, The OAuth Authorization Protocol, is an authorization protocol that allows one service provider (aka a third-party client, e.g., an insurance company), who is not an owner, to access the resources from another service provider (aka a resource server, e.g., a hospital data centre) on behalf of the owner (e.g., a patient or doctors) without exposing the owner's secret credentials to the third-party client. This is a centralized system, in which all those entities have to trust an authentication server, which is extremely challenging to implement this trust role. OpenID is an open authentication protocol that works on top of the OAuth framework.

We summarize those security mechanisms in the following figure.



Are they enough? Absolutely they are NOT; due to the facts that IoMT, smart devices, and information systems will form the backbones of healthcare systems, and AI will eventually dominate the health data system for providing data analytics and cloud service for sustaining storage. The existing security mechanisms cannot solve those problems.

1.2 Problems and challenges in securing health data

In the following, we outline specific challenging problems in security and privacy in health data.

(1) Health data has a vast degree of diversity, since it is created from IoMT, smart devices, and AI.

(2) How to implement security and privacy in the framework of Person-Centred Data by Design [10] enabled healthcare systems?

(3) How to retrieve and share them across multi-stakeholders (for individual, clinical, and analytical access and use) and across provinces?

– Health data themselves must be encrypted during storage and transfer, since encryption is necessary for confidentiality of sensitive data, so that the contents of those data are secure from corruption or damage, or unauthorized access, and from malware and cybersecurity threats.

(4) How to preserve patients' privacy (such as location, load, traffic of interaction with doctors/care providers, etc.) or doctors' privacy (e.g., proprietary treatments), to resist various cyberattacks?

The above enquiries constitute extremely challenging problems, which cannot be solved by existing security technologies! New security and privacy protection mechanisms are needed for implementing those functionalities.

In the remainder of this short survey, we introduce some basic concepts of blockchain technologies and Bitcoin blockchain in Section 2. In Section 3, we address the problem of privacy in blockchain systems and introduce the cryptographic tool known as zero-knowledge (ZK) proofs for preserving blockchain privacy. In Section 4, we present our new findings for tackling those challenges when applying to health care. Finally, we conclude the survey by providing remarks on the adoption of blockchain technology in health data in Section 5.

2 Some basic concepts of blockchain networks

The implementation of blockchain technology is based on peer-to-peer transmission, consensus mechanisms, and cryptographic algorithms. A smart contract will have an additional capability to automatically process small pieces of code. Most of the contents of this section are taken from [9].

2.1 Transactions, blocks, and chains

A blockchain can be described as a database shared among multiple distributed nodes, and the data stored on it is highly unforgeable, verifiable, access traceable, and transparent, without the need for a public trusted authority. Blockchains can provide trusted consensus, computation, and immutable data between untrusted entities. There are three basic components in the blockchain structure:

(1) Transaction: refers to all operations on the ledger.
(2) Block: It is used to record all the transactions and states within a specified time interval and can be considered as a consensus result for the current state of the ledger.
(3) Chain: A set of blocks which are listed in chronological order and can be considered as a historical log of all the time nodes of the whole ledger.

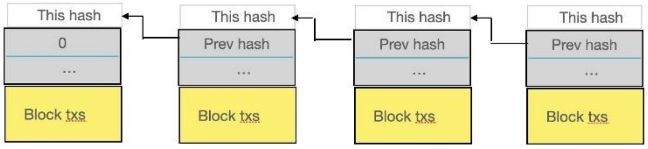


Fig. 1. A blockchain structure

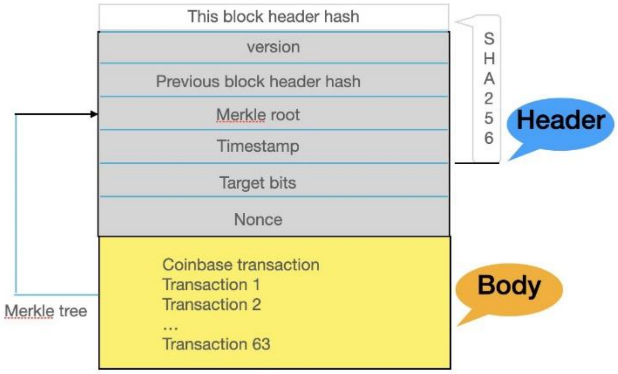


Fig. 2. A block structure in Bitcoin blockchain

In summary, a blockchain is a singly linked list of blocks by a hash function succeedingly applied to each block, sketched in Figure 1, and each block contains multiple transactions. It contains 64 transactions as shown in Figure 2.

How does each node agree on the block? This is resolved by running a consensus protocol (e.g., PBFT, PoW, PoS, etc.).

2.2 Types of blockchains

There are four types of blockchains; i.e., public chain; permissionless or permissioned; private chain; and consortium chain [1]. The core difference between them is the read and write access and the degree of decentralization. Public chains are open to everyone, private chains are open to individuals who meet certain conditions, and consortium chains are open to authorized organizations or institutions. In essence, consortium chains also belong to private chains, but the degree of privatization is different. Generally speaking, the higher the degree of decentralization, the higher the trustworthiness, and the lower the transaction efficiency. This is summarized in Table 1.

Table 1. Comparison of Different Types of Blockchain

1. Type	**2. Ledger-keeping Rights**	**3. Degree of Decentralization**	**4. Performance**
5. Public	6. All	7. High	8. Low
9. Permissioned	10. Partial	11. High	12. Low
13. Consortium	14. Partial	15. Medium	16. Medium
17. Private	18. Private	19. High	20. Low

2.3 Bitcoin blockchain

Bitcoin blockchain was proposed by Satoshi Nakamoto in 2008. Satoshi Nakamoto is the name used by the presumed pseudonymous person or persons who developed Bitcoin, authored the bitcoin white paper, and created and deployed bitcoin's original reference implementation, entitled as Bitcoin: A Peer-to-Peer Electronic Cash System, which can be accessed at https://bitcoin.org/bitcoin.pdf.

In order to have digital cash systems, we need to investigate the following questions.

(1) How to perform digital cash transactions? Traditionally, this is conducted by the bank.

(2) How to replace banks? In general, basic functions that a bank provides are as follows.
– Identity management
– Transactions
– Prevents double spending

(3) How can we enforce these properties through cryptography?

A Bitcoin blockchain is a distributed network system consisting of computers (miners) which maintains a collective bookkeeping via the Internet. Each miner node owns an exact full copy of the bookkeeping, and the bookkeeping is neither closed nor controlled by one party. This fully distributed publicly available ledger (bookkeeping) is referred to as a blockchain.

2.4 Transactions in Bitcoin blockchain

How are transactions formed in Bitcoin blockchain? Each user, say user X, has its private-key and public-key pair (sk_X , pk_X) where pk_X serves as its (wallet) address (in order to achieve privacy, each entity can have multiple public-keys or pseudo public-keys).

An electronic coin is defined as a chain of digital signatures. Each owner transfers the coin to the next by digitally signing a hash of the previous transaction and the public-key of the next owner and adding those to the end of the coin. A payee can verify the signatures to verify the chain of ownership.

For example, when user A with her public-key pk_A wants to send amt=$3 in Bitcoin currency to B with his public-key sk_B, she forms a transaction T x = (pk_A, pk_B, 3) and digitally signs that, i.e., Sig_A(hash(previous Tx), T x).

The signed data is broadcast to the network and, if valid, it ends up in a block in the blockchain with the consensus among all the miners! When the transaction is in the network, each miner will validate a transaction using cryptography (i.e., signature, authorization, and spending once). Specifically, it needs:

- validating the sender's signature,
- validating spending authorization, and
- validating that the entity A referenced the input transaction once.

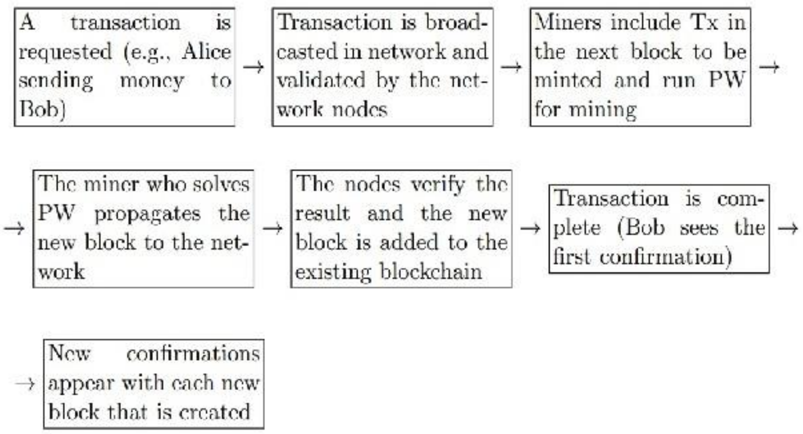
Thus, a Bitcoin is a chain of transactions tracing its flow from mining up until the current owner address, i.e., a coin is a chain of digital signatures. By the property of a digital signature, no one can forge it, but everyone can verify that!

2.5 Proof-of-Work (PW) - Consensus protocol in Bitcoin

How to prevent double spending? How does a payee get proof that at the time of each transaction, the majority of nodes agreed that it is the first received? Those tasks are achieved by implementing a distributed timestamp server using the concept of Proof-of-Work (PW).

Basically, PW is a mining process, i.e., a miner (e.g., a computer) finds a Nonce (i.e., pre-image) in the block such that its hash value has the required number of zeros (e.g., currently it requests 74 consecutive zeros in the hash value). This is an extremely expensive computation and costs massive electrical energy for doing this computation.

When the first miner finds the pre-image, he broadcasts it, and the other miners verify its validity, i.e., the authenticity of the hash-chain. If it is true, a block with signed transaction and PW is added, and this miner will be rewarded for some Bitcoins. In this way, transaction data is permanently recorded in blocks, which is un-fungible. The complete process of a transaction from its formation to its completion is shown in Figure 3.



3 Blockchain privacy and zkSNARKs

3.1 Blockchain privacy

Bitcoin blockchain has its feature of transparency for the entire transaction progress. This is due to Bitcoin blockchain as well as most cryptocurrencies, including permissioned and consortium blockchains, publishing all transaction data to all miners or validators to enable consensus validation.

Due to the fact they can provide trusted consensus, computation, and immutable data between untrusted entities, blockchain technologies are gaining interest as a possible solution to many applications. A few examples include:

- decentralized finance (DeFi)
- decentralized identity
- supply chain management, healthcare
- private data management

However, those applications need privacy! In other words, for any application with sensitive data, blockchain privacy is needed. Thus, the problem is how to provide privacy for senders, transaction data, and receivers for applications with sensitive data? The current effective tool to provide blockchain privacy is to apply zero-knowledge proof systems in cryptography!

3.2 Zero-knowledge proofs and zkSNARK

Loosely speaking, zero-knowledge proofs are proofs that yield nothing beyond the validity of the assertion, as shown in Figure 4.

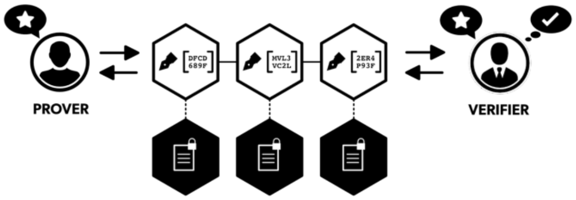


Fig. 4. A zero-knowledge proof system can convince a verifier of the validity of a statement without revealing it! For example, if we would like to prove any transaction in Zcash, it suffices to show the miners that SHA256 circuit y = SHA256(x) without giving the values in x = transaction (i.e., who is payer, payee, and the amount of Bitcoin is transferred). Zcash implemented Groth16 to achieve this goal.

Formally, a statement which can be proved without yielding anything beyond its validity is called a zero-knowledge (ZK) proof. In other words, ZK proofs are proofs that yield nothing (i.e., “no knowledge”) beyond the validity of the assertion.

A special case, called zero-knowledge -Succinct -Non-interactive -ARgument of -Knowledge, shortened as zkSNARK, is specially designed for arithmetic circuit proof systems.

4 Our Solutions

We have tackled the problem of blockchain privacy for different application scenarios. We have approached this challenge from different angles. First, in order to have an efficient zkSNARK system, we proposed a new proof system, called Polaris, and we found a new algorithm for speeding up the computation in pairing based zkSNARK systems. Secondly, we attempted to design a blockchain based health data system with an efficient consensus protocol.

4.1 Polaris: A New zkSNARK scheme

We have proposed Polaris [8], a new zkSNARK with the following properties.

(a) It is currently the most efficient zkSNARK system.

(b) It is post-quantum secure, compared with those based on hard problems: GGPR13, Groth16, Bullet proof (BCCGP2016), Marlin (CHMMVW20), etc.

(c) Polaris has improved verifier’s efficiency compared with the most promising ones: Stark (BBHR18), Ligero/Ligero++ (AHIV17, 2021), Aurora (BCRSVW19), Spartan$_{OR}$ (Setty20), etc.

4.2 A new algorithm for speeding up multi-scalar multiplication towards efficient zkSNARKs

We have proposed a new efficient algorithm [11] to speed-up the computation paring based zkSNARK schemes, e.g., Groth16, implemented in Zcash and adopted by Ethereum. Our new algorithm can gain from 3%-20% improvement over the current systems.

4.3 Blockchain healthcare model with an efficient consensus protocol

We have investigated a blockchain healthcare model and proposed a design with an efficient consensus protocol. A system model for blockchain healthcare is shown in Figure 5. In our initial work [7], we use a consortium chain to design a healthcare blockchain application with Hyperledger fabric as the underlying architecture, but it can be extended to other blockchains.

The model has the following five entities: patients, hospitals, research institutions, government medical institutions and insurance companies. Data Operation Module can be a proxy server used to deploy and coordinate the code between front-end calls and platform services, blockchain services are called through chain code services, and Decentralized Storage System is another chain composed of a group of nodes for storing health data with large data volumes such as a Filecoin System.

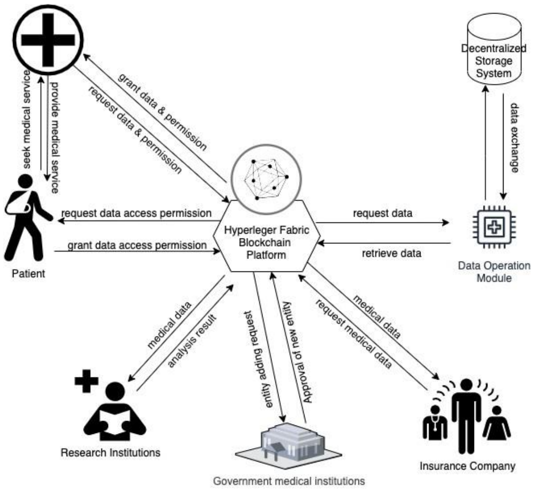


Fig. 5. Blockchain healthcare model [6] [7]

In our model, hospitals and medical institutions will be entities (i.e., miners/validators) in the system, while these institutions have to collaborate with insurance companies and scientific institutions, in addition it may be regulated by government sectors. Furthermore, each of these identities has the risk of being a Byzantine fault node, for example:

(1) The hospital node may be the main target of an attack since it hosts massive patient information and medical resources.
(2) The information recording process of the clinic node may be inaccurate and non-compliant, resulting in incorrect information.
(3) The insurance company may deny the facts and reject an insurance claim for its own benefit.
(4) An adversary may use the fake patient nodes for a Sybil attack.

We have investigated the design of a blockchain based health data management system with an efficient consensus protocol building on the practical Byzantine fault tolerance (PBFT) protocol [3]. We have proposed three different ways to speed-up the PBFT type of consensus protocols, i.e., optimization by role grouping, mary tree structured PBFT, and a new node rating mechanism.

5 Concluding remarks

We have briefly introduced our work:

– Polaris is a new tool to achieve practical blockchain privacy.
– A new algorithm is proposed to speed-up zero-knowledge proofs in Zcash and Ethereum.
– A design of Blockchain healthcare model with an efficient consensus protocol is presented.

5.1 Privacy is a need in all types of blockchains

None of the existing work on applying blockchain networks to health data, either in theory or in practice, has considered the privacy problem. This resembles the situation from 40 years ago for the password login system. The system stored all the passwords in plaintext, which caused the majority of break-ins. Tremendous efforts have been made in the last 20 years. The password protection schemes have been proposed and the standardizations have been conducted. So, currently, none of the secure systems stores passwords in plaintext.

5.2 Suitability

The utilization of smart contracts for transferring and sharing authentication and authorization credentials in health data systems is becoming increasingly apparent. The automation capabilities of smart contracts make them well-suited for these tasks. In the context of Person-Centred Data by Design [10], blockchain- enabled health record systems are more efficient compared to centralized systems. This efficiency stems from the unique characteristics of blockchain technology, such as enabling mutually distrusting entities to agree on a shared history and providing trusted consensus and immutable data among untrusted entities.

Blockchain technology proves suitable for sharing data in health data systems among multiple stakeholders, even in heterogeneous environments. To achieve this, we believe that a joint design of a consortium chain and a public chain would be well-suited. This combination allows for the benefits of both private and public blockchains, ensuring secure and efficient sharing of health data while maintaining privacy and accessibility as required.

There are many publications on blockchain enabled health data management. In academic research, each year's number of publications is more than double that of the previous year, since 2014. In practice, two examples are as follows:

The first scheme, MedRec [2], proposed by MIT group, is implemented on the public chain, Ethereum (v1. 2015, v2. 2020).

The first trial system in Canada is done by Canada's University Health Network (UHN), and it is built on the Hyperledger Fabric blockchain network (2019), permissioned blockchain.

5.3 Barriers to implement blockchain privacy in health data

ZK proofs were proposed only a few years later than public-key cryptography, but they remained primarily of theoretical interest in the field of cryptography for over 40 years due to their computational complexity. However, with the emergence of blockchain networks, people recognized the privacy challenges and realized that ZK proof systems could effectively address privacy concerns within blockchain systems. Nevertheless, this effectiveness comes at a cost.

In the past three years, there has been rapid development in implementing zkSNARK systems for blockchain privacy to meet the growing demand. The transfer or sharing of authentication and authorization for health data can be facilitated through smart contracts, streamlining the process once the permissions and rights are defined within the contract.

The barrier to the adoption for heath data is the efficiency of ZK proofs; i.e., how to efficiently implement and optimize those proofs at hardware, software, and protocol layers (it is much more complicated compared with a simple encryption)!

References

1. Mohiuddin Ahmed and Al-Sakib Khan Pathan. Blockchain: Can it be trusted? Computer, 53(4):31–35, 2020.

2. Asaph Azaria, Ariel Ekblaw, Thiago Vieira, and Andrew Lippman. Medrec: Using blockchain for medical data access and permission management. In 2016 2nd international conference on open and big data (OBD), pages 25–30. IEEE, 2016.

3. Miguel Castro and Barbara Liskov. Practical byzantine fault tolerance and proactive recovery. ACM Transactions on Computer Systems (TOCS), 20(4):398–461, 2002.

4. Randall D Cebul, James B Rebitzer, Lowell J Taylor, and Mark E Votruba. Organizational fragmentation and care quality in the us healthcare system. Journal of Economic Perspectives, 22(4):93–113, 2008.

5. Lidong Chen and Guang Gong. Communication System Security. CRC, 2012.

6. Aiden Feng. A design of electronic medical record system based on permissioned blockchain. Master Thesis, University of Waterloo, Jan. 2022.

7. Aiden Feng and Guang Gong. A blockchain based health data system with efficient mary tree consensus protocols. Internal Report, University of Waterloo, May 2022.

8. Shihui Fu and Guang Gong. Polaris: Transparent succinct zero-knowledge arguments for R1CS with efficient verifier. the Proceedings on Privacy Enhancing Technologies, 2022(1):544 – 564, Jan. 2022.

9. Guang Gong. Lecture notes on ECE 628 computer network security. Lecture Notes, University of Waterloo, Jan. 2023.

10. Expert Advisory Group. Pan-Canadian health data strategy: Toward a world-class health data system. Final report, May 2022.

11. Guiwen Luo, Shihui Fu, and Guang Gong. Speeding up multi-scalar multiplication over fixed points towards efficient zkSNARKs. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2023(2):358–380, Mar. 2023.

12. Rachel Pryor, Connie Atkinson, Kaila Cooper, Michelle Doll, Emily Godbout, Michael P Stevens, and Gonzalo Bearman. The electronic medical record and covid-19: Is it up to the challenge? American journal of infection control, 48(8):966, 2020.

Modernizing Canada Health Data Systems to Improve Capacity for Evidence-informed Decision-Making: Lessons from the Long-term Care Pandemic

John P Hirdes, PhD FCAHS,
University Professor, School of Public Health Sciences University of Waterloo &
Fellow, Balsillie School of International Affairs

Change in health care

According to Heraclitus, the only constant in life is change[1], but this phrase applies equally well to health care in Canada. Some changes occur in a long-term predictable manner and others are more abrupt and difficult to anticipate. Decision makers at all levels of health care must be prepared to respond to both types of changes in a proactive, timely, and appropriate manner. Scientifically rigorous health information systems are essential tools to improve the effectiveness of the decision-making process. However, it is not enough to simply *gather* data. The management, mobilization, interpretation, dissemination, and application of data all affect the extent to which health system decisions translate to better population health outcomes.

Long term predictable societal changes can necessitate substantial changes in the organization and delivery of health care services. Their implications are often clear, but when changes occur over decades the demands pose more acute challenges that can delay responses to long term concerns.

Some examples of important long-term trends affecting the health and health care of Canadians include:

1) **Population aging** - the increased percentage of persons age 65 years or more is an accelerating global phenomenon[2]. Population aging has itself accounted for only modest increases in growth of health expenditures compared with technological change, intensification of clinical interventions, and wage inflation[3]. However, older persons have more complex health needs than younger individuals and have less certain health outcomes. The greater concentration of older persons with complex, resource intensive needs in health care settings has reduced the margins for “error” in health policy and service delivery;
2) **Female labour force participation** - for more than a century, the increased rates and duration of participation by women in the paid labour force has been a well-known transition[4]. However, changes in childcare policies and services have lagged. This transition is also likely to affect the capacity and nature of informal caregiving responsibilities that women can take on due to competing time demands. Given that over 70% of the support received by home care clients comes from informal sources[5,6], a reduced ability of caregivers to provide instrumental support may make current home care solutions inviable;
3) **Urbanization** - in most countries, there has been a long-term trend of migration from rural to urban settings. This has been particularly true for younger individuals resulting in a greater concentration of older persons who have remained in rural settings. The resulting localized reductions in available informal supports and increased old age dependency ratios can place increased pressures on formal care systems to provide care to a widely dispersed population without the economies of scale needed to provide a full range of health services[7,8];
4) **Geographic mobility** - a variety of factors including job opportunities and cost of housing continue to drive increased rates of geographic mobility, particularly among younger individuals. This change can also reduce informal supports available to older persons and persons with disabilities or complex medical needs[9].

These four abovementioned long-term changes require a thoughtful approach to how health systems are managed and organized. The future of the health care system is inextricably linked to the policy and service delivery choices made to respond to the needs of older persons[10]. We cannot assume that models of care configured for social structures and processes 50 years ago will be viable for the coming half century. One of the sectors that urgently requires effective and innovative new approaches to service delivery is long-term care.

It would be incorrect to claim that the COVID-19 pandemic beginning in 2020 was an unexpected event[11]. Infectious disease epidemiologists have long warned of the potential for global pandemics. Perhaps the relatively successful management of SARS and other earlier outbreaks led to some complacency or overconfidence at all levels of society. Nonetheless, COVID-19 resulted in a dramatic rise in GDP expenditures in health care, massive numbers of death among persons in long-term care, social and economic upheaval, and an unprecedented crisis in health human resources. This "unexpected" event had serious adverse consequences in virtually every aspect of health care in Canada. It magnified the existing vulnerabilities in the health system and delayed access to health services for other urgent health needs (e.g., surgery).

Despite having highly sophisticated clinical information systems in place, the long-term care sector experienced among the most severe consequences of the first two years of the COVID-19 pandemic.

The Long-term Care Pandemic

The earliest evidence about at-risk populations at the outset of the COVID-19 pandemic came mostly from acute hospitals in China [12-16]. Advanced age and multimorbidity were identified as major risk factors for COVID-19 deaths in those reports[17], but the only evidence available for long-term care at the start of the pandemic came from small case studies in a limited number of US homes[18-22]. However, preliminary Canadian analyses that considered age, multimorbidity, and pneumonia as a proxy for COVID-19 using data from the Canadian Institute for Health Information's (CIHI) Continuing Care Reporting System suggested that deaths due to COVID-19 in long-term care could be substantial [23]. The almost 15,000 LTC deaths that occurred in the first two waves [24,25] was 10 times greater than pessimistic predictions put forth at the start of the pandemic. Outbreaks in LTC were shown to be a function of a diverse range of factors, including pre-pandemic quality of care, staffing levels, use of agency staff, facility and chain size, infrequent presence of medical directors, and rates of COVID-19 in local communities[26-29].

All of this transpired despite the presence of an advance health data system[30] providing extensive information about the health needs of millions of individuals in over 1,200 long-term care homes, covering most regions of Canada (www.yourhealthsystem.cihi.ca)[31]. By January 2020, 8 provinces/territories had already implemented interRAI assessment systems[32-37] that provided person level clinical information on over 300 health conditions (e.g., cognition, functional status, frailty, mental health, continence, sensory impairments, medication use, treatment, and interventions) with longitudinal records based on updated assessments every three months. The data from these assessments have been shown to be robust, reliable, and valid [38-43]; they are compatible with other interRAI systems used in other health sectors (e.g., home care) [44,45]; there are sophisticated information technology systems in place to support their clinical, managerial, and policy-level uses; and they are supported by Pan-Canadian public reporting systems managed by CIHI to report on their quality of care[46,47]. To date, over 20 million interRAI assessments had already been completed on over 6 million unique Canadians in LTC, complex continuing care hospitals, home care, mental health facilities, palliative care, and acute hospitals, with some implementations covering full clinical populations for over 20 years.

Given that LTC is a health data "rich" environment, why then was the impact of COVID-19 so severe in this sector?

Weaknesses in LTC Information Systems

Despite having a world-class data system in place for health assessment of LTC residents, there are numerous important gaps and limitations that affected Canada's decision-making abilities for this sector. A significant number of these challenges remain unaddressed. The following examples highlight some of the important limitations of existing LTC information systems.

Public health reporting in LTC facilities – Perhaps, the most severe informational problem in LTC was the inadequacy of the system for reporting outbreaks and COVID-19 related deaths in that sector. In the earliest stages of the pandemic, the most complete source of information about LTC outbreaks came from media sources who polled individual homes about their outbreak status. Since then, several provinces have compiled data sets to allow for retrospective analyses of home outbreaks, but a Pan-Canadian outbreak reporting system is still not in place for this sector and such data are not linked routinely to clinical information in CIHI's reporting system for LTC.

Quality of life and resident experience – There is ample evidence that COVID-19 caused excess mortality in LTC homes. The effects on health outcomes such as cognitive decline, reduced functional status, and mood disturbances have also been demonstrated [48-50]; however, these effects have been more modest than would be anticipated, based on popular media reports. In addition, recent evidence shows that the primary consequence for worsened quality of care has been in the form of increased antipsychotic use after the pandemic [47]. It has been argued that the quality-of-life consequences of the lock down strategies used by LTC homes to contain outbreaks could be pervasive and severe [51,52]. However, there is no Pan-Canadian standard in place for a common measure of quality of life in LTC[53]. Results from one multifacility study run counter to the assumption that quality of life was severely compromised by these measures [50]; however, that study cannot be assumed to be generalizable to all homes nationally.

Process measures - The interRAI assessment systems used by most provinces and territories in Canada includes a host of clinical outcome measures that provide a comprehensive picture of the health and well-being of residents in LTC and home care settings. However, these systems provide incomplete information about processes of care, including vaccination data and medication use (other than general categories of psychoactive medications).

Structural measures – Structural indicators relate to the organization and resources allocated to health services. The recently updated Nursing and Residential Care Facility Survey (NRCFS) will fill multiple gaps in structural indicators, but such data were not readily available during the first two years of the pandemic. They are also not routinely linked to the clinical outcome data managed by CIHI. As part of the Ontario Long-term Care COVID Commission review of the underlying factors affecting LTC homes during the pandemic, such data were gathered through surveys of individual homes, but they were not retained for uses beyond the report of the Commission.

Data sharing across health sectors – Despite the implementation of interRAI assessment systems across multiple care settings to form the basis for an integrated health information system with longitudinal interoperable records[44], the actual clinical use of these data across care settings remains relatively limited. These assessments are not routinely shared across the circle of care, and they were generally not used to manage care with transitions across health settings. This has been a consequence of inadequate cross-sector collaboration rather than a problem of compatibility or relevance of the clinical information.

A broader view of challenges in Canadian health data systems

Unmeasured populations in need

The abovementioned challenges affected the relatively robust information systems in place for LTC and home care in Canada. However, the situation in other sectors is much more problematic. There are almost complete gaps in health information for other vulnerable populations with complex health needs. For example, there are limited to no data for older adults living in retirement homes or assisted living settings, and clinical records for acute care and primary care are relatively sparse and rudimentary, with measures mostly limited to diagnosis, treatments, and procedures, as well as medication use. Few provinces have adopted common standards for inpatient psychiatry beyond the Discharge Abstract Database (DAD) and no national standards are in place for community mental health services. Similar, there is no pan-Canadian standard for clinical assessment of children or youth with mental health needs, intellectual or physical disabilities, or complex medical needs. Each of those settings has *less* clinically actionable information than LTC does, yet they serve highly vulnerable populations with complex needs.

Inconsistent data standards

Enacting national data standards has been remarkably difficult in most countries. The DAD is a widely used administrative database for acute care settings, but it is limited in its clinical content to mainly information about diagnosis and interventions. After over 25 years of effort, Canada has *almost* achieved a common data standard for LTC. The only outliers are Quebec and Nunavut. A similar pattern is evident for home care services across the country. In other settings, a mixture of measures is used; some of which are scientifically based, others are "home grown" non-standardized measures.

An important question is whether diverse measures can somehow be harmonized to be used to make national inferences to inform health decisions[54,55]. When it comes to health is a "rose still a rose by any other name"? If two different measures are used for the same underlying construct, can we still make conclusions about health needs, effectiveness of resource allocation, or quality of care between jurisdictions using different standards?

The answer is *probably not,* for two reasons. First, although it may be able to make some crosswalks between different measures of the same latent construct, these will almost always involve conversion of more granular to simplified, rudimentary measures. One might convert two different scales of functional loss to become cruder binary measures of *any* versus *no* impairment; however, not all types of impairment are equivalent. Measures involving different time frames, performance, or capacity are not comparable, even in their simplified forms. Historical measurement differences may also be problematic within the same region. This may affect estimates of prevalence and incidence of health conditions, and comparisons between pre-pandemic and post-pandemic performance may be difficult.

Aside from the scientific challenges of making inferences from non-equivalent measures, the social and political challenges may be greater. A first line of defense used by agencies with poor performance in quality reports is that health outcomes were measured in a different way. Standards can negate that argument, but use of inconsistent measures gives it life. The second line of defense is "my patients are different." The solution of using standardized data for clinical covariates that serve as risk adjusters for quality indicators [46,56] is unavailable when disparate measures are used. The consequence is that performance measures based on non-standardized data will almost never be accepted by organizations or governments with problematic results based on those data. Trust in the veracity of performance measures depends on the belief that everyone is measuring the same things in the same way.

The denominator problem

Information about the health of populations generally comes from two sources: sampled data and routinely collected data sources. Sampled data include surveys of the general population or specific subpopulations of interest. Non-response bias has become a severe problem for most survey research with low response rates posing a major threat to the representativeness and generalizability of findings. In addition, periodic sampling of health data with panel surveys may not yield information about populations *when* they are needed, even if they are representative at the time of data collection.

Routine data collection at the point of care can yield fully representative population level data for service recipients. For example, all residents admitted to LTC homes in most provinces and territories in Canada, as well as all long-stay home care clients in those regions, are assessed with interRAI systems. The attrition due to death or discharge *prior to* assessment is modest, making these data sets equivalent to population-level information. However, the definition of the population that has been measured can sometimes still be unclear. For example, there are dozens of different names in Canada for what would otherwise be called "nursing homes" in international usages [57]. What constitutes the "home care" population is even more unclear within and outside of Canada. Similar challenges would apply to mental health services, palliative care, and services for persons with disabilities. We can define populations based on service recipients, but not all persons with similar needs receive those services.

Inertia in the mobilization of evidence

When emergencies threaten population health, the rapid mobilization of health evidence is essential. Paper records retained in filing cabinets have no utility in a crisis. They cannot be accessed, so they cannot be reported. Non-standardized data may not be interpretable, and they may not be accepted by decision-makers or other stakeholders. Narrative notes in electronic health records have been shown to have limited sensitivity and validity as data sources for clinical problems[58].

A particularly important problem that affected the data rich LTC environment was the use of *batch submissions* for clinical records. In most regions of Canada, data were reported to CIHI as batch files on a quarterly basis. The next quarter was used to resolve submission errors or data quality problems, meaning that data about a given month of the pandemic may not have been available before the next wave of COVID-19 hit. Provinces like New Brunswick adopted CIHI's newer Integrated interRAI Reporting System (IRRS) which has *near real-time* functionality. This meant that New Brunswick had population level health data for LTC by mid-summer of 2020 whereas similar data for other provinces were only available after the start of Wave 2.

An additional challenge that hinders the mobilization evidence is Canada's limited supply of health data scientists. The volume of work needed to analyze all health data during the height of the COVID-19 pandemic simply exceeded the supply of people with sufficient expertise to analyze those data. This was exacerbated by delays in access to relevant data by people who had that expertise (e.g., for the reasons noted above). Moreover, the complexity of the analyses required to understand the many undulations and unexpected turns of the pandemic led to considerable risk aversion among some researchers who might otherwise have been engaged in this type of research.

Privacy concerns and limitations on data use

A major privacy concern with large complex health datasets is the reidentification of individuals through the examination or reporting of sets of personal attributes. This is particularly a risk when precise geospatial data are linked to person-level records and when individuals are part of small populations (e.g., gender identity, race/ethnicity). However, it is also the case that the experience of COVID-19 was highly variable between regions, and it had a greater impact on multiple marginalized or vulnerable populations. The problem of small subsample sizes meant that many affected subgroups were either excluded from analyses or collapsed into ambiguous residual groups.

Linked data sets provide powerful, cost-effective means to rapidly examine broad ranges of predictors and consequences of events like the COVID-19 pandemic. They can also be used to solve methodological challenges (e.g., tracking use of multiple health care services, mortality). Linked data can provide decision makers with system-level views rather than sector specific views only. However, such linked data sets can also magnify concerns related to reidentification risks because the sheer size and diversity of measure can increase dramatically for each person.

While there are mechanisms to allow for examination of rare populations or to use linked data sets, the time lags and availability of these data often mean that their analysis provides evidence *after the fact*, rather than providing contemporaneous evidence to decision-makers.

Data driven vs evidence-informed decisions

Despite the appeal of the phrase "data driven decision-making" it should be recognized that data *do not* speak for themselves. The questions that policy makers ask can be formulated into research questions that can be evaluated through careful, systematic examination of high-quality data. However, the formulation of questions, interpretation of the results, and the transformation of evidence into policy and practice are ultimately value-based propositions. The choices we face are often ambiguous and the actions we take in response to a given problem usually involve opportunity costs related to other problems. For example, at the start of the pandemic the decision was made in some provinces to protect the supply of personal protective equipment for hospitals at the expense of providing those to long-term care settings. This came despite evidence that long-term care homes cared for the highest risk populations according to evidence provided by WHO at that time.

To make full use of health data to inform clinical, service delivery, and policy decisions we must apply value frameworks that are explicitly defined through social and political processes. As was clearly stated in the Romanow report on the future of health care in Canada[59], health data systems should help us understand whether our choices support progress toward health care goals valued by Canadians. That principle can be readily applied to long-term predictable changes in health care. When the time frame for change is decades, there should be ample opportunity for meaningful public input and dialogue about health system responses. On the other hand, the intense pace of change and unpredictable nature of the COVID-19 pandemic left little time for public discourse about policy choices in response to each new phase of the pandemic.

The promise and challenges of artificial intelligence in health care

Machine learning and artificial intelligence (AI) systems have already begun to transform the analysis of complex health data sets. The capacity to analyze large longitudinal data sets with numerous covariates can easily outstrip the abilities of most seasoned data scientists using conventional techniques. However, the interpretability of those analyses is sometimes unclear and the applicability of results from one dataset to another (or another population) is not always certain. "Black box solutions" that are not interpretable may not be replicated with other settings or populations. More importantly, they may not provide scientific evidence that helps us understand underlying truths in the relationships between variables of interest. That may reduce acceptance of innovations by clinicians or policy makers. It may also hinder rather than help identifying effective treatments if underlying mechanisms are not understood.

A newer consideration is the use of generative AI systems to automate clinical, managerial, or policy decisions. Without a doubt, generative AI systems can provide insights yielding better information to reduce errors in the health system that harm health and well-being. However, such systems could also be used to enhance non-health priorities (e.g., profitability) that could come at the expense of health. Such systems could also be used inappropriately to further other interests (e.g., gaming of case-mix systems). Moreover, generative AI systems can be subject to systemic biases because they draw from widely available information that is itself biased. For example, frail older persons are typically excluded from clinical trials making such evidence inapplicable to them[60].

The ultimate implication of these risks is that health system decision-makers must become competent curators of evidence from machine learning and AI solutions[61]. They must have the capacity to engage patients and other stakeholders in discussions about their strengths, preferences, and needs to provide a context to which evidence from these solutions can be applied. They must also gain an understanding of the underlying scientific basis for the evidence emerging from those solutions. AI systems will certainly become valuable sources of evidence for health decisions. However, the decisions themselves must remain a product of a collaborative dialogue between patients, their support network, and the team of clinicians with a full understanding of the implications of alternative care choices and how those choices relate to their personal priorities.

AI systems have the potential to free clinicians' time to allow for greater interpersonal engagement in health decision making. On the other hand, there will be temptations to gain efficiency in health care with automated decision making that reduces rather than enhances the humanity of health care decision making.

Specific practical recommendations

1) **Build trust and confidence** – as problems become more challenging and health data solutions more complex, the challenge for the general population may be that difficulty in comprehending evidence may lead to recommendations not being accepted. Therefore, it is essential that education and public engagement strategies be combined with transparency in reporting to ensure public trust in health information systems. Canadians must have a sense that their values underpin all levels of health decision-making.
2) **Fill the gaps** – massive efforts have been undertaken to provide comprehensive data for some aspects of some health sectors (e.g., long term care, home care). We must continue efforts to establish standardized, interoperable, valid, reliable, relevant, and actionable health data for missing populations (e.g., persons with mental health needs) and health sectors (e.g., retirement homes, assisted living).
3) **Move to real time solutions** – the pace of change in health care is unlikely to slow in the foreseeable future. Therefore, rapid access to near real time information will continue to grow in importance. For long term care and home care this practically means that current legacy systems (e.g., CIHI's HCRS and CCRS reporting systems) must transition to near real time solutions. New systems implemented to fill system gaps should be designed from the outset to have near real time functionality.
4) **Strengthening Canada's capacity in health data science** – the need for persons with sophisticated analytic skills will *increase* at all levels of the health system because of advances in health data systems, machine learning, and artificial intelligence. Moreover, all decision makers in every facet of the health system must increase their capacity to evaluate, interpret, communicate, and act on evidence from diverse sources.
5) **Develop professional standards and accreditation processes for health data scientists** – to ensure the trust of Canadians and to enhance the use of data while reducing reidentification risk for small but important populations, mechanisms should be put in place to apply regulatory professional standards and accreditation to scientists and data analysts who work with health data.
6) **Use synthetic data where they are applicable** – synthetic data may not be usable for person-level clinical decision-making, organizational performance measurement, resource allocation, or policy development. However, they can help to accelerate the modernization of Canada's health data systems in other ways, including education and training, code verification, exploratory analysis, hypothesis generation, and preliminary algorithm development.

Concluding Comments

Canada has made remarkable strides in the breadth, quality, relevance, and utility of health information systems for every health care sector in the last 40 years. For example, as limited as the funding allocated to long-term care has been, Canada has a highly sophisticated, “world class” system to measure and publicly report on the performance of that sector. Other sectors like home care bear similar near-term promise. That said, improving the quality of health information systems should be a continuous priority as part of on-going efforts to modernize the Canadian health data system. Some components are already in place to realize the objectives of the Pan Canadian Health Data Strategy to allow all Canadians to “benefit from a fully integrated and continuously optimized health system” with “person-centric health information that ensures health data follows individuals across points of care”[62]. However, the practical steps described here should be undertaken to enhance our capacity to respond effectively to both the predictable long-term changes and the unexpected rapid changes in health and health care that will inevitably be faced by future decision-makers.

References

1. Müller-Merbach H. Heraclitus: philosophy of change, a challenge for knowledge management? Knowledge Management Research & Practice 2006;4:170-171.
2. He W, Goodkind D, Kowal PR. An aging world: 2015. United States Census Bureau Washington, DC; 2016.
3. Kinsella KG, Velkoff VA. An aging world: 2001: Bureau of Census, 2001.
4. Montcho G, Carrière Y, Mérette M. Population Aging and Work Life Duration in Canada. Canadian Public Policy 2023;49(S1):32-47.
5. Health Council of C. Seniors in need, caregivers in distress: What are the home care priorities for seniors in Canada? : Health Council of Canada Toronto; 2012.
6. Betini RSD, Hirdes JP, Lero DS, Cadell S, Poss J, Heckman G. A longitudinal study looking at and beyond care recipient health as a predictor of long term care home admission. BMC Health Serv Res 2017;17(1):709. DOI: 10.1186/s12913-017-2671-8.
7. Wiese LAK, Gibson A, Guest MA, et al. Global rural health disparities in Alzheimer's disease and related dementias: State of the science. Alzheimer's & Dementia 2023.
8. Haan M, Jin H, Paul T. The Geographical Mobility of Journeypersons in Canada: Evidence from Administrative Data. Population Research and Policy Review 2023;42(2):19.
9. Li L, Wister A. Geographic distance and social isolation among family care-givers providing care to older adults in Canada. Ageing & Society 2023;43(2):298-323.
10. MacDonald B-JW, Michael; Hirdes, John P. The Future Co$t of Long-Term Care in Canada. Toronto, Canada: Ryerson University, 2019. (https://static1.squarespace.com/static/5c2fa7b03917eed9b5a436d8/t/5dbadf6ce6598c340ee6978f/1572527988847/TheFutureCostofLong-TermCareinCanada.pdf).
11. Morse SS, Mazet JAK, Woolhouse M, et al. Prediction and prevention of the next pandemic zoonosis. The Lancet 2012;380(9857):1956-1965.
12. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med 2020;382(8):727-733. DOI: 10.1056/NEJMoa2001017.
13. Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med 2020. DOI: 10.1056/NEJMoa2002032.
14. Xie J, Tong Z, Guan X, Du B, Qiu H. Clinical Characteristics of Patients Who Died of Coronavirus Disease 2019 in China. JAMA Netw Open 2020;3(4):e205619. DOI: 10.1001/jamanetworkopen.2020.5619.
15. Liu K, Chen Y, Lin R, Han K. Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. J Infect 2020. DOI: 10.1016/j.jinf.2020.03.005.
16. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395(10223):497-506. DOI: 10.1016/S0140-6736(20)30183-5.
17. World Health O, World Health O. Report of the who-china joint mission on coronavirus disease 2019 (covid-19). 2020.
18. Crimmins EM. Age-related vulnerability to coronavirus disease 2019 (COVID-19): biological, contextual, and policy-related factors. Public Policy & Aging Report 2020;30(4):142-146.
19. Kimball A, Hatfield KM, Arons M, et al. Asymptomatic and Presymptomatic SARS-CoV-2 Infections in Residents of a Long-Term Care Skilled Nursing Facility - King County, Washington, March 2020. MMWR Morb Mortal Wkly Rep 2020;69(13):377-381. DOI: 10.15585/mmwr.mm6913e1.
20. Gardner W, States D, Bagley N. The coronavirus and the risks to the elderly in long-term care. Journal of Aging & Social Policy 2020;32(4-5):310-315.
21. McMichael TM, Clark S, Pogosjans S, et al. COVID-19 in a Long-Term Care Facility - King County, Washington, February 27-March 9, 2020. MMWR Morb Mortal Wkly Rep 2020;69(12):339-342. DOI: 10.15585/mmwr.mm6912e1.
22. Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 - Studies Needed. N Engl J Med 2020;382(13):1194-1196. DOI: 10.1056/NEJMp2002125.
23. Hirdes JP, Declercq A, Finne-Soveri H, et al. The Long-term Care Pandemic: International Perspectives on COVID-19 and the Future of Nursing Homes. Balsillie Papers 2020;2(4).
24. CIHI. Pandemic Experience in the Long-Term Care Sector: How Does Canada Compare With Other Countries? Ottawa: Canadian Insititute for Health Information, 2020 2020. (https://www.cihi.ca/sites/default/files/document/covid-19-rapid-response-long-term-care-snapshot-en.pdf).
25. CIHI. COVID-19's impact on long-term care. Canadian Institute for Health Information. (https://www.cihi.ca/en/covid-19-resources/impact-of-covid-19-on-canadas-health-care-systems/long-term-care).
26. Stall NM, Jones A, Brown KA, Rochon PA, Costa AP. For-profit long-term care homes and the risk of COVID-19 outbreaks and resident deaths. Cmaj 2020;192(33):E946-E955.
27. Long-Term Care COVID-19 Commission Meeting: Canadian Institute for Health Information. Long Term Care COVID-19 Commission. Toronto, Canada: Neesons; 2021:105.
28. Abrams HR, Loomer L, Gandhi A, Grabowski DC. Characteristics of US nursing homes with COVID-19 cases. Journal of the American Geriatrics Society 2020;68(8):1653-1656.
29. Turcotte LuD, Natalie. Factors associated with COVID-19 outbreaks, transmission and poor outcomes in Ontario's LTC Homes. CIHI. (http://ltccommission-commissionsld.ca/presentations/pdf/Canadian_Institute_for_Health_Information_March_3_2021.pdf).
30. Hirdes JP, Mitchell L, Maxwell CJ, White N. Beyond the 'iron lungs of gerontology': using evidence to shape the future of nursing homes in Canada. Can J Aging 2011;30(3):371-90. DOI: 10.1017/S0714980811000304.

31. CIHI. Your Health System: In Depth - Technical Notes for Contextual Measures. Ottawa: Canadian Insitute for Health Information, 2023 2023. (https://www.cihi.ca/sites/default/files/document/yhs-technical-notes-contextual-measures-meth-notes-en.pdf?_gl=1*ss6s6q*_ga*MTY4NTc0MTIwOC4xNjczMDIwMzY5*_ga_44X3CK377B*MTY5NDk0NzE1Mi4zMC4xLjE2OTQ5NDc2NDMuMC4wLjA.&_ga=2.31586109.2095070510.1694947152-1685741208.1673020369).
32. Carpenter GI, Hirdes JP. Using interRAI assessment systems to measure and maintain quality of long-term care. A good life in old age? Monitoring and improving quality in long-term care. Paris: OECD Publishing; 2013:2017.
33. Steel K, Jonsson PV, Dupasquier JN, et al. Systems of care for frail older persons. InterRAI. Trans Am Clin Climatol Assoc 1999;110:30-5; discussion 35-7. (https://www.ncbi.nlm.nih.gov/pubmed/10344004 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2194306/pdf/tacca00006-0086.pdf).
34. Hirdes JP, Fries BE, Morris JN, et al. Integrated health information systems based on the RAI/MDS series of instruments. Healthcare Management Forum 1999;12(4):30-40.
35. Hirdes JP. Addressing the health needs of frail elderly people: Ontario's experience with an integrated health information system. Oxford University Press; 2006.
36. Bernabei R, Landi F, Onder G, Liperoti R, Gambassi G. Second and third generation assessment instruments: the birth of standardization in geriatric care. J Gerontol A Biol Sci Med Sci 2008;63(3):308-13. DOI: 10.1093/gerona/63.3.308.
37. Heckman GA, Gray LC, Hirdes JP. Addressing health care needs for frail seniors in Canada: the role of interRAI instruments. 2013.
38. Foebel AD, Hirdes JP, Heckman GA, et al. Diagnostic data for neurological conditions in interRAI assessments in home care, nursing home and mental health care settings: a validity study. BMC Health Serv Res 2013;13:457. DOI: 10.1186/1472-6963-13-457.
39. Poss JW, Jutan NM, Hirdes JP, et al. A review of evidence on the reliability and validity of Minimum Data Set data. Healthc Manage Forum 2008;21(1):33-9. DOI: 10.1016/S0840-4704(10)60127-5.
40. Hirdes JP, Poss JW, Caldarelli H, et al. An evaluation of data quality in Canada's Continuing Care Reporting System (CCRS): secondary analyses of Ontario data submitted between 1996 and 2011. BMC Med Inform Decis Mak 2013;13:27. DOI: 10.1186/1472-6947-13-27.
41. Hogeveen SE, Chen J, Hirdes JP. Evaluation of data quality of interRAI assessments in home and community care. BMC Med Inform Decis Mak 2017;17(1):150. DOI: 10.1186/s12911-017-0547-9.
42. Hirdes JP, Ljunggren G, Morris JN, et al. Reliability of the interRAI suite of assessment instruments: a 12-country study of an integrated health information system. BMC Health Serv Res 2008;8:277. DOI: 10.1186/1472-6963-8-277.
43. Turcotte LA, Poss J, Fries B, Hirdes JP. An Overview of International Staff Time Measurement Validation Studies of the RUG-III Case-mix System. Health Serv Insights 2019;12:1178632919827926. DOI: 10.1177/1178632919827926.
44. Gray LC, Berg K, Fries BE, et al. Sharing clinical information across care settings: the birth of an integrated assessment system. BMC Health Serv Res 2009;9:71. DOI: 10.1186/1472-6963-9-71.
45. Morris JN, Fries BE, Steel K, et al. Comprehensive clinical assessment in community setting: applicability of the MDS-HC. J Am Geriatr Soc 1997;45(8):1017-24. DOI: 10.1111/j.1532-5415.1997.tb02975.x.
46. Jones RN, Hirdes JP, Poss JW, et al. Adjustment of nursing home quality indicators. BMC Health Serv Res 2010;10:96. DOI: 10.1186/1472-6963-10-96.
47. Turcotte LA, McArthur C, Poss JW, et al. Long-Term Care Resident Health and Quality of Care During the COVID-19 Pandemic: A Synthesis Analysis of Canadian Institute for Health Information Data Tables. Health Services Insights 2023;16:11786329231174745.
48. Webber C, Myran DT, Milani C, et al. Cognitive decline in long-term care residents before and during the COVID-19 pandemic in Ontario, Canada. Jama 2022;328(14):1456-1458.
49. McArthur C, Saari M, Heckman GA, et al. Evaluating the effect of COVID-19 pandemic lockdown on long-term care residents' mental health: a data-driven approach in New Brunswick. Journal of the American Medical Directors Association 2021;22(1):187-192.
50. McArthur C, T. Mulla R, Turcotte LA, Chu JC-Y, Jantzi M, Hirdes JP. Mental Health and Quality of Life in Long-Term Care During the Pandemic. In: Horney JA, ed. COVID-19, Frontline Responders and Mental Health: A Playbook for Delivering Resilient Public Health Systems Post-Pandemic: Emerald Publishing Limited; 2023:97-117.
51. Barnett ML, Waken RJ, Zheng J, et al. Changes in health and quality of life in US skilled nursing facilities by COVID-19 exposure status in 2020. JAMA 2022;328(10):941-950.
52. Anderson DC, Grey T, Kennelly S, O'Neill D. Nursing home design and COVID-19: balancing infection control, quality of life, and resilience. Journal of the American Medical Directors Association 2020;21(11):1519-1524.
53. HSO. HSO Long-Term Care Services: National Standard of Canada. Ottawa: Standards Council of Canada, 2023. (https://longtermcarestandards.ca/).
54. Carpenter GI, Bernabei R, Hirdes JP, Mor V, Steel K. Building evidence on chronic disease in old age. Standardised assessments and databases offer one way of building the evidence. BMJ 2000;320(7234):528-9. DOI: 10.1136/bmj.320.7234.528.
55. Heckman GA, Turcotte L, Hirdes JP. The crying need for validated consensus frailty measurement standards: will the real frailty please stand up? Canadian Journal of Cardiology 2021;37(11):1702-1704.
56. Frijters DH, van der Roest HG, Carpenter IG, et al. The calculation of quality indicators for long term care facilities in 8 countries (SHELTER project). BMC Health Serv Res 2013;13:138. DOI: 10.1186/1472-6963-13-138.
57. Hirdes JP, Berg K. Care of the Elderly in the Canadian Health Care System. In: Carpenter I, Challis D, Hirdes JP, Bernabei R, Ljunggren G, eds. Care of Older People in North America, Europe and Japan. London: Farrand Press; 1999:17-32.

58. Jonsson PV, Finne-Soveri H, Jensdottir AB, et al. Co-morbidity and functional limitation in older patients underreported in medical records in Nordic Acute Care Hospitals when compared with the MDS-AC instrument. Age Ageing 2006;35(4):434-8. DOI: 10.1093/ageing/afj060.
59. Romanow RJ, Qc C. Building on values: the future of health care in Canada. 2002.
60. Evans JG. Evidence-based and evidence-biased medicine. Age and Ageing 1995;24(6):461-464.
61. Heckman GA, Hirdes JP, McKelvie RS. The role of physicians in the era of big data. Canadian Journal of Cardiology 2020;36(1):19-21.
62. Canada PHAo. Moving Forward on a Pan-Canadian Health Data Strategy. Government of Canada. (https://www.canada.ca/en/public-health/programs/pan-canadian-health-data-strategy.html).

The Dynamic Shift and Transformative Impacts of Big Data on the Canadian Health System Landscape

Pedro Miranda, PhD(c), MSc
Plinio Morita, PEng, PhD
Arlene Oetomo, PhD(c), BSc
Shahan Salim, PhD(c), MSc

Ubiquitous Health Technology Lab (UbiLab)
School of Public Health Sciences
University of Waterloo

Abstract

Health systems in high income countries, including Canada, are rapidly evolving, and transforming. However, this progress has been accompanied by a growing burden on the system and funding cuts, further exacerbated by the COVID-19 pandemic revealing weaknesses. Moreover, the system is currently grappling with addressing systemic inequalities and inequities. A notable concern is the system's inherent reactivity and slowness, primarily focusing on treatment rather than a preventative approach. In this context, data is pivotal in modern health systems, aiding medical professionals, public health officials, and healthcare administrators in informed decision-making, effective planning, and resources allocation. With the advent of advanced data collection technologies, data availability has reached unprecedented levels, making it imperative to integrate these datasets into healthcare and public health systems to maximize their potential. Failure to do so would be a lost opportunity. This chapter delves into the dynamic realm of Big Data, advocating its incorporation into the healthcare and public health system and exploring how this can be achieved. The exploration encompasses novel data sources applicable to public health systems and examines their interaction within healthcare and public health contexts. Ultimately, this chapter paints a comprehensive picture of the future of a holistic health system, where the integration of diverse data pipelines emerges as a key determinant of progress and success.

Current State of Canadian Health System

The state of health systems worldwide and rise of neoliberal policies presents a critical epidemic that necessitates a paradigm shift in our approach. A crucial aspect of medical decision-making revolves around patient data, pivotal in determining diagnoses and prognoses. However, patients often navigate through a convoluted process of seeking opinions and feedback from multiple specialists, a journey that may take days or months to reach a diagnosis. Each expert provides their piece of the puzzle, but the holistic picture only emerges once the data is transmitted; often through outdated technological means such as fax or email threads. The siloed approach within the health system results in sluggish performance and fragmentation[1]. The resulting delays can be detrimental to a patient's wellbeing, as critical changes in their condition during this time may go unnoticed and untreated, slipping through the cracks. Consequently, the health system is in a perpetual state of catch-up, prioritizing treatments over preventative measures, further compounding the issue[2]. Moreover, the overburdened state of the health system only serves to intensify the impact of such delays. For instance, postponed surgeries, brought about by the system's backlog, contribute to further deterioration in patients' health, compounding the adverse effects on their well-being.

This leads to the second pressing concern—the health system reaching a tipping point. As the health system grapples with the compounding effects of COVID-19, the strain on its resources and personnel has been unprecedented. Hospital workers on the front lines of the battle against the pandemic have borne the brunt of the crisis, resulting in burnout and fatigue among healthcare professionals[3]. The record number of professionals exiting the healthcare workforce or facing physical and mental health challenges represents a highly concerning trend that demands immediate attention, as Emergency Rooms around the country cut hours due to staff shortages[4,5].

The strain caused by the current healthcare challenges extends far beyond the burden placed on the healthcare staff; it also profoundly impacts the entire health system and has significant economic implications. Post COVID-19 healthcare spending in Canada ballooned to $331 billion in 2022 and is expected to continue growing, exacerbated by an aging population and overall population growth[6]. However, the health system's ability to improve health and wellbeing remains limited, with significant gaps failing to address critical social determinants of health such as poverty, low income, and low education, coupled with soaring inflation[7–9]. Gaps in coverage for dental care, eye care, and pharmacy care impact the most marginalized disproportionately[10,11]. Additionally, the pandemic has exposed and amplified the pre-existing inequalities entrenched within the health system. Certain vulnerable populations, including minority groups, recent immigrants and refugees, low-income individuals, and marginalized communities, have faced disproportionate impacts of infection rates, less access to healthcare services, and overall poorer health outcomes.

Amidst these mounting challenges, the strain on resources and the dire need for comprehensive reforms underscore the urgency to address the systemic issues head-on. While some have advocated privatizing healthcare as a potential solution, such a strategy requires careful consideration, as it may not offer a one-size-fits-all answer to the complex array of issues the health system confronts and would ultimately increase the disparities.

Upcoming Healthcare Challenges

Looking ahead, the challenges faced by Canada's health system are poised to intensify, necessitating proactive measures to address forthcoming complexities. An imminent concern revolves around the significant aging population, with adults aged 65 years or over expected to surge by 68% over the next two decades[12]. This demographic shift will inevitably drive-up healthcare costs as the system grapples with meeting the escalating care needs of older adult individuals. It is imperative to prepare for this demographic shift to ensure that healthcare services remain accessible and responsive to the evolving requirements of our aging population.

Simultaneously, the looming spectre of the climate crisis presents another pressing health system challenge. Canada's population is increasingly exposed to extreme weather events, such as heatwaves, flooding, and brief yet intense exposures to air pollution due to wildfires[13–15]. The ongoing global climate crisis will likely exacerbate these exposures, amplifying the demand for health system services and resources. Alarmingly, these climate change-related events disproportionately affect low-income communities and communities of colour, accentuating existing health disparities[16]. As a result, addressing climate change impacts necessitates a health system and supporting infrastructure that is responsive to increased utilization and actively prioritizes equity in its approach.

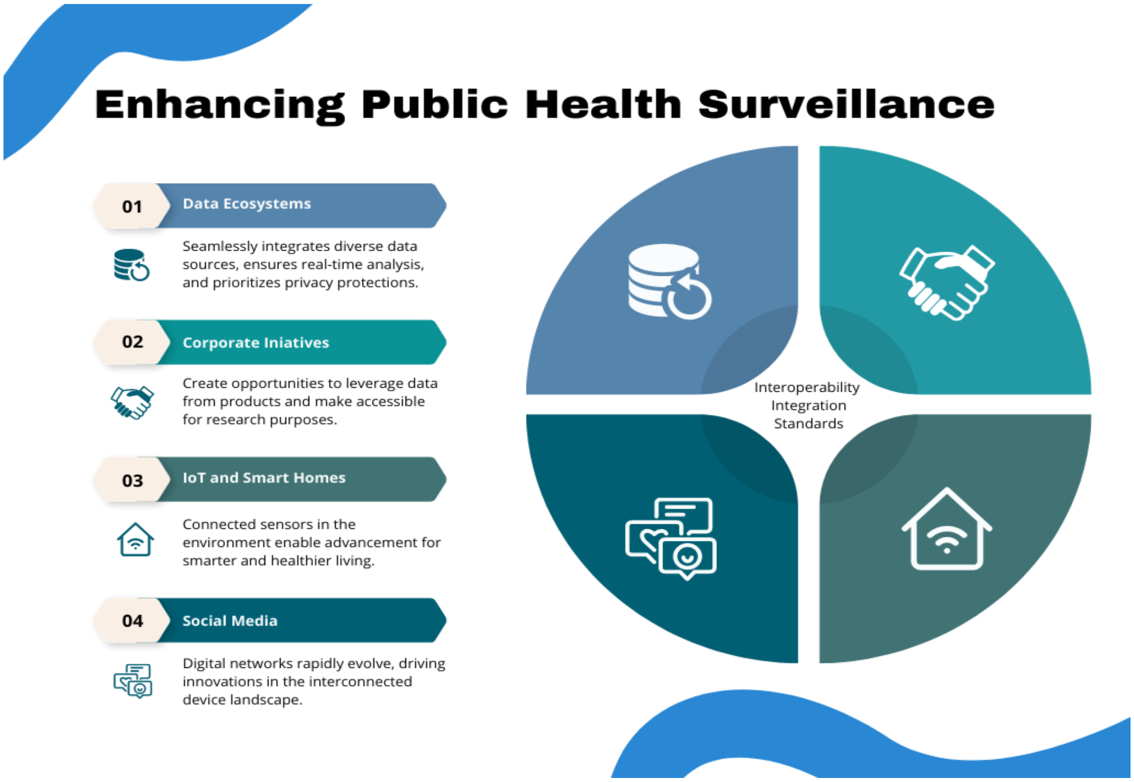
In light of these looming issues, the future of Canada's health system calls for comprehensive and adaptive strategies. Focused efforts should be directed toward preparing for the aging population's demands and mitigating the health consequences of climate change. Proactive resource allocation, technology integration, and interdisciplinary collaborations will be critical components of a robust health infrastructure that addresses these multifaceted challenges. Additionally, embracing and actively integrating principles of equity and inclusivity will be indispensable in ensuring that all population segments receive fair and equitable healthcare services.

We advocate for a fundamental shift in the health system, prioritizing prevention over treatment. By integrating traditional and emerging data sources, a comprehensive understanding of individuals' health can be achieved, allowing for real-time health monitoring. Embracing such an approach not only promises to curtail healthcare costs but, more importantly, can potentially diminish the need for frequent access to healthcare for treatment as proactive health maintenance becomes the norm. Consequently, this shift would lead to reduced health system utilization, alleviating the strain and financial burden it currently faces. Such preparedness will be vital as we brace ourselves for the inevitable surge in healthcare demands brought about by the aging population and the impacts of climate change.

Central to this proposal for Next-Gen Public Health Surveillance is incorporating novel data streams alongside traditional sources. Utilizing novel data technologies, social media data, wearables, and other health and environmental monitoring devices will contribute to a more holistic and timely assessment of individuals' well-being. Harnessing these innovative data sources will empower healthcare professionals and public health officials to make informed decisions, guiding interventions to prevent health issues before they escalate. By focusing on prevention, the health system can address health concerns at their nascent stages, reducing chronic illnesses' incidence and associated costs. Moreover, empowering individuals to monitor their health will foster a proactive self-care and wellness culture, ultimately leading to improved health outcomes.

Throughout the rest of this chapter, we will delve into the diverse data sources that can be integrated into this transformative framework. Furthermore, we will explore the dynamic interactions between healthcare and public health within this context while examining the challenges and considerations that must be carefully weighed before implementing this innovative system. Finally, we will present our model of what this transformed healthcare framework might look like, paving the way for a resilient and equitable health system that is better prepared to meet future demands.

New Sources of Data for Public Health



01 **Data Ecosystems**

Seamlessly integrates diverse data sources, ensures real-time analysis, and prioritizes privacy protections.

02 **Corporate Iniatives**

Create opportunities to leverage data from products and make accessible for research purposes.

03 **IoT and Smart Homes**

Connected sensors in the environment enable advancement for smarter and healthier living.

04 **Social Media**

Digital networks rapidly evolve, driving innovations in the interconnected device landscape.

Figure 1. Enhancing Public Health Surveillance through data ecosystems, corporate initiatives, IoT and Smart Home integration, and social media.

In recent years, the landscape of public health surveillance has undergone a transformative shift, driven by rapid advances in information technology. This paradigm shift has ushered in new sources of data collection that hold great promise for enhancing public health surveillance research. Among the most prominent of these sources are Internet of Things (IoT) sensors, which are ubiquitous devices embedded in various objects and environments, capable of capturing real-time health-related data. IoT sensors offer new opportunities to monitor vital signs, air quality, temperature and humidity, air pollution, and weather, thereby enabling early detection of disease outbreaks and environmental hazards. Integrating IoT data into public health surveillance frameworks allows for a comprehensive understanding of population health trends as well as more timely and targeted interventions.

Corporate initiatives, such as corporate social responsibility (CSRs) programs are examples of data collection and sharing with researchers. Examples include the use of ecobee for sleep and heatwave monitoring, Fitbit for sleep data, and Google Mobility data during the pandemic. Additionally, the integration of mobility data has become another useful data-source in public health surveillance, thanks to the widespread use of smartphones and location-based services. Mobility data provides valuable insights into population movement patterns, facilitating the tracking of infectious disease spread and predicting potential outbreaks. By understanding population mobility, health authorities can implement interventions, such as targeted vaccination campaigns and travel advisories, to curb disease transmission and safeguard public health on both local and global scales.

Other sources of data include social media platforms. Social media has revolutionized global communication, but also through publicly available Application Programming Interfaces (API), public health researchers can collect user data (to an extent) for research purposes. Social media data mining presents an invaluable tool for understanding health behaviors, sentiments, and trends in real-time. Researchers can tap into ultra-large repositories of user-generated content to track the spread of infectious diseases, assess public attitudes towards vaccination, and identify emerging health concerns. Though these data sources pose challenges related to privacy and data accuracy, innovative methodologies for sentiment analysis and machine learning hold the potential to extract valuable insights from this digital landscape and enhance public health surveillance efforts.

Environmental data has long been recognized as a crucial component in understanding the complex interplay between the environment and human health. Traditional data collection methods have been complemented by available remote sensing technologies and spatial data analysis techniques. These advancements allow for continuous monitoring of environmental factors such as air and water quality, temperature variations, and pollutant levels. Incorporating environmental data into public health surveillance systems empowers researchers to discern patterns in disease transmission, identify vulnerable communities, and implement targeted preventive measures to mitigate the impact of environmental hazards on public health.

Moreover, to overcome challenges related to data privacy and access, the concept of synthetic data has gained traction in public health research. Synthetic data involves generating artificial datasets that mimic real-world data while protecting individual privacy. State-of-the-art research on synthetic data claims that researchers can simulate scenarios, conduct robust analyses, and develop predictive models without compromising sensitive information; although, synthetic data is a relatively new approach, and should be continuously monitored by those involved. The incorporation of synthetic data into public health surveillance research has the potential to facilitate PHI sharing on a national or global scale and ensure data-driven decision-making while "*possibly*" upholding the highest standards of privacy protection.

As these new sources of data continue to evolve, a continuous effort in improving ethical considerations, data security, and the establishment of robust data sharing frameworks is needed. Nonetheless, the integration of IoT sensors, social media, environmental monitors, mobility data, and the strategic use of synthetic data into public health surveillance holds the potential to improve public health policies and interventions outcomes, fostering more precise, responsive, and data-driven approaches to safeguarding public health and improving overall population well-being.

Perhaps, now is an opportune moment to delve into the critical necessity for a collaborative effort between corporate entities, along with regulatory authorities and/or governments. Frequently, when novel technologies emerge, lawmakers find themselves trailing behind such rapid advancements. They grapple with comprehending the full extent of these innovations' implications, leading to the establishment of barriers that hinder both growth and seamless integration. On the other hand, corporations often move at a fast pace, and in their eagerness to innovate, they may sometimes overlook or unintentionally ignore existing norms of technology implementation. In an ideal world, these two entities should join forces, working hand in hand. Corporations must actively engage with regulatory authorities to ensure their technologies conform to existing frameworks, ensuring harmonious coexistence. And when necessary, they should actively participate in the development of new frameworks that guarantee the sustainability and suitability of their technologies. On the other end, policymakers should actively collaborate with technology companies to incorporate their advancements into health practices. By empowering these companies, they can play a crucial role in strengthening the overall health systems. This alliance fosters an environment where innovation can flourish while ensuring that public health remains at the forefront of technological progress.

Public Health interactions with the Data

Public health service delivery is dependent on data, which serves as the basis for comprehending and enhancing population health outcomes. As public health experts, we engage with data regularly throughout the entire process, from collecting, to processing and cleaning, through analysis and decision-making. The importance of this data cannot be overstated, as it plays a crucial role in detecting health trends, identifying health risk factors, tracking disease patterns, as well as creating effective interventions and policies[17].

By using data-driven insights, public health professionals have a better understanding of the social, environmental, and commercial factors that impact health[18–21]. This knowledge enables the identification of disparities and development of targeted interventions to address health outcomes. Reliable data is crucial to assess the effectiveness of health intervention programs, track risk factors and indicators, or to forecast and monitor disease control efforts[22]. Without reliable data, public health initiatives may suffer from inefficiencies, wasted resources, and missed opportunities to improve health outcomes.

Data collection methods have existed to serve public health practice, including surveys, surveillance systems, data registries, and repositories. Traditional collection methods, such as paper-based or telephone surveys, provide valuable insights into the population's health status, behaviours, and needs. Regional and national surveillance systems can continually monitor disease trends and provide early warnings of novel disease outbreaks. Registries and repositories can help to track specific conditions or populations over time, supporting ongoing research and monitoring efforts. Newer technologies, such as wearables and IoT sensors that collect data 24/7, have created data lakes that can be leveraged for new insights and novel public health surveillance methods[23].

Nevertheless, all data collection methods come with their unique challenges and opportunities. For example, surveys may suffer from response bias, performance bias or recall bias, and limited representativeness (i.e. gender, race, sexual orientation, etc.). A surveillance system must balance sensitivity and specificity to avoid over-reporting or missing cases, and the generalizability of results must be considered. Registries and repositories also require careful design and maintenance to ensure data completeness and accuracy. Embracing novel technological solutions, such as mobile data collection, consumer-grade wearable devices and digital health platforms, presents opportunities to improve efficiency, representativeness, and data accuracy[24].

Once the data is collected, efficient data management pipelines and ecosystems are essential for data flow and sharing with public health professionals and corporations[25]. Data-sharing agreements enable stakeholder collaboration, allowing a comprehensive and holistic understanding of population-level health[26]. However, stringent protocols are vital to ensure data privacy and cybersecurity, especially with the rise of digital health and electronic medical and health records[27,28]. Public perception of public health and trust must also be managed and respected, as demonstrated with COVID-19[29,30].

As digital health data becomes more pervasive in public health, ethical considerations and bias from training data in AI must be critically examined[31]. Public health professionals and corporate entities must navigate the ethical dilemmas associated with data collection, storage, and usage. User rights to privacy and security must be met and protected to prevent harm. Additionally, while powerful tools, AI systems must be designed and trained with safeguards to avoid perpetuating existing systemic biases that would impact health outcomes.

With the rapid evolution of technology, data analytic skills development necessitates continuous training and education for public health professionals. Public health professionals could effectively interpret and communicate their findings with data literacy, statistical analysis skills, and proficiency in data visualization tools. Thus, it is very important that public health professionals remain informed and that training programs and core curriculum competencies are updated routinely with emerging data sources, analysis techniques, and innovative approaches[32].

Public health institutes, especially in Canada, depend on the trust and confidence of their communities. The COVID-19 pandemic and numerous climate crisis events have underscored the importance of nurturing and respecting public trust. Transparent data governance, ethical data practices, and clear communication channels are essential to maintaining the public's trust. Nefarious actors benefit from the noise and chaos created by sowing seeds of distrust in an increasingly polarized climate. Public health practitioners must continue to engage with communities, address these concerns, and build partnerships to foster trust and cooperation.

Unfortunately, challenges remain in low and middle-income settings despite the immense potential that data-driven public health practices offer. Limited resources, remote locations, physical infrastructure, and access to technology can hinder effective data collection, access, and analysis. If properly addressed, these communities will continue to be included in the benefits of digital health technologies, further narrowing the digital divide. Public health practitioners must leverage low-cost data development and computational tools to promote sustainability and reach these populations. Innovative approaches, such as digital literacy programs, funding and incentives, leveraging mobile technology, and community health workers can help bridge gaps and ensure that data-driven insights are accessible and actionable in resource-poor settings.

The key to the future of public health is its ability to utilize data effectively. This can be achieved through reliable data collection methods, ethical data management practices, and responsible use of AI systems. It is crucial for public health professionals to continually educate themselves and stay up to date with the latest new technologies to make data-driven decisions. However, building and maintaining trust in public health institutions is equally important as data becomes more integrated in the field. Public health has been underfunded and systematically dismantled by neoliberal governments in a move towards privatization[33,34]. This is a dangerous precedent as it would only weaken the overall healthcare system and the most marginalized, women and children especially will face greater inequities. Additionally, it is vital to promote sustainability in low-income areas to ensure that all communities benefit from data-driven public health initiatives. By prioritizing trust, responsible technology use, and equity and sustainability, Next-Generation Public Health Surveillance can make a significant positive impact on global health.

Big Data and Public Health

The Public Health Agency in Canada is responsible for promoting health, preventing and controlling infectious diseases, chronic diseases, injuries, knowledge transfers between Canada and the rest of the world, preparing and responding to public health emergencies, research and development of Canada's public health programs, to strengthen international collaboration on public health and coordinate national strategies to public health and planning[35]. Recently, due to the advances in Big Data and IoT technologies, public health agencies have the opportunity to conduct research on large amounts of data, often times Big Data, generated from smart devices (e.g., smartphones, smartwatches), wearables (e.g., fitness bands and watches, smart clothing), and environmental sensors (e.g., smart thermostats, air pollution sensors, door sensors)[36–38]. Big Data's definition has been discussed by a number of authors[39–41]; which the chosen definition is based on, in which the author proposes a formal definition of Big Data based on its essential features:

"Big Data is the Information asset characterized by such a High Volume, Velocity, and Variety to require specific Technology and Analytical Methods for its transformation into Value." [42]

In the above definition, high volume refers to the high volume of data being generated from all interconnected devices[42]. Statista reports that data generated from IoT and smart devices will go from 2 Zettabytes in 2010 to around 131 Zettabytes in 2025[43]. This includes data created, captured, copied, and consumed globally by consumers, researchers, private business, public institutions, and governments. Based on the International Data Corporation (IDC) 2020 report, from all current data measured, the ratio of unique data created and captured to data copied and consumed is 1:9, with an expectation of reaching 1:10 by 2024[44]. Such change reflects how much generated data from IoT and non-IoT devices are desired by third-parties data processors.

Moreover, another IDC report from 2021 showed that due to the COVID-19 pandemic[45], the world experienced an increase in data generated, reaching 64.2 Zettabytes in 2020. Nevertheless, less than 2% of this data was preserved due to being temporary data for research, being overwritten with new data or ephemeral data. Velocity refers to how fast new data is generated, replicated, processed, and consumed[42]. Traditional software tools are not capable of handling the huge datasets generated from different data sources with a great number of columns. Variety is related to the differences between big data sources, such as different devices, data formats, transport protocols, types of encryptions, overarching laws and regulations, semantics, structured, semi-structured, and unstructured data are some of the aspects of big data that data processors must take in consideration[42].

Big data in public health generally originates from Electronic Health Records (EHR), the Personal Health Record (PHI), Medical Practice Management Software (MGM), healthcare provider data, genomics data, and data collected from IoT and smart devices. Mainly, EHR, PHI, self-reported data (e.g., surveys), and IoT data are often the most common and present in public health research. Big data has been applied in public health to help with better identification, monitoring, assessment, prevention and treatment, and outcomes of diseases; it is also used for generating insights to help the creation of better public health interventions. For example, big data can provide national or global scale information about smoking, drinking, drugs, exercise, chronic health conditions and mental health, which when provided to public health researchers can provide invaluable behavioral information in large scales that was not possible before big data was widely available[46–52]. Big data also provides the public health means to accelerate the timing from acquiring knowledge to translation and transferring of knowledge to the proper target audience[53].

Challenges with Big Data

Big data is on the front line of the new age of public health research; however, challenges to leverage such data are still barriers for most researchers due to a lack of specific skills and infrastructure. Concerns regarding ethical utilization of big data and knowledge translation from methods used in the entire life cycle of research are also relevant challenges, since public trust is imperative for efficient and effective (e.g., fit for purpose) big data collection.

Furthermore, if public health agencies and industry intend to work together to achieve seamless data collection and aggregation, the challenges of interoperability need to be addressed. Currently, in the Canadian landscape, there are no mandatory laws or regulations demanding adoption of interoperability technologies, which hinders the progress of big data collection and public health innovation. It is worth mentioning that there are initiatives currently ongoing trying to promote interoperability and its benefits in Canada, such as the Canada Health Infoway[54] and Digital Health Canada[55].

Data collection from different devices and sources is challenging because not all devices give researchers direct access to data. Some devices require integration with third parties API and the writing of scripts to collect data from private clouds automatically. Moreover, not all devices' data stewards have the appropriate authority within the organization to authorize access to their client's collected data. The low levels of maturity[56] of data governance programs [57], specifically inter-organizational data governance programs [58] in the public health domain's stakeholders at an organizational and infrastructure level, create barriers to achieving data sharing agreements for research. For example, the lack of inter-organizational data governance creates barriers to interoperability because each organization silos their big data in different data structures, granularity, and types of storage which causes challenges when these silos need to be integrated. In other cases, the quality of generated data is not optimal, having low accuracy or unacceptable reliability, thus creating precisely inaccurate outcomes[59,60].

The universal quality dimensions utilized to evaluate big data datasets are commonly defined as Availability, Usability, Reliability, Relevance, and Presentation.

- Availability concerns challenges regarding how efficient and effective the big data's life cycle is. For example, how long does it take for the data collected from end-users to be available to a researcher, or what type of authorization is needed to access the data?
- Usability regards information about the data and how credible their sources are.
- Reliability concerns are how accurate and appropriate a big data dataset is for research.
- Relevance is centred on the appropriateness of the data that will be used for research. For example, is the data directly related to the problem or question at hand? Is it capable of providing valuable insights and information for the specific analysis or application?
- Presentation quality regards how data and information extracted from big data is communicated to stakeholders and how effective such communication is in positively impacting the general population's health.
- Figure 2 below illustrates these dimensions and their sub-domains:

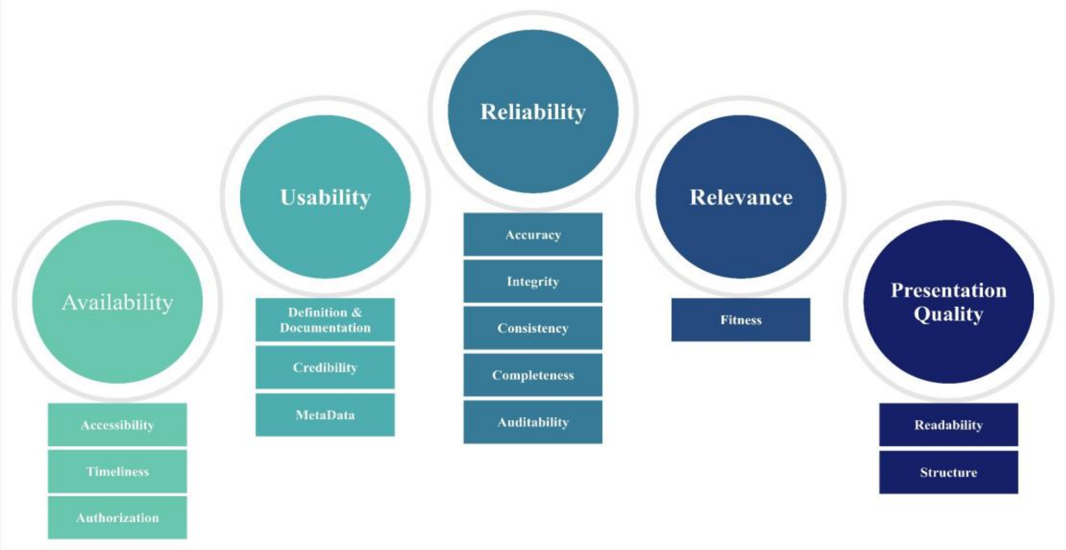


Figure 2. Dimensions used to evaluate big data's life-cycle quality[59] .

Data storage challenges include space and privacy protection. Space is related to the constantly ever-growing datasets prevalent in big data research. Although storage space is currently cheaper, larger datasets with terabytes, petabytes or even exabytes can incur excessive costs if stored in an inefficient and inadequate environment. More sophisticated storage for large datasets in private cloud storage is still costly and requires specific knowledge of the cloud services they are inserted into. With regards to privacy protection, although a good portion of research over EHR and PHI is performed utilizing time series data previously anonymized, de-anonymization is still a reality that should not be underestimated. For example, one case that gained notoriety on research over big data was a data set provided in 2006 by Netflix, which offered a prize for research that could best predict its users' movies ratings. They replaced sensitive information such as usernames with random numbers and scrambled personal details as the strategy for anonymization of the dataset. The author linked Netflix's dataset with Internet Movie Database (IMDB) non-anonymized users' movies reviews, and were able, with relatively little information from IMDB, to re-identify users with 68% of success[61].

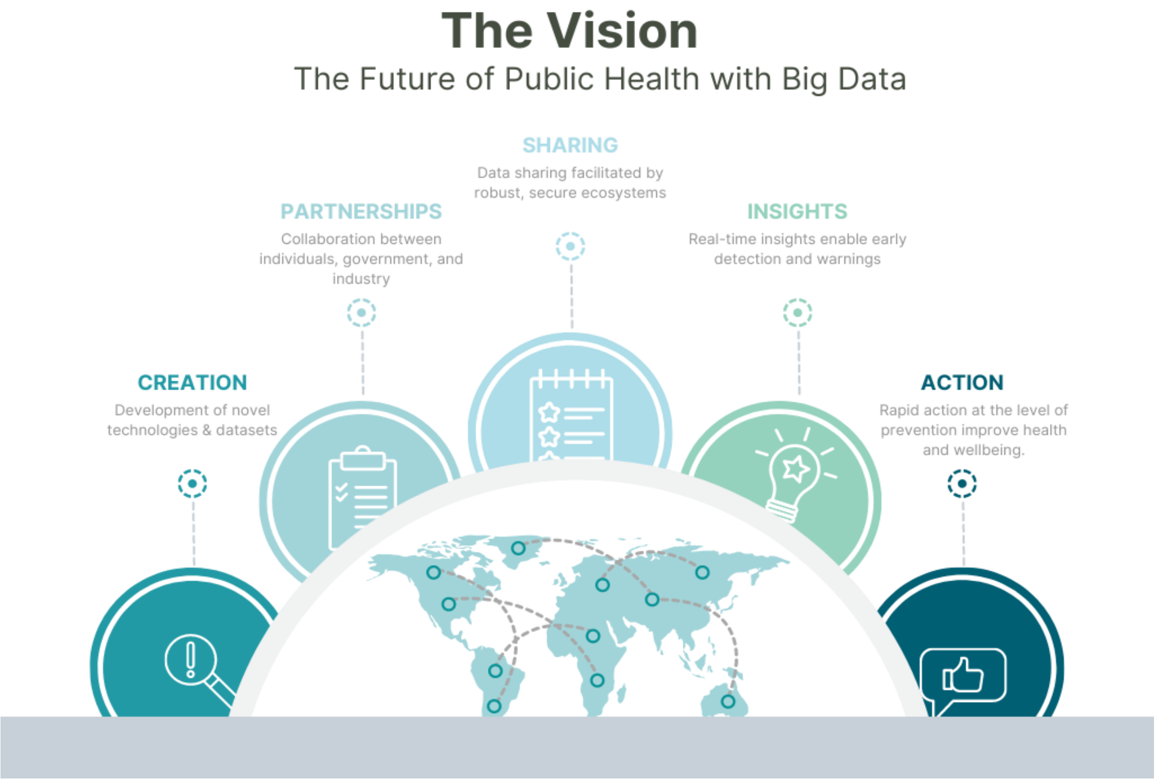
With the rise of Artificial Intelligence (AI) and Large Language Models (LLMs), it is also important to be aware of the bias within the datasets used for training. There can be serious consequences, as a report from the Center for Countering Digital Hate found that Google's Bard generated misinformation and disinformation, including racism, misogyny, and antisemitism (even creating fake evidence)[62]. Governments and private companies must work together to create laws that prevent and protect against information that could undermine public health efforts, the democratic process, and the foundational truths held by society.

Data processing offers challenges involving appropriate data pre-processing, classification, clustering, association, and other steps, depending on the method being utilized for the data analysis. The processing of large datasets requires large amounts of processing power and are often executed under a private cloud environment using third party clusters, which requires relevant funding for the research. The bigger and lower quality the dataset, the harder and more costly the data processing will be[36,63–66].

In summary, although significant advances in big data analysis have been consistently improving and benefiting society[67], the backbone infrastructure and local infrastructure still needs to be improved and scaled even further to facilitate the collection of Personal Health Information (PHI) in nationwide scales.

The Vision

In an ideal world, Next-Generation Public Health Surveillance could revolutionize the delivery of public health by leveraging technological advancements and data-driven insights to tackle future infectious disease outbreaks, chronic conditions prevention, and promote health equity[68,69]. Optimistic about this future, we envision a world where early detection and prevention is where the focus is, healthcare is personalized, and global cooperation fosters sustainability and reduces global inequities[70].



The Vision

The Future of Public Health with Big Data

SHARING

Data sharing facilitated by robust, secure ecosystems

PARTNERSHIPS

Collaboration between individuals, government, and industry

INSIGHTS

Real-time insights enable early detection and warnings

CREATION

Development of novel technologies & datasets

ACTION

Rapid action at the level of prevention improve health and wellbeing.

Figure 3. The components required to deliver this vision of Public Health with Big Data.

With the invention of highly effective vaccines and antibiotics, eradicating infectious diseases has been a key achievement in public health. With early detection of environmental and lifestyle risk factors, potential illnesses would be met with timely interventions, thereby arresting disease progression, and reducing costs for treatment both for the individual and the system.

The health systems of tomorrow will prioritize addressing behavioural and lifestyle risk factors to educate and inform individuals. Widespread consumer-driven technology adoption has produced vast amounts of data that public health professionals can harness to target interventions which encourage healthier choices and inform policy decisions. The major developments in consumer-driven technologies will empower individuals to take charge of their health. Self-monitoring through commercially available devices means that the greatest number of individuals have easy access to data about their health, and individuals can become active advocates for personalized healthcare, leading to better health outcomes.

Global agreements facilitating seamless health data exchanges through interoperability standards are a cornerstone of enhanced Public Health Surveillance. It benefits the entire population through positive action and cooperation between healthcare providers, corporations, and technology companies that would enable comprehensive and real-time health information sharing. Most importantly, these advancements create an opportunity to prioritize health equity, tip the scale of inequity back, and give equitable access to healthcare resources, technology, and health. Opportunities for leapfrog technologies enable bridging the digital divide to bring transformative healthcare solutions to the Global South and other regions lacking physical infrastructure and resources.

With humanity standing well in the path of the climate crisis, addressing the disproportionate burden of diseases and consequences for vulnerable populations must be a priority. With this kind of data, direct region-specific solutions may be tailored to the unique challenges faced by these communities. We must also acknowledge the environmental impacts of technology and the additional resources required to power Next-Gen Surveillance. By developing energy-efficient data warehouses and employing green or nature-based cooling and backup systems, these solutions will contribute minimally or, at best, not at all to total GHGs[71–73]. Additionally, mining practices for metals needed for new technologies, such as semiconductors and batteries, should ensure fair and safe labour practices, such as not employing children[74–76].

The current health system is built upon reactionary care delivery and could be improved by embracing a fully preventive system. Focusing on upstream risk factors through predictive modeling and early warning systems, environmental and public health professionals could intervene before outbreaks to save lives and resources[77].

This must be achieved through strong public-private partnerships; collaboration must be transdisciplinary. The result will be innovative solutions, access to resources, collective action toward the common good, and recognition of health as the highest form of wealth[78].

While ambitious and optimistic, our vision of Next-Generation Public Health surveillance is not impossible. In its pursuit, we must embrace rapid technological advancements, empower individuals, and foster global cooperation. In preparing for the future, this approach has the potential to transform public health outcomes, eliminate infectious diseases, reduce harmful environmental exposures, and address the epidemic of chronic diseases. It requires dedication, collaboration, and a collective commitment to ensuring health equity and sustainability, which align with the climate targets such as the Climate Paris Agreement and the UN Sustainable Development Goals[79]. As public health practitioners, it's vital to prioritize the well-being of our communities, bridge the gaps, and build a future that supports the universal right to health[80,81].

Charting a path forward in Canada

The current state of our health system is undoubtedly complex and precarious. We find ourselves standing at the precipice of a time where the system faces increasing strain due to its slow and fragmented nature, resulting in financial burdens and inequitable care for many. Furthermore, the challenges are only expected to become exacerbated as the country grapples with an aging population and the effects of climate change. Yet, amidst these challenges, there lies an opportunity to revolutionize healthcare by embracing novel and innovative data that can significantly bolster our health systems. This paper sets forth a compelling vision for integrating new sources of data and proposes a viable pathway forward. In this concluding section, several crucial recommendations are outlined.

Firstly, there is an urgent need to enhance literacy, knowledge, and public awareness surrounding health data. Safeguarding the use of this data is paramount. Secondly, a centralized big data strategy must be formulated and implemented. As we contemplate dealing with massive data sets, it becomes evident that they demand distinct computational power and storage solutions. Thirdly, we must invest in training public health professionals and future students in the nuances of handling and utilizing big data. Updating training programs is essential to equip them with the necessary skills to work effectively with such data. Lastly, fostering collaborative efforts between corporate and public sectors is crucial. While individual data may not be the focus, population-level data holds immense power in the realm of public health. Establishing mechanisms that promote collaboration between different entities is vital for the successful implementation of these ideas.

This chapter outlines a comprehensive vision for the incorporation of these strategies. Embracing this transition will ensure that we not only address the needs of the Canadian population but also alleviate the strain on our health system. By leveraging the potential of big data, we can forge a path towards a more robust and responsive health system.

References

1. Martin, D. *et al.* Canada's universal health-care system: achieving its potential. *The Lancet* **391**, 1718–1735 (2018).
2. Levine, S. Health Care Industry Insights: Why the Use of Preventive Services Is Still Low. *Prev. Chronic. Dis.* **16**, (2019).
3. Duong, D. & Vogel, L. Overworked health workers are "past the point of exhaustion". *Can. Med. Assoc. J.* **195**, E309–E310 (2023).
4. Zaidi, D., Journalist, Ctvn. ca D. & Contact, F. |. Nurses more likely to quit in next 3 years, health worker survey finds. *CTVNews* https://www.ctvnews.ca/health/nurses-more-likely-to-quit-in-next-3-years-health-worker-survey-finds-1.5932193 (2022).
5. How Canada's emergency rooms are faring as major staffing shortages persist. *The Globe and Mail* (2023).
6. National health expenditure trends, 2022 — Snapshot | CIHI. https://www.cihi.ca/en/national-health-expenditure-trends-2022-snapshot.
7. Williamson, D. L. *et al.* Low-income Canadians' experiences with health-related services: Implications for health care reform. *Health Policy* **76**, 106–121 (2006).
8. Kasman, N. M. & Badley, E. M. Beyond access: who reports that health care is not being received when needed in a publicly-funded health care system? *Can. J. Public Health Rev. Can. Sante Publique* **95**, 304–308 (2004).
9. Wilson, K. & Rosenberg, M. W. Accessibility and the Canadian health care system: squaring perceptions and realities. *Health Policy Amst. Neth.* **67**, 137–148 (2004).
10. Wallace, B. B. & MacEntee, M. I. Access to Dental care for Low-Income Adults: Perceptions of Affordability, Availability and Acceptability. *J. Community Health* **37**, 32–39 (2012).
11. Bryant, T., Leaver, C. & Dunn, J. Unmet healthcare need, gender, and health inequalities in Canada. *Health Policy Amst. Neth.* **91**, 24–32 (2009).
12. Infographic: Canada's seniors population outlook: Uncharted territory | CIHI. https://www.cihi.ca/en/infographic-canadas-seniors-population-outlook-uncharted-territory.
13. Heat wave continues to sizzle across Canada with weather warnings | Globalnews.ca. *Global News* https://globalnews.ca/news/9814271/canada-heat-wave-july-2023/.
14. Nova Scotia wrestles with aftermath of devastating floods | Reuters. https://www.reuters.com/business/environment/atlantic-canada-province-wrestles-with-aftermath-devastating-floods-2023-07-23/.
15. Reporter, M. V. S. Toronto's air quality is among the worst in the world due to wildfire smoke. *Toronto Star* https://www.thestar.com/news/gta/toronto-s-air-quality-is-among-the-worst-in-the-world-due-to-wildfire-smoke/article_81645fb7-20ed-5a77-9979-33a23bba85f1.html (2023).
16. Hayes, K., Berry, P. & Ebi, K. L. Factors Influencing the Mental Health Consequences of Climate Change in Canada. *Int. J. Environ. Res. Public. Health* **16**, 1583 (2019).
17. Mooney, S. J. & Pejaver, V. Big Data in Public Health: Terminology, Machine Learning, and Privacy. *Annu. Rev. Public Health* **39**, 95–112 (2018).
18. Williams, D. R., Costa, M. V., Odunlami, A. O. & Mohammed, S. A. Moving Upstream: How Interventions that Address the Social Determinants of Health can Improve Health and Reduce Disparities. *J. Public Health Manag. Pract. JPHMP* **14**, S8-17 (2008).
19. Public Health Agency of Canada. Social determinants of health and health inequalities. https://www.canada.ca/en/public-health/services/health-promotion/population-health/what-determines-health.html (2020).
20. Prüss-Ustün, A. *et al.* Diseases due to unhealthy environments: an updated estimate of the global burden of disease attributable to environmental determinants of health. *J. Public Health* **39**, 464–475 (2017).
21. Schulz, A. & Northridge, M. E. Social Determinants of Health: Implications for Environmental Health Promotion. *Health Educ. Behav.* **31**, 455–471 (2004).
22. Soyiri, I. N. & Reidpath, D. D. An overview of health forecasting. *Environ. Health Prev. Med.* **18**, 1–9 (2013).
23. Zagan, E. & Danubianu, M. Data Lake Approaches: A Survey. in *2020 International Conference on Development and Application Systems (DAS)* 189–193 (2020). doi:10.1109/DAS49615.2020.9108912.
24. Ming, D. K. *et al.* Continuous physiological monitoring using wearable technology to inform individual management of infectious diseases, public health and outbreak responses. *Int. J. Infect. Dis.* **96**, 648–654 (2020).
25. van Panhuis, W. G. *et al.* A systematic review of barriers to data sharing in public health. *BMC Public Health* **14**, 1144 (2014).
26. Abernethy, A. *et al.* The Promise of Digital Health: Then, Now, and the Future. *NAM Perspect.* **2022**, 10.31478/202206e.
27. Birkhead, G. S., Klompas, M. & Shah, N. R. Uses of Electronic Health Records for Public Health Surveillance to Advance Public Health. *Annu. Rev. Public Health* **36**, 345–359 (2015).
28. Kolasa, K., Mazzi, F., Leszczuk-Czubkowska, E., Zrubka, Z. & Péntek, M. State of the Art in Adoption of Contact Tracing Apps and Recommendations Regarding Privacy Protection and Public Health: Systematic Review. *JMIR MHealth UHealth* **9**, e23250 (2021).
29. Siegrist, M. & Bearth, A. Worldviews, trust, and risk perceptions shape public acceptance of COVID-19 public health measures. *Proc. Natl. Acad. Sci.* **118**, e2100411118 (2021).
30. Swire-Thompson, B. & Lazer, D. Public Health and Online Misinformation: Challenges and Recommendations. *Annu. Rev. Public Health* **41**, 433–451 (2020).

31. Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D. & Tzovara, A. Addressing bias in big data and AI for health care: A call for open science. *Patterns* **2**, 100347 (2021).
32. Benke, K. & Benke, G. Artificial Intelligence and Big Data in Public Health. *Int. J. Environ. Res. Public. Health* **15**, 2796 (2018).
33. Guyon, A. & Perreault, R. Public health systems under attack in Canada: Evidence on public health system performance challenges arbitrary reform. *Can. J. Public Health.* **107**, e326–e329 (2016).
34. Hemenway, D. Why We Don't Spend Enough on Public Health. *N. Engl. J. Med.* **362**, 1657–1658 (2010).
35. Government of Canada. About the Agency - Canada.ca. (2019).
36. Firouzi, F. *et al.* Internet-of-Things and big data for smarter healthcare: From device to architecture, applications and analytics. *Future Gener. Comput. Syst.* **78**, 583–586 (2018).
37. Zheng, Z., Xie, S., Dai, H., Chen, X. & Wang, H. An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. in *Proceedings - 2017 IEEE 6th International Congress on Big Data, BigData Congress 2017* 557–564 (Institute of Electrical and Electronics Engineers Inc., 2017). doi:10.1109/BigDataCongress.2017.85.
38. Bagot, K. S. *et al.* Current, future and potential use of mobile and wearable technologies and social media data in the ABCD study to increase understanding of contributors to child health. *Developmental Cognitive Neuroscience* vol. 32 121–129 (2018).
39. Dash, S., Shakyawar, S. K., Sharma, M. & Kaushik, S. Big data in healthcare: management, analysis and future prospects. *J. Big Data* **6**, 1–25 (2019).
40. Hajjaji, Y., Boulila, W., Farah, I. R., Romdhani, I. & Hussain, A. Big data and IoT-based applications in smart environments: A systematic review. *Comput. Sci. Rev.* **39**, 100318 (2021).
41. Wolfert, S., Ge, L., Verdouw, C. & Bogaardt, M. J. Big Data in Smart Farming – A review. *Agric. Syst.* **153**, 69–80 (2017).
42. De Mauro, A., Greco, M. & Grimaldi, M. A formal definition of Big Data based on its essential features. *Library Review* vol. 65 122–135 (2016).
43. Arne Holst. • Total data volume worldwide 2010-2025. *Statista* (2021).
44. Shirer, M. & Rydning, J. IDC's Global DataSphere Forecast Shows Continued Steady Growth in the Creation and Consumption of Data. *Int. Data Corp. IDC* (2020).
45. David Reinsel, John, R. & John, F. G. Data Creation and Replication Will Grow at a Faster Rate than Installed Storage Capacity, According to the IDC Global DataSphere and StorageSphere Forecasts. *Worldwide Global DataSphere Forecast, 2021–2025: The World Keeps Creating More Data — Now, What Do We Do with It All?* (2021).
46. Barrett, M. A., Humblet, O., Hiatt, R. A. & Adler, N. E. Big Data and Disease Prevention: From Quantified Self to Quantified Communities. *Big Data* vol. 1 168–175 (2013).
47. Kramer, A. D. I., Guillory, J. E. & Hancock, J. T. Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8788–8790 (2014).
48. Gamache, R., Kharrazi, H. & Weiner, J. P. Public and Population Health Informatics: The Bridging of Big Data to Benefit Communities. *Yearb. Med. Inform.* **27**, 199–206 (2018).
49. Martin-Sanchez, F. J., Aguiar-Pulido, V., Lopez-Campos, G. H., Peek, N. & Sacchi, L. Secondary Use and Analysis of Big Data Collected for Patient Care. *Yearb. Med. Inform.* **26**, 28–37 (2017).
50. Ajunwa, I., Crawford, K. & Ford, J. S. Health and big data: An ethical framework for health information collection by corporate wellness programs. *J. Law. Med. Ethics* **44**, 474–480 (2016).
51. Chen, M., Hao, Y., Hwang, K., Wang, L. & Wang, L. Disease Prediction by Machine Learning over Big Data from Healthcare Communities. *IEEE Access* **5**, 8869–8879 (2017).
52. Balan, S., Shristiraj, N., Shah, V. & Manjappa, A. Big data analysis of youth tobacco smoking trends in the United States. *Proc. - 2017 IEEE Int. Conf. Big Data Big Data 2017* **2018-January**, 4727–4729 (2017).
53. Khoury, M. J. & Ioannidis, J. P. A. Big data meets public health. *Science* vol. 346 1054–1055 (2014).
54. Home | Canada Health Infoway. https://www.infoway-inforoute.ca/en/.
55. Digital Health Canada: Connect. Inspire. Empower. *Digital Health Canada* https://digitalhealthcanada.com/.
56. Tse, D., Chow, C., Ly, T., Tong, C. & Tam, K. The Challenges of Big Data Governance in Healthcare. in *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)* 1632–1636 (2018). doi:10.1109/TrustCom/BigDataSE.2018.00240.
57. Hafen, E. Personal Data Cooperatives – A New Data Governance Framework for Data Donations and Precision Health. in *The Ethics of Medical Data Donation* (eds. Krutzinna, J. & Floridi, L.) 141–149 (Springer International Publishing, 2019). doi:10.1007/978-3-030-04363-6_9.
58. Da S. E S. Miranda, P. A. & Morita, P. P. A Proposed Personal Data Sovereignty Inter-Organizational Governance Framework for Public Health Research. in *2022 IEEE International Conference on Big Data (Big Data)* 1–1 (2022). doi:10.1109/BigData55660.2022.10020348.
59. Cai, L. & Zhu, Y. The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Sci. J.* **14**, (2015).
60. McFarland, D. A. & McFarland, H. R. Big Data and the danger of being precisely inaccurate: *http://dx.doi.org/10.1177/2053951715602495* **2**, (2015).
61. Narayanan, A. & Shmatikov, V. How To Break Anonymity of the Netflix Prize Dataset. (2006).
62. CCDH. Google's new Bard AI generate lies. *Center for Countering Digital Hate | CCDH* https://counterhate.com/research/misinformation-on-bard-google-ai-chat/.
63. García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M. & Herrera, F. Big data preprocessing: methods and

prospects. *Big Data Anal. 2016 11* **1**, 1–22 (2016).
64. Manogaran, G. *et al.* Big Data Analytics in Healthcare Internet of Things. *Underst. Complex Syst.* 263–284 (2017) doi:10.1007/978-3-319-55774-8_10.
65. Tchagna Kouanou, A. *et al.* An optimal big data workflow for biomedical image analysis. *Inform. Med. Unlocked* **11**, 68–74 (2018).
66. Marjani, M. *et al.* Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges. *IEEE Access* **5**, 5247–5261 (2017).
67. The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns - Nir Kshetri, 2014. https://journals.sagepub.com/doi/full/10.1177/2053951714564227.
68. Willis, V. C. *et al.* Digital Health Interventions to Enhance Prevention in Primary Care: Scoping Review. *JMIR Med. Inform.* **10**, e33518 (2022).
69. Wienert, J., Jahnel, T. & Maaß, L. What are Digital Public Health Interventions? First Steps Toward a Definition and an Intervention Classification Framework. *J. Med. Internet Res.* **24**, e31921 (2022).
70. Wise, A., MacIntosh, E., Rajakulendran, N., Khayat, Z., & 2016. Transforming health: Shifting from reactive to proactive and predictive care. *MaRS Discovery District* https://www.marsdd.com/news/transforming-health-shifting-from-reactive-to-proactive-and-predictive-care/.
71. Lucivero, F. Big Data, Big Waste? A Reflection on the Environmental Sustainability of Big Data Initiatives. *Sci. Eng. Ethics* **26**, 1009–1030 (2020).
72. Brierley, C. Big data's hidden cost. *University of Cambridge* https://www.cam.ac.uk/stories/green-algorithms (2023).
73. Rong, H., Zhang, H., Xiao, S., Li, C. & Hu, C. Optimizing energy consumption for data centers. *Renew. Sustain. Energy Rev.* **58**, 674–691 (2016).
74. McKinsey Global Institute. *Beyond the Supercycle - How Technology is reshaping resources*. https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Sustainability/Our%20Insights/How%20technology%20is%20reshaping%20supply%20and%20demand%20for%20natural%20resources/MGI-Beyond-the-Supercycle-Full-report.pdf (2017).
75. Allan-Blitz, L.-T., Goldfine, C. & Erickson, T. B. Environmental and health risks posed to children by artisanal gold mining: A systematic review. *SAGE Open Med.* **10**, 20503121221076936 (2022).
76. Amnesty International. Exposed: Child labour behind smart phone and electric car batteries. *Amnesty International* https://www.amnesty.org/en/latest/news/2016/01/child-labour-behind-smart-phone-and-electric-car-batteries/ (2016).
77. Kong, E. Moving from a reactive to a proactive society: recognizing the role of environmental public health professionals. *Environ. Health Rev.* **65**, 41–43 (2022).
78. Bloom, D. E. & Canning, D. The Health and Wealth of Nations. *Science* **287**, 1207–1209 (2000).
79. Hassani, H., Huang, X., MacFeely, S. & Entezarian, M. R. Big Data and the United Nations Sustainable Development Goals (UN SDGs) at a Glance. *Big Data Cogn. Comput.* **5**, 28 (2021).
80. Nunes, R., Nunes, S. B. & Rego, G. Health care as a universal right. *J. Public Health* **25**, 1–9 (2017).
81. Pillay, N. Right to health and the Universal Declaration of Human Rights. *The Lancet* **372**, 2005–2006 (2008).

Making Canadian Healthcare Systems “AI READY”: What do we Need to Build AI-Powered Trustworthy Primary Healthcare Solutions?

Sirisha Rambhatla*
Management Sciences Department
University of Waterloo
Waterloo, ON N2L 3G1
sirisha.rambhatla@uwaterloo.ca

ABSTRACT

With the double whammy of aging populations and healthcare worker shortages sending shockwaves throughout the healthcare systems in Canada, there is an urgent need to develop and deploy AI-powered healthcare solutions to assist and support our healthcare workers. While AI for healthcare solutions have made tremendous progress in recent years, translating these gains to improve quality of service (QoS) for real-world healthcare systems has been extremely slow despite widespread interest from researchers, hospitals, and policymakers. This is primarily because our healthcare systems are not “Artificial Intelligence (AI) Ready”. In this paper, we first outline various ways AI-powered solutions can revolutionize primary healthcare based on our research in liver transplantation and burn surgical candidacy forecasting. Second, via an example of AI-based surgical skill assessment in radical prostatectomy, we highlight how AI can play a key role in tooling and training our healthcare workforce in highly skilled tasks. Next, we delve into the data aspects of machine learning for trustworthy and fair AI modeling in healthcare systems and hospitals, calling attention to the need to tackle the new wave of large language model-driven misinformation. We then discuss the current challenges faced by hospitals and conclude with recommendations to channelize the wealth of information in our healthcare systems to positively impact the lives of Canadians and people across the globe.

Keywords AI for healthcare · Survival Analysis · Deep Learning · LLMs More

1 Introduction

Timely access to healthcare is critical for prevention of disease and in saving lives. As healthcare systems struggle to operate on reduced resources and personnel, the access to quality healthcare for patients is being jeopardized. This decrease in quality of service (QoS) is a serious concern for both primary and public healthcare outcomes, since receiving sound medical advice is key to maintaining trust in healthcare systems and to curb medical misinformation [De Figueiredo et al., 2020, Xiang and Lehmann, 2021]. Moreover, with Canada set to experience an unprecedented increase in its senior population, who are set to comprise 25% of its people by 2040 [Wong, 2014], there is a dire need to dramatically increase the healthcare capabilities to ensure quality healthcare for all. However, Canada’s shrinking nursing population, which was suffering from acute shortages even before the pandemic, has only been exacerbated by the pandemic, further stressing the healthcare systems [CFNU, 2022]. As a result, there is an urgent need to build back the healthcare workforce and develop tools to assist them in delivering quality healthcare for all.

Artificial Intelligence (AI)-powered machine learning (ML) models can exactly help achieve this goal by assisting in a wide range of tasks such as disease diagnosis, forecasting need for surgery (surgical candidacy) post-operative

---

[*] Invited paper at the Cybersecurity, Privacy, and Artificial Intelligence in Health Data: Advancements and Challenges Conference, 2023 For more information about the papers discussed in the work please visit https://sirisharambhatla.com

outcomes, waitlist outcomes for transplantation, donor-recipient matching, skilled medical professional training, and many other application domains. It is, however, important to note that AI/ML models rely on historical data, and that means that these models can reinforce the very systemic biases the healthcare systems are trying to mitigate. In other words, if a particular bias impacted the outcome for a patients’ belonging to a subgroup in the past, the ML models which learn from such data can end up learning a correlation between patient’s protected attributes and the outcome; in effect, propagating any inequities and policies in historical data. We address these aspects in Sec. 4. We also cover the trustworthiness from a recent shift towards large language models (LLMs) for healthcare in Sec. 5.

The reckoning across different organizations in Canada regarding the racial, gender, ethnicity, and other biases must also be framed in the context of healthcare. For instance, systemic racism in the healthcare system significantly impacts Indigenous peoples’ health outcomes, which has been found to plague the entire Canadian healthcare system [Gunn, 2016, Blake, 2018]; similar biases are known to exist across sex and gender identities [Kim et al., 2022]. Furthermore, the issue of privacy and responsible use of healthcare data has also emerged as an important consideration, especially in the context of increased cybersecurity attacks. As a result, any application of AI/ML modeling must consider these factors while designing models for healthcare applications.

While increased AI/ML awareness and education in the medical community has driven the initial wave of collaborations, these often lead to an idealized view of both healthcare and ML practitioners. Specifically, the work of a) identifying a particular problem, and b) data corresponding to it is usually undertaken by the healthcare expert, while the ML practitioner c) understands the task at hand, d) processes the data, and e) builds models to accomplish a particular prediction task. This requires both parties to essentially speak the same technical language and to some extent understand each others’ techniques – which works in an academic setting, but in a number of real-world scenarios is extremely hard to achieve. A lack of understanding about ML capabilities can often prevent us from identifying opportunities to develop solutions that can positively impact healthcare outcomes in the real-world healthcare systems, which underscores the need for data literacy. This also includes handling of sensitive data in a secured way, and developing models which can learn in a privacy preserving fashion, as well as establishing synthetic healthcare data networks.

Another important issue that stems from this knowledge gap is that while healthcare systems may be collecting data, they may lack the capability to track and log their data for ML modeling. This is especially challenging given the sheer volumes of patients they might be handling. This data organization and maintenance takes a significant amount of time and financial investment; however, it may be unclear how to effectively triage and access this data. This, and other data-related factors, impact how quickly ML models can be developed and deployed in real-world systems. For Canada to emerge as a leader in healthcare, there is an urgent need to create a blueprint for healthcare organizations to effectively harness the potential of their data. In this paper, we aim to lay down the foundations of this strategy which we refer to as AI Readiness. We discuss specific recommendations in Sec. 6.

Transparency and interpretability of the ML models are also important considerations. Adoption of ML models in critical healthcare decision making is contingent on the medical professionals being comfortable with relying on model predictions. While explainability of deep learning-based models offers some solutions revealing potential reliance on sensitive attributes (gender, race, etc.), they may prioritize human understanding over actual model decision-making. Given that humans are prone to finding patterns in seemingly unrelated phenomenon, explainability which focuses on humans to make sense of model decisions raises additional reliability concerns [Rudin, 2019]. On the other hand, explainability which reveals reliance on Furthermore, rigorous testing of blackbox models and monitoring performance metrics while inherently opaque raises another aspect of trust. We cover this aspect via the example of explainability for burn surgical candidacy in Sec. 2.2 where we leverage our work on explainability to inform clinicians about the model decision making [Tsang et al., 2020].

2 AI for Forecasting Patient Outcomes in Primary Healthcare Applications

In this section we outline ways in which AI-powered machine learning models can revolutionize primary healthcare by forecasting patient outcomes. We consider specific examples from our previous and current work on a) forecasting the trajectory of patients waitlisted for liver transplantation, and b) predicting the burn surgical candidacy for plastic surgery. These provide a comprehensive view of data and models used to provide key insights to clinicians to aid their decision making.

2.1 Liver Transplantation

Background. Clinicians often rely on their experience to manage patients waitlisted for liver transplantation (LT) and prioritize them based on their functional status. The Model for End-Stage Liver Disease (MELD) and MELD-Na scores (number between 6 and 40) are an important factor in assessing a patient’s candidacy for transplantation

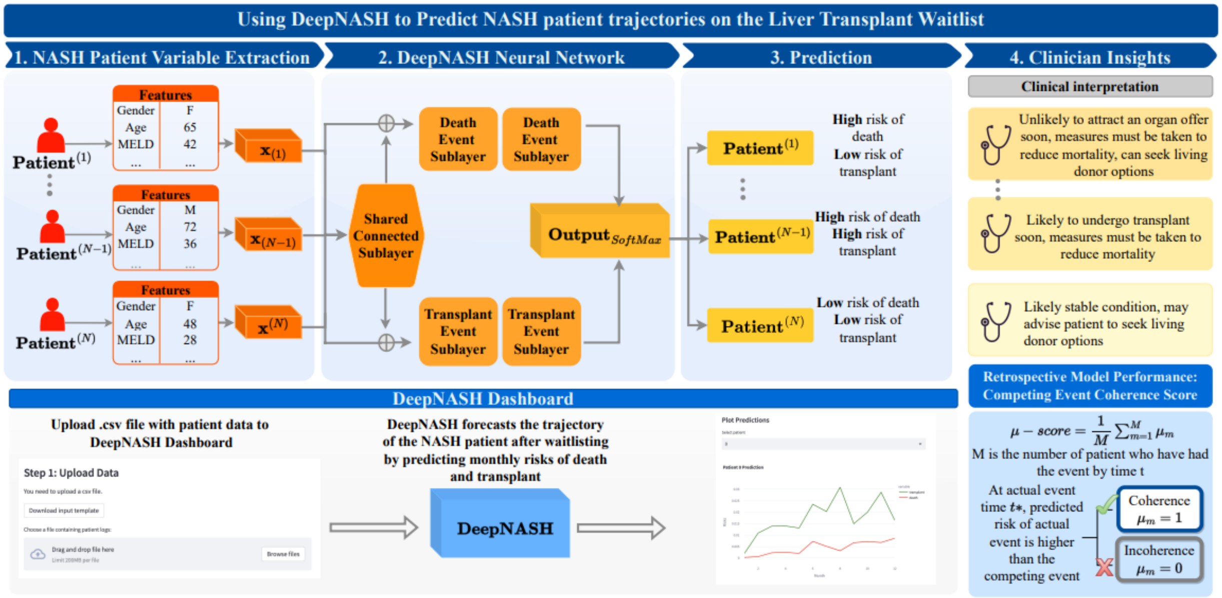


Figure 1: Using DeepNASH to forecast NASH patient trajectories on the waitlist; Figures reproduced from following works [Punchhi et al., 2022a,b,c, 2023].

[Wiesner et al., 2001], where a higher MELD results in a greater priority for organ offers. However, MELD and MELD-Na's predictive ability is relatively low for Non-alcoholic steatohepatitis (NASH) cirrhosis patients [Sheka et al., 2020], [Godfrey et al., 2019], which is rapidly becoming the leading indication for LT [Sheka et al., 2020], [Kwong et al., 2020], [Karnam et al., 2022]. Moreover, since they also experience other comorbidities, and deterioration within two-months of waitlisting, clinicians are often hesitant to consider them for LT [Kwong et al., 2020], [Haldar et al., 2019], [Vanwagner et al., 2012], [Charlton et al., 2011], [Henson et al., 2020]. This section covers our work from [Punchhi et al., 2022a,b,c, 2023].

Understanding clinical events via data-driven machine learning (ML) models. Given this poor performance in modeling NASH patient outcomes, coupled with the fact that NASH is set to become the top indication for LT, there is a need to build data-driven models that can tackle this task to aid clinicians in their decision making. To this end, time-to-event models based on survival analysis can be used to predict the time until a specific event of interest (such as death or transplant) occurs. The Cox Proportional Hazards (CoxPH) model are a popular choice for this task but are limited since these cannot capture complex covariate interactions and dynamics [Spooner et al., 2020]. Other models that aim to capture complex dynamics include random survival forests (RSF), a random forests-based ML model, and DeepSurv, a neural network-based model [Jain et al., 2021], [Nagai et al., 2022], [van Geloven et al., 2022], [Ishwaran et al., 2008], [Katzman et al., 2018]. However, in the real-world, patients can often experience multiple outcomes, which necessitates the use of modeling under concurrent competing risks. Recent deep learning-based works – DeepHIT – build an assumption-free model for tackling such a case [Lee et al., 2018].

Key Contributions. Our work developed a deep learning-based competing risk model to understand and compare the risk of death and transplant for waitlisted NASH patients in the Scientific Registry of Transplant Recipients (SRTR) data in United States [Punchhi et al., 2022a,b,c, 2023]. With its ability to model nonlinear relationships between covariates and outcomes in a time-varying manner, our model – DeepNASH – captures the complex interaction between patient covariates and to allow for patient-specific variations in the trajectory. An overview of the proposed DeepNASH workflow is shown in **Fig. 1.**

Our critical analysis reveals that the two popular metrics: the **Concordance index (C-index)** and the **Brier score**, may not be appropriate to analysis to measure the model's ability to predict a certain type of event, and to be compared across multiple events. To address these challenges, we develop a new metric – **competing event coherence (CEC) score**, which assesses patient-level model performance by matching model event prediction which the actual event for each patient. Therefore, giving a deeper insight into patient-level predictions.

Further, we interpret DeepNASH predictions by identifying features associated with death and transplant on the LT waitlist [Lee et al., 2018], and developing a DeepNASH dashboard: https://deepnash.streamlit.app/ that allows clinicians to upload patient features and predict waitlist trajectory. This model can therefore be used to understand patient trajectories and assist the clinical decision making to manage this specialized patient group.

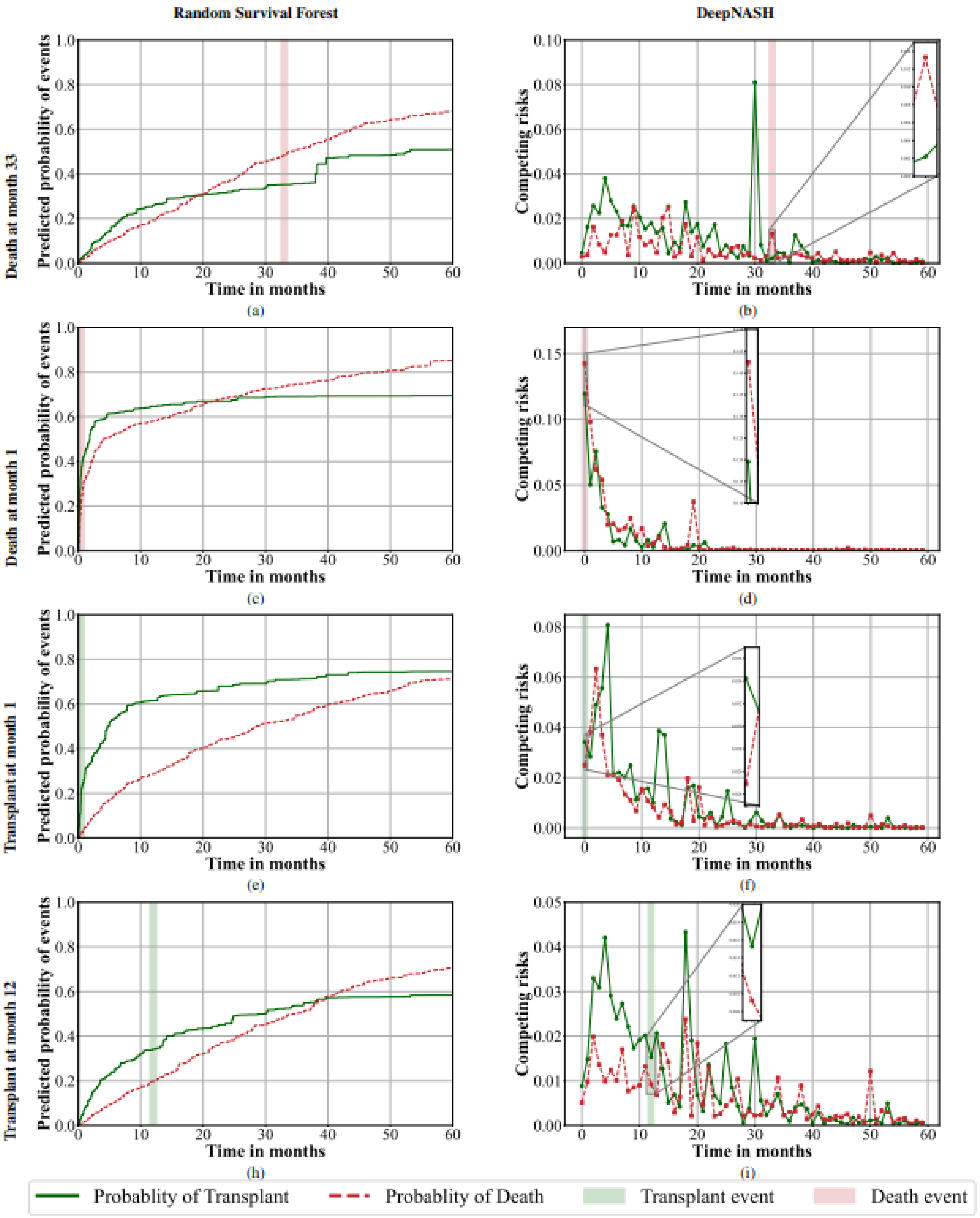


Figure 2: Prediction of waitlist death (red) and transplant (green) using RSF (left) and DeepNASH (right) for four sample patients. Actual event type and time indicated by shaded boxes. Figures reproduced from following works [Punchhi et al., 2022a,b,c, 2023]

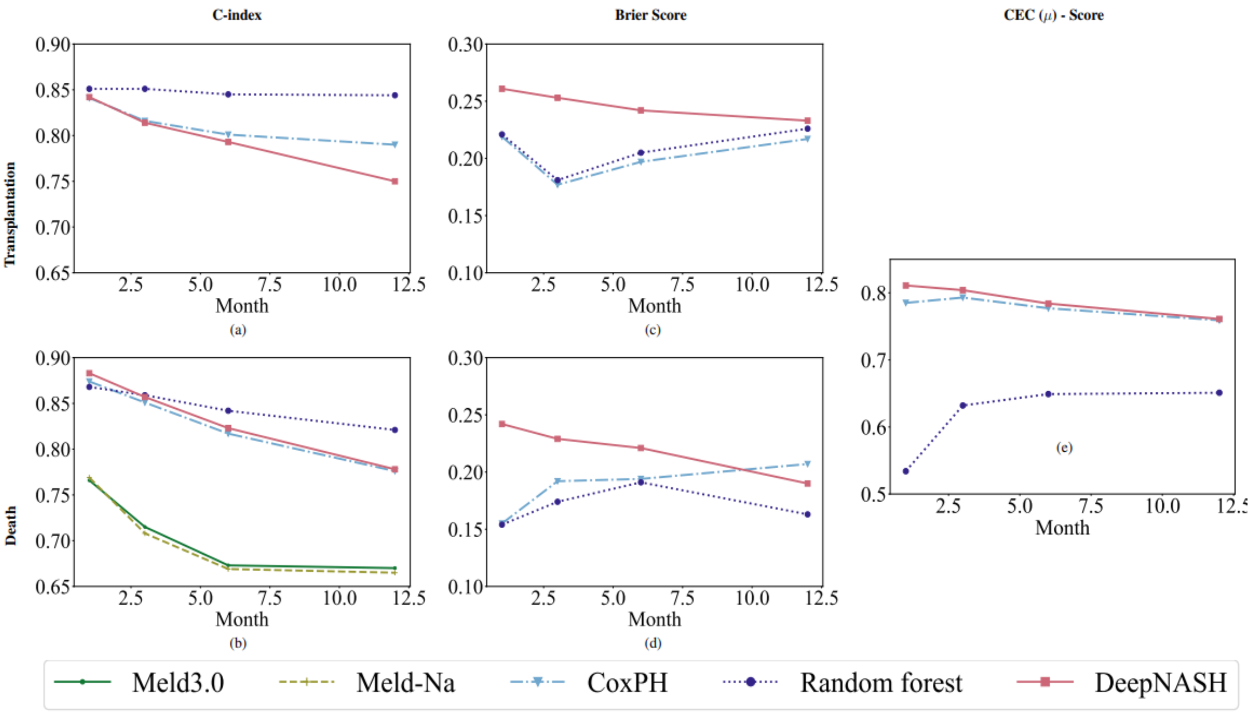
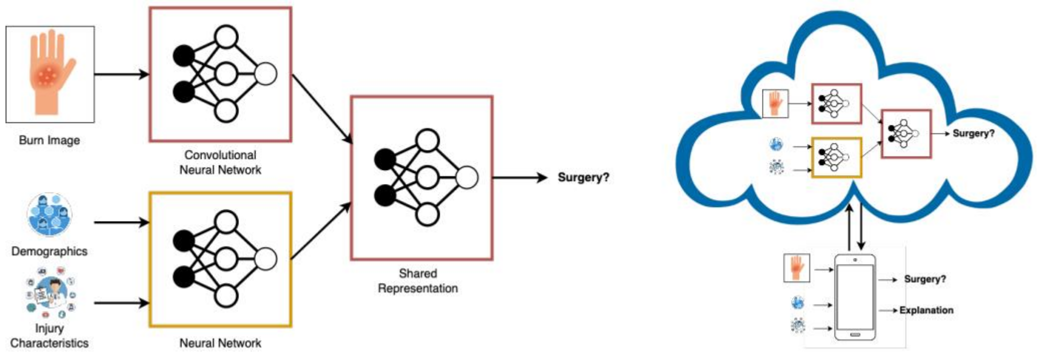


Figure 3: Comparison of C-index (a,b) and Brier (c,d) scores of all models to predict transplant (a,c) and death (b,d) for up to 12 months on the waitlist. Panel (b) also shows MELD-Na and MELD 3.0 compared with all models for death event (only applies to this event). In panel (e) we show the CEC score calculated for the RSF, CoxPH and DeepNASH models. As opposed to C-index and Brier score, CEC is suitable for competing risk analysis since it can compare the relative risk between different types of events. Figures reproduced from Punchhi et al. [2022a,b,c, 2023]

2.2 Burn Surgical Candidacy

Background. Visual inspection is the primary way to evaluate if a certain burn wound requires surgery or not (Burn Surgical Candidacy). This decision making involves analysis of burn depth and other patient characteristics. This process is extremely subjective, and in addition to the proclivity to overestimate burn depth, the accuracy varies widely based on surgeon experience; so much so that experienced burn surgeons have an accuracy of 64-76% while novice surgeons have 50% accuracy [Acha et al., 2013], [Cirillo et al., 2019], [Chauhan and Goyal, 2020], [Martínez-Jiménez et al., 2018], [Li et al., 2015], [Pham et al., 2019]. At a time where health systems across the world are facing worker shortages, these unnecessary surgical interventions are counterproductive for patients as well as the hospitals [Pham et al., 2019], [Acha et al., 2013].



(a) (b)

Figure 4: Overview of DL4Burn – a Multimodal Deep learning approach for burn surgical candidacy prediction.
Panel a) shows the multimodal deep learning model specifics. In panel (b) we illustrate the cloud-based workflow of the iOS application being used in the field. Figures reproduced from [Rambhatla* et al., [2021].

Table 1: Comparing model performance for burn wound surgical candidacy prediction in terms of area under the curve (AUC) and Accuracy metrics.

Method	AUC	Accuracy
Surgeons	–	0.50–0.76
HOG-SVM	0.595	0.587
ORB-SVM	0.548	0.563
CNN	0.637	0.655
ResNet-50-based	0.854	0.810
DL4Burn	**0.975**	**0.938**

Table 2: Performance of ResNet-50 and DL4Burn (Neural Network (NN) based methods) across 5-fold cross-validation in terms of area under the curve (AUC) and Accuracy metrics.

NN Method	AUC		Accuracy	
	Avg.	Best	Avg.	Best
Surgeons	–		0.50–0.76	
ResNet-50-based	0.833	0.854	0.784	0.810
DL4Burn	**0.907**	**0.975**	**0.858**	**0.938**

On the other hand, missing burn wounds which require surgical intervention are at higher risk of slow healing, hypertrophic scarring, and contractures [Pham et al., 2019], [Acha et al., 2013].

Data-driven burn wound analysis via deep learning. Given this variability in assessing burn surgical candidacy and its adverse impact on the patient outcomes, there is an urgent need to develop automated data-driven models that can assist clinical decision making. The existing body of work utilizes machine learning models to predict burn depth, which in turn is used to determine surgical candidacy. This approach includes works such as [Serrano et al., 2015], [Yadav et al., 2019], that employ support vector machine (SVM) trained on features extracted from burn images. On the other hand, works such as [Wang et al., 2020] aim to classify burn wounds based on their healing time.

While these studies report promising results and accuracies of 79%-100%, they focus on relatively small dataset, and predict burn depth and healing times to inform the downstream burn surgical candidacy. In fact, burn depth is one such consideration, and other patient factors, as well as surgical risks due to comorbidities, also play a major role [Browning and Cindass, 2019]. For example, patient age directly impacts skin thickness and therefore thermal protection, and pre-existing conditions influence the injury pathophysiology [Daugherty, 2017]. Furthermore, factors such as total burn surface area (%TBSA) and sepsis can also effect inflammatory responses and thermoregulation, ultimately informing surgical decisions [Browning and Cindass, 2019], [Daugherty, 2017], often requiring multiple initial assessments to monitor wound healing [Devgan et al., 2006]. As a result, an accurate data-driven approach utilizing contemporary deep learning models to leverage patient characteristics along with burn wound images (multiple modalities) can greatly alleviate these pain-points and ultimately improve patient outcomes.

Multimodal Deep Learning for Burn Wound Evaluation. Considering multiple clinical indicators and information sources to mimic disease diagnosis in the real world can overcome the aforementioned challenges [Ngiam et al., 2011]. Referred to as multimodal models, these have shown promising results in other medical domains to analyze cervical dysplasia [Song et al., 2014], [Xu et al., 2016], Alzheimer's [Lee et al., 2019], [Zhang et al., 2019], and skin lesions [Gessert et al., 2020]. Motivated from these, we developed a multimodal deep learning architecture for burn evaluation – DL4Burn – shown in **Fig. 4 (a)**, which utilizes two neural networks to build image and the tabular data representations separately [Rambhatla* et al., 2021]. Specifically, we fuse representations from a ResNet-50-based deep convolutional neural network (CNN) to process the images and a multi-layer perceptron for the patient characteristics, to arrive at a probability indicating whether a surgical intervention would ultimately be recommended or not (in hindsight).

We learn this model on wound images as well as individualized patient characteristics collected at the Southern California Regional Burn Center at LAC+USC on patients admitted between January 2015 and December 2016. Further, we also use state-of-the-art interpretability techniques to analyze the model's decision-making process to assist clinicians [Tsang et al., 2020]. Complementary to ML modeling, we also develop an iOS mobile application to assist the clinicians in the hospital for data acquisition, triage, and predictions.

We compared the area under the receiver operating characteristic curve (AUC) and accuracy performance of DL4Burn with support vector machine (SVM) and neural network-based baselines, including Histogram of Oriented Gradients (HOG) features [Dalal and Triggs, 2005], SVM trained on Oriented FAST and Rotated BRIEF (ORB) descriptors [Rublee et al., 2011], and a 2-layer CNN as shown in **Table 1**. We also perform further analysis of the multi-modal approach as compared to a ResNet-50 neural network baselines in **Table 2**; results from [Rambhatla* et al., 2021]. These results clearly indicate the advantages of using a deep learning-based multi-modal approach to model surgical decision making. **Fig. 5** shows the explanations generated by Archipelago [Tsang et al., 2020] to understand DL4Burn's decision making. Refer to [Rambhatla* et al., 2021] for details.

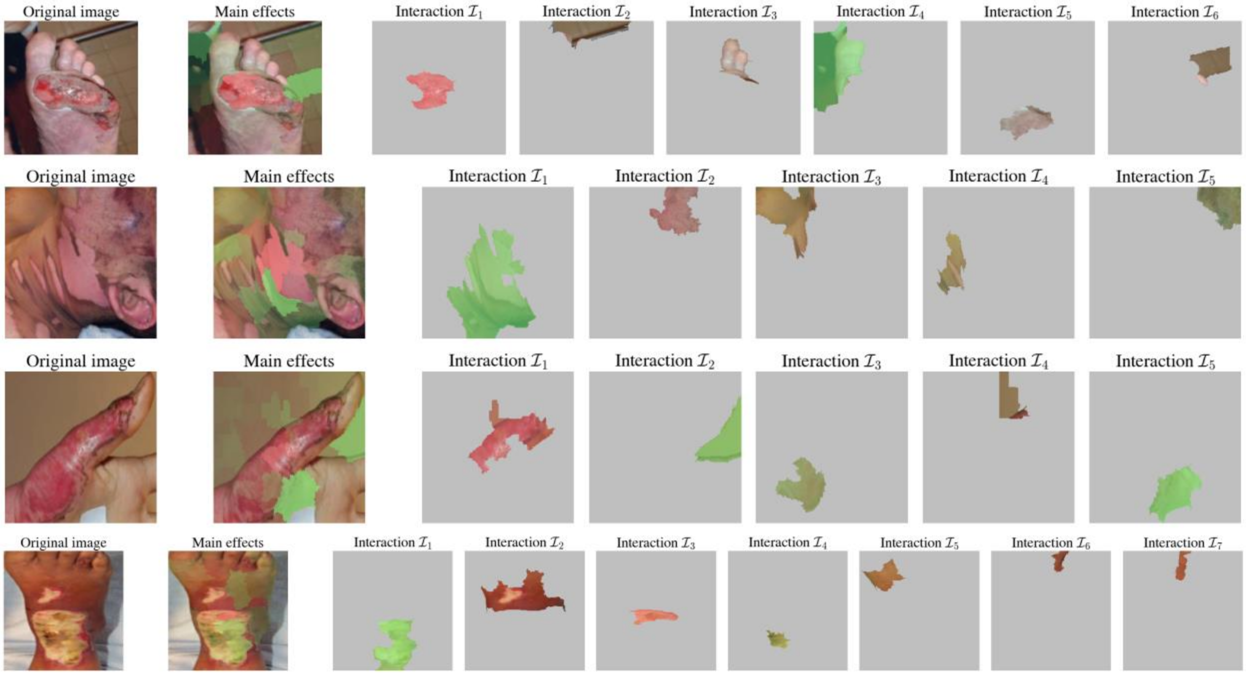


Figure 5: Sample explanations of the image classification results using Archipelago. Here, the colored patches show if they support (green), are against (red), or are neutral (no color), for a given classification result.

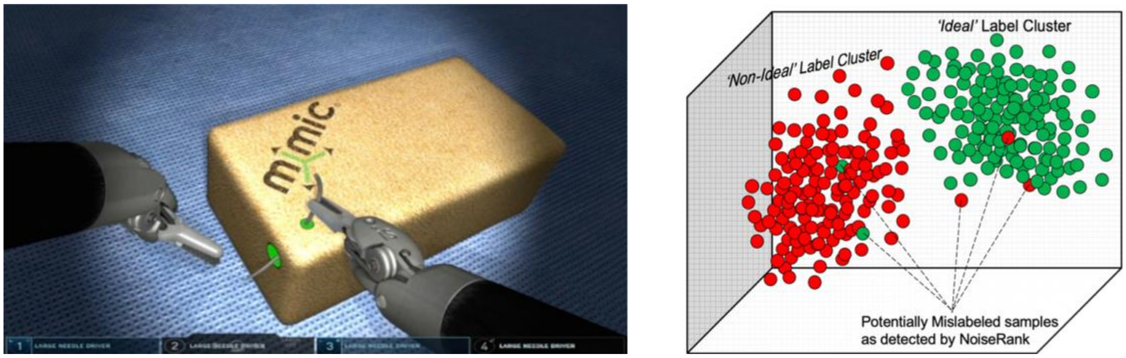
Conclusions. Multiple factors and their complex interactions impact burn surgical candidacy decisions, which are often analyzed on criterion which are highly subjective. Data-driven models can assist surgeons to draw insights from historical cases to improve patient outcomes. While our results are extremely promising, we did observe slight overfitting which needs to be countered with adding more data (enabled by our iOS app). In this study we use the final assessment, i.e., if they would ultimately require surgery or not instead of the initial assessment. It would be valuable to record initial assessments as a way to analyze surgeon decision making. Another line of work could also study the counterfactuals and effect of interventions. The proposed DL4Burn model and mobile application leverage the expert analysis along with burn images to build a multimodal deep learning-based tool to improve decision making and patient care.

3 AI for Medical Training and Skill Assessment

Powered by augmented reality and virtual reality (AR/VR) infrastructure, AI models can also play a critical role in training the next generation of surgeons by assisting with their a) automated skill assessment while also providing b) fine-grained feedback to improve their skills. In this exposition we will see how automated skill assessment can help surgeons practice their skills, and ultimately improve outcomes for patients undergoing radical prostatectomy. This section draws from our following works: [Hung et al., 2021a,b], [Chen et al., 2022], refer to these for additional details.

Background. Leveraging data collected during robot-assisted surgery, AI-powered models have had significant success in predicting surgeon experience and surgical action recognition using the kinematics data corresponding to the surgical actions [Lane, 2018], [Hung et al., 2019a, 2018]. These surgical motions have also been used to predict the subsequent patient outcomes in the context of radical prostatectomy [Hung et al., 2019b]. While manual feedback is extremely important and is considered the gold standard during surgical training [Goh et al., 2012], [Martin et al., 1997], [Raza et al., 2015], it is also extremely slow and time-consuming, prompting the need to develop automated ways to evaluate surgical skills. However, the current automated metrics are somewhat crude since they only focus on metrics such as time-to-completion [Birkmeyer et al., 2013], [Trehan et al., 2015], [Lendvay et al., 2015]. As a result, there is a need to develop AI/ML models that can automate surgical skill assessment while providing meaningful feedback to surgeons to aid their training at smaller timescales.

Neural Networks-based Automated Surgical Skill Evaluation. In this work, we aimed to train a neural network model on kinematic time-series data to directly predict the expert evaluation. To be able to provide fine-grained feedback, we divided the suturing step into four technical skill assessment tasks – **Needle positioning**, **Needle entry angle**, **Needle driving**, and **Needle withdrawal** – and collected the synchronized instrument kinematic data at this sub-stitch level, and the corresponding labels generated by human graders. Specifically, the urologic surgeons performed standard Basic Suture Sponge exercise thrice on the Mimic™ FlexVR robotic simulator **Fig.6(a)**.



(a) (b)

Figure 6: Automated Surgical Skill Evaluation. Panel (a) shows a screenshot of basic suture sponge exercise on mimic FlexVR, while panel (b) shows how NoiseRank [Sharma et al., 2020] works: NoiseRank computes the probability of a sample being potentially mislabeled given its neighbors. Figures reproduced from [Hung et al., 2022].

These suturing skills are also relevant across different specialities, and surgeon skills at this granularity have also been identified to be key markers of expertise, impacting patient outcomes [Trinh et al., 2022], [Chen et al., 2021a]. For each step, we used this data to train a two-layer LSTM-based (long short-term memory) model to classify whether the step was performed correctly or not. For additional details refer to [Hung et al., 2022]. Due to the inherent unequal distribution of the positives (Ideal) and negatives (Non-Ideal) labels, we use the best test AUC performance averaged across three Monte-Carlo runs.

Battling Uncertainty in Real-World Data. While we sought to utilize human feedback to ascertain the correctness of a suturing step, we encountered a very practical dilemma – there seemed to be a subjective bias in the human graders which led to uncertainty in the subsequent label. We addressed this in two ways. First, we developed a systematic consensus procedure to arrive at final label (refer to [Hung et al., 2022] for details). Next, we adapted the label uncertainty detector – NoiseRank [Sharma et al., 2020] – to identify and rank the final labels (after the consensus process) according to the degree of their uncertainty. NoiseRank uses the kinematic tracing representation and the final labels to estimate the probability that a given sample is mislabeled using its neighbors **Fig.6(b)**. We then retrain the model after removing different percentages of potentially mislabeled data. Here we developed a NoiseRank variant that can handle imbalanced data. Results of the analysis are shown in Table 3. We observe that different suturing steps benefit from different levels of (projected mislabeled) data removed.

Conclusions. Automating surgical skill assessment requires faithfully reproducing expert opinion about each sub-step of a suturing step using ML models, which require supervision in the form of labels obtained from human experts. However, we found that even at a fine-grained scale, the assessment by human graders can be subjective which ultimately leads to uncertainty in the labels and consequently mislabeling in the final assessment labels after consensus. Since the final labels may contain mislabeled data, we used identified potential mislabeling and systematically studied the impact of removing them and pave a way to battle this inherent uncertainty in practical healthcare applications.

We observed that the performance of the machine learning model varied by skill domain. While predictions for the needle positioning skill were mediocre, which can be attributed to the fact that there is no tracking of the needle-to-instrument interaction, the results for other more important skill domains were all greater than 0.70, which underscores the viability of AI-based automated skill assessment. In addition, when it comes to identification of potential mislabeling, we had to discard a variable number of samples to improve model performance. Understandably, the mislabeling varies with the skill domain, but more importantly the proverbial ‘throwing out of the baby with the bath water’ can occur if labels which define the decision boundary are also thrown away. In other words, our results indicate that it is important to preserve the nuanced information about each skill domain. As such, this analysis showcases the potential and challenges involved in incorporating AI-based assistive technologies to help critical applications and healthcare skill development.

Table 3: Model performance in automating suturing technical skills assessment, Performance (AUC) Mean ± SD

Technical Skill Domain	**Percent of Projected Mislabeled Data Removed**					
	Full Data	Top 15%	Top 25%	Top 50%	Top 75%	100%
Needle positioning	0.532 ± 0.030	0.498 ± 0.063	0.518 ± 0.111	0.538 ± 0.136	0.517 ± 0.054	**0.551 ± 0.035**
Needle entry angle	0.705 ± 0.026	0.715 ± 0.055	0.733 ± 0.061	**0.766 ± 0.069**	0.751 ± 0.083	0.719 ± 0.097
Needle driving	0.704 ± 0.025	0.642 ± 0.078	0.671 ± 0.030	0.650 ± 0.096	**0.715 ± 0.054**	0.699 ± 0.022
Needle withdrawal	0.698 ± 0.072	**0.761 ± 0.038**	0.722 ± 0.086	0.735 ± 0.032	0.741 ± 0.099	0.714 ± 0.109

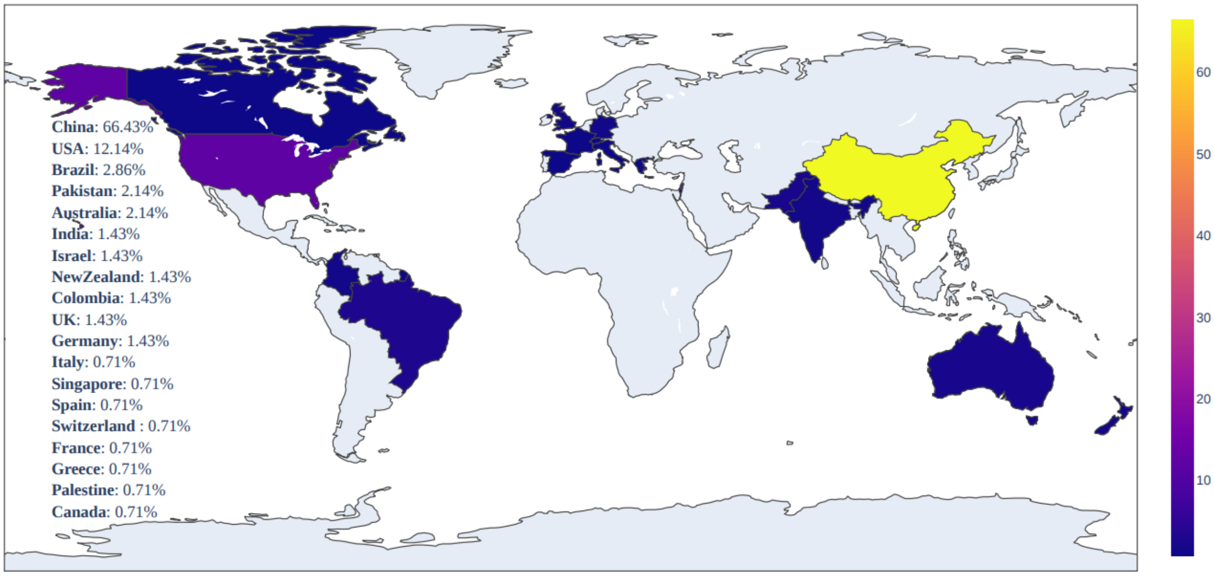


Figure 7: Distribution of healthcare prediction models using MIMIC data across the world from 2018 to 2022. The % value indicates the amount of research carried out by each respective country using MIMIC data. The color meter reflects the level of research activity. Figure reproduced from [Murugan et al., 2023].

4 Trustworthy and Fair AI modeling for healthcare

Building trustworthy and fair machine learning models critically relies on the properties of underlying data. Since the healthcare data collected is representative of the demographic group it serves, any biases in the training data often tend to get reflected in the trained models. Therefore, it is critical to consider the fairness properties during training. This is especially important in the healthcare settings where the outcomes depend on individual patient characteristics, as the existing biases in the dataset can adversely impact the predictions for different patient groups.

Background. While algorithmic bias in ML models, where model performance differs by demographic groups is an active area of research [Zou and Schiebinger, 2018], [Chen et al., 2018, 2019], [Rajkomar et al., 2019], [Obermeyer et al., 2019], [Rajkomar et al., 2018], [Chen et al., 2021b], its counterpart – data bias - has not been extensively studied. Data bias can lead to subsequent algorithmic bias and is caused due to the inherent lack of data for demographic groups and can encode any systemic bias which lead to different outcomes. As a result, to understand “what are our models good for?” should be accompanied by “who does our model represent?”. In the context of healthcare this becomes especially important since popular clinical research databases (CRD) are being used across the globe for downstream healthcare tasks. However, devoid of the context, such as policy-level factors like insurance and demographics, these models may lead to unintended consequences. This section draws from our work [Murugan et al., 2023].

By analyzing Medical Information Mart for Intensive Care (MIMIC) [Goldberger et al., 2000] – one of the most popular CRD [Johnson et al., 2016], [Japkowicz, 2000], [Johnson et al., 2023] – and half a decade of healthcare ML models, we demonstrate a noteworthy association between ethnic demographics, insurance status, and outcomes. To better document the data characteristics, we develop a comprehensive datasheet [Gebru et al., 2021] for MIMIC IV v2.0 CRD to aid in the understanding of this massive database.

Where is MIMIC IV v2.0 being used, and how? The MIMIC IV v2.0 data originates from Beth Israel Deaconess Medical Centre in Boston, MA, U.S.A. As a result, it represents, and is impacted by, the local demographics and policies. Interestingly, our systematic review revealed that approximately 67% of the publications using this CRD originated from research publications in China, with significant contributions from USA, Brazil, Pakistan, Australia, and India as shown in **Figure 7.** Moreover, our analysis reveals that almost 80% of the studies either did not consider performance across ethnic demographics and/or left these features out without explanation. This raises two concerns a) the model may have varying performance for different demographic groups, and b) models learned for this data for a new hospital may not be reliable due to differences in demographic make-up.

Association between demographic variables and the predicted outcomes. We analyzed the association between the ethnic/racial features, insurance status, and the outcomes such as readmission rates, and mortality prediction for sepsis, heart failure, and chronic kidney disease. Our analysis revealed a significant association between these features.

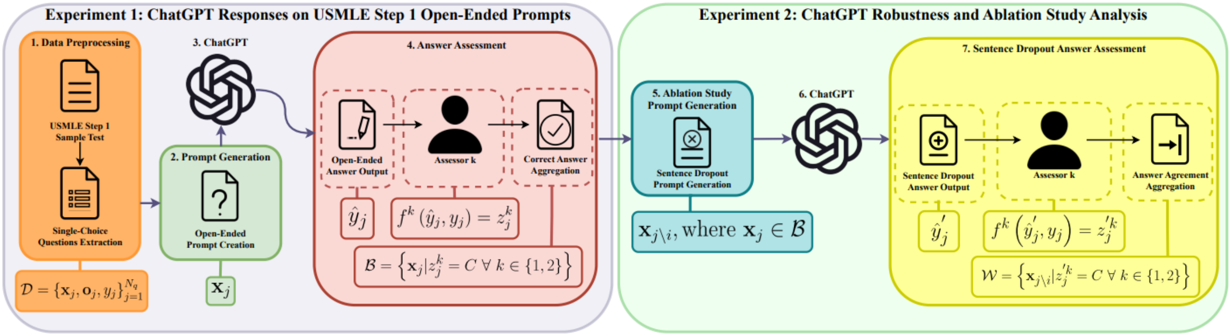


Figure 8: Summary of Methodology. 1) We select 94 USMLE Step 1 [USMLE, 2022] questions consisting of multiple-choice, single-answer questions. 2) These questions are used to generate open-ended prompts. 3) Open-ended prompts are fed into ChatGPT API [OpenAI, 2023]. 4) ChatGPT’s answers are shown to **k** independent non-expert assessors, who categorize them as “Correct” (**C**), “Partially Correct” (**P C**), “Incorrect” (**I**), or “Ambiguous” (**A**). 5) Data corresponding to the case where both assessors marked the answer as “Correct” are collected as ***B*** and prepared for sensitive analysis via iterative sentence dropout. 6) The prompts from 5) are fed to ChatGPT. 7) The assessors categorize ChatGPT’s answers again, and we collected those for which both assessors categorized the answer as “Correct.” (dataset ***W***). Figures reproduced from [Barnard et al., 2023].

Specifically, for readmission, we observe an association between Black-Medicare and Black-Other insurance subgroup, and Asian-Medicaid patients, exhibiting higher readmission rates regardless of their ailments (χ^2 = 141.89, p < 0.001) and the ANOVA (F = 3.62, p = 0.019). For sepsis mortality, other ethnic groups and Asian consistently had higher proportions of deaths, regardless of their insurance status (χ^2 = 975.185, p < 0.001). For heart failure mortality Other ethnic group consistently exhibited higher mortality rates, irrespective of their insurance (χ^2 = 105.107, p < 0.001). For CKD mortality, Other-ethnic groups consistently exhibit higher proportions of deaths, regardless of their insurance type. Medicare-insured Blacks experience the second-highest mortality rates (χ^2 = 106.578, p < 0.001) and ANOVA (F = 2.203, p = 0.098). Following this analysis, we develop a comprehensive datasheet for this database.

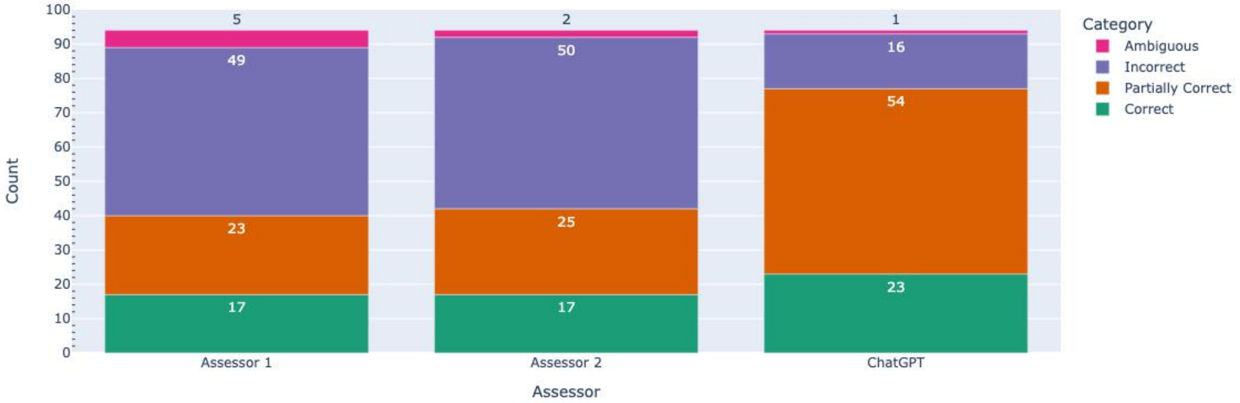
Conclusion. Demographic imbalance-induced data bias is inevitable, since it originates from the patient group a hospital serves. However, when using data derived from a health system in other contexts (say for transfer learning), it is extremely important to understand the similarities and differences. Our analysis reveals that the healthcare outcomes can depend on the ethnicity and insurance status, which is especially concerning if reinforced by a machine learning model. On the other hand, our analysis can also be used to understand the systemic biases and inefficiencies in a hospital and calls for data-aware modeling for development of generalizable and trustworthy models.

5 Self-diagnosis and a new wave of misinformation with large language models (LLMs)

Background. Recent popularity of large language models (LLMs) and their perceived human-like capability had led to a number of industries to explore LLM applications in their domains [Taecharungroj, 2023], and users placing inordinate amount of trust in them [Salah et al., 2023]. Since web-search based self-diagnosis is already very common and very valuable in a number of cases [Goyder et al., 2009, White and Horvitz, 2009], people turning to LLMs for this application is inevitable. In the backdrop of healthcare worker shortages [Kuehn, 2022, Boniol et al., 2022, Michel and Ecarnot, 2020, Turale and Nantsupawat, 2021], both public and private entities are looking to LLMs for improving healthcare workflows [Horesh, 2023, Lee, 2023, Sedaghat, 2023, Kung et al., 2023, Iftikhar et al., 2023, Primack, 2023].

More recently, researchers have started analyzing the capability of ChatGPT – a popular LLM – to answer medical exams [Holmes et al., 2023, Kung et al., 2023, Gilson et al., 2022, Sallam et al., 2023, Sallam, 2023]. Recent works such as [Kung et al., 2023] reported scores of 60% on United States Medical Licensing Exam [USMLE, 2022], arguing use of LLMs in medical student training, thus creating an inaccurate sense of ChatGPT’s capabilities. Specifically, these works required ChatGPT to pick-out an option from a list, which is not indicative of real-world clinical decision making. Moreover, they also used a less granular scale to assess ChatGPT responses, ignoring the cases where the responses are ambiguous [Kung et al., 2023]. Motivated from these analyses, we aimed to evaluate the performance of ChatGPT in a realistic setting on a granular scale from the perspective of a non-expert hoping to use it for self-diagnosis. This section covers our work from [Barnard et al., 2023].

Analyzing ChatGPT’s response and sensitivity to open-ended medical questions. To critically analyze ChatGPT’s capabilities in real-world, we develop a testing methodology to evaluate its performance on open-ended questioning. Specifically, we base our analysis on a subset of single-answer questions from the United States Medical Licensing



Panel (a)

Panel (b)

Figure 9: Experimental results. Panel (a) shows the visualization of non-expert assessor and ChatGPT self-assessment categorization. Panel (b) shows a visualization of non-expert assessor and ChatGPT self-assessment categorization for ablation study. Note that this experiment analyzes the correct responses shown in panel (a), considering the impact of sentence drop-out on each question. Figures reproduced from [Barnard et al., 2023].

Examination (USMLE) [USMLE, 2022], and perform a two-stage evaluation (which includes sensitivity analysis) with two non-expert assessors, also asking ChatGPT to evaluate its responses as shown in **Fig 8.** Corresponding to the two experiments in **Fig 8, Fig 9 (a)** and **(b)**, respectively, show the results of our analysis, respectively. This reveals that not only, a) these models perform worse than previously reported, but they also exhibit b) overconfidence when stating incorrect recommendations and are also less confident when stating the correct answer.

Discussion. Our results indicate that ChatGPT's ability to answer exams may not translate to real-world understanding; in fact, we see that it is only able to answer a small subset correctly (17/94); see **Fig 9(a)**. Since it is unclear what resources ChatGPT is using, it is also difficult to understand which type of questions it is more likely to fail on. Next, our granular scale which clearly distinguishes "Correct" and "Partially Correct," shows that the majority of ChatGPT responses are incorrect, **Fig. 9 (a)**. This is understandable since ChatGPT cannot reason; however, this is still extremely concerning, given its popularity.

Qualitatively, we also observed that ChatGPT fails to communicate uncertainty with generally using qualifiers such as "most likely" or "likely", aligning with concerns raised by [De Angelis et al., 2023] regarding nuanced ways in which LLMs can spread misinformation. ChatGPT also exhibited overt optimism in two ways. In the first experiment, when asked to evaluate its own responses, it marked 57.45% of the answers as "Partially Correct". This trend continues in the second experiment where we analyze the subset of answers which were answered correctly by ChatGPT. Here, we see a reversal. ChatGPT only marked 50% of its answers as "Correct". Overall, we see that ChatGPT is more optimistic when it is more likely to be incorrect, and cautious when it is likely to be correct. This is the inverse of the tendency that we would require of a healthcare provider. Given this peculiar behavior, at this time LLMs can significantly increase the risk of spreading medical misinformation.

6 Recommendations and Roadmap

With populations aging and a shortage of healthcare workers, Canada stands at a very critical juncture, where scaling the reach of the existing healthcare systems is of utmost importance. While translating the advances in data-driven healthcare research to real-world is an urgent need, there are a number of factors and infrastructural requirements that need to be met to unleash the power of AI/ML-based healthcare solutions. Based on our experience, working with hospitals and facing roadblocks along the way, we now outline some of the challenges that are faced by our healthcare systems today, and provide recommendations to ensure we are equipping our healthcare systems with the necessary tools to help them transition and make Canadian healthcare systems "AI Ready".

Our recommendations also complement those put forward by the 2022 Expert Advisory Report on "The pan-Canadian Health Data Strategy: Charting a Path toward Ambition" [Goel et al., 2022]. The report shows a requirement for concerted efforts from different organizations. Specifically, it demonstrates that the challenges that we face are not due to a lack of technical expertise but due to the fact that Canadian health data is highly fragmented, and in recent times this has foreshadowed healthcare efforts such as Canada's response to COVID-19.

- **Need to create a strong Canadian healthcare data network.** Creating a corpus of Canadian healthcare data is extremely important. Most of the current research focusses on data from the United States; these results, however, cannot be used in the Canadian context due to different demographics and policies (e.g., insurance).
- **Need to collect demographic information.** AI/ML models learn from historical data, and therefore can reinforce any biases present in the data. Hospitals need to track demographic information to understand these biases, and in order to rectify them. This is an urgent need; we can't fix what we don't know!
- **No data. No AI/ML.** Data is the pre-requisite for any kind of data-driven modeling, and we need a unified effort to ensure that hospitals are equipped to collect clinical data accurately.
- **Annotating healthcare data is expensive.** Supervised ML modeling requires data annotations. There are ways to address lack of annotations, but we still need clinician's time to ensure we have labeled data for training.
- **Reality is often nuanced.** Doctor/surgeon's/medical professional's opinion are subjective, and we need to develop ML models that can handle such inherent ambiguity.
- **We are, in fact, never done with model training.** AI/ML models need to be monitored and continuously trained on new data to be reliable otherwise models can be outdated very quickly. Since a lot of this work has to be done in-house in hospitals, this calls for building the Canadian workforce in these areas.
- **AI/ML infrastructure.** Just collecting data is not enough. To channel this data for real-world impact, hospitals need strategic support (infrastructure and investments) for AI/ML model training and maintenance (mlOps).
- **It takes a village and data literacy.** Setting-up the infrastructure for a Canadian healthcare AI revolution requires coordination and knowledge building between departments, both at the administrative and the hospital level. It is therefore extremely important to have data literacy for organizations to ensure reliable application of AI/ML techniques and mitigate inaction.
- **Privacy and Cybersecurity.** Health data is extremely sensitive and needs to be protected. We can build privacy preserving ML models and synthetic data networks to achieve these goals; data stewardship is key.
- **Need for inter-operable blueprint for hospitals and health systems.** A standard and unified way of recording various types of patient information is absolutely essential for speedy transition. Hospitals need guidance.
- **Sharing is caring.** Likewise, hospitals will have to come together and share their data in a privacy preserving and secured way. We need to tackle data inequity from the get-go to ensure that far-off communities (who need these interventions the most) are not adversely impacted.

7 Conclusions

Advances in ML modeling can tackle challenges ranging from disease diagnosis to personnel training, in order to revolutionize the quality of care for patients and improved decision-making for medical personnel. While organizations may be eager to adopt AI/ML solutions, translating these gains to real-world healthcare systems has been slow due to data governance and collection challenges. Therefore, we end with an urgent call to bring together hospitals, industry experts, academic researchers, and policymakers to act upon a unified healthcare strategy for building AI ready healthcare systems. This will lead to faster adoption of research results, data transparency initiatives for patients, and provide a blueprint for AI-centric data organization in the health systems in Canada and abroad.

References

Alexandre De Figueiredo, Clarissa Simas, Emilie Karafillakis, Pauline Paterson, and Heidi J Larson. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. The Lancet, 396(10255):898–908, 2020.

David Xiang and Lisa Soleymani Lehmann. Confronting the misinformation pandemic. Health Policy and Technology, 10(3):100520, 2021.

Alice Wong. Actions for seniors report, 2014. URL https://www.canada.ca/en/news/archive/2014/09/ harper-government-launches-new-government-canada-action-seniors-report.html.

CFNU. Canada’s nursing shortage at a glance, 2022. URL https://nursesunions.ca/wp-content/uploads/ 2022/07/nurses_shortage_media_ref_guide_comp.pdf.

Brenda L. Gunn. Ignored to death: Systemic racism in the Canadian healthcare system. Human Rights Council, Expert Mechanism on the Rights of Indigenous Peoples, United Nations, 2016. URL https://www.ohchr.org/sites/ default/files/Documents/Issues/IPeoples/EMRIP/Session9/A.HRC.EMRIP.2016.CRP.1.pdf.

Emily Blake. People ’dying unnecessarily’ because of racial bias in Canada’s health-care system, researcher says. Canadian Broadcasting Corporation, 2018. URL https://www.cbc.ca/news/canada/north/ health-care-racial-bias-north-1.4731483.

Isabel Kim, Thalia S Field, Darryl Wan, Karin Humphries, and Tara Sedlak. Sex and gender bias as a mechanistic determinant of cardiovascular disease outcomes. Canadian Journal of Cardiology, 2022.

Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature machine intelligence, 1(5):206–215, 2019.

Michael Tsang, Sirisha Rambhatla, and Yan Liu. How does this interaction affect me? interpretable attribution for feature interactions. Advances in Neural Information Processing Systems, 33, 2020.

Russell H Wiesner, Sue V McDiarmid, Patrick S Kamath, Eric B Edwards, Michael Malinchoc, Walter K Kremers, Ruud AF Krom, and Ray W Kim. Meld and peld: application of survival models to liver allocation. Liver transplantation, 7(7):567–580, 2001.

A. C. Sheka, O. Adeyi, J. Thompson, B. Hameed, P. A. Crawford, and S. Ikramuddin. Non-alcoholic steatohepatitis: A review. JAMA, 323(12):1175–1183, 2020. ISSN 1538-3598. URL https://www.ncbi.nlm.nih.gov/pubmed/ 32207804.

E. L. Godfrey, T. H. Malik, J. C. Lai, A. L. Mindikoglu, N. T. N. Galván, R. T. Cotton, C. A. O’Mahony, J. A. Goss, and A. Rana. The decreasing predictive power of meld in an era of changing etiology of liver disease. Am J Transplant, 19(12):3299–3307, 2019. ISSN 1600-6135. doi:10.1111/ajt.15559.

A. J. Kwong, D. Devuni, C. Wang, J. Boike, J. Jo, L. VanWagner, M. Serper, L. Jones, R. Sharma, E. C. Verna, J. Shor, M. N. German, A. Hristov, A. Lee, E. Spengler, A. A. Koteish, G. Sehmbey, A. Seetharam, N. John, Y. Patel, M. R. Kappus, T. Couri, S. Paul, R. J. Salgia, Q. Nhu, C. T. Frenette, J. C. Lai, A. Goel, and Re-Evaluating Age Limits in Transplantation (REALT) Consortium. Outcomes of liver transplantation among older recipients with non-alcoholic steatohepatitis in a large multicenter us cohort: the re-evaluating age limits in transplantation consortium. Liver Transpl, 26(11):1492–1503, 2020. ISSN 1527-6473. doi:10.1002/lt.25863. URL https: //www.ncbi.nlm.nih.gov/pubmed/33047893.

R. S. Karnam, N. Mitsakakis, G. Saracino, L. Lilly, S. K. Asrani, and M. Bhat. Predicting long-term survival after liver transplantation in patients with nash cirrhosis. Clin Gastroenterol Hepatol, 20(3):704–705, 2022. ISSN 1542-3565.

D. Haldar, B. Kern, J. Hodson, M. J. Armstrong, R. Adam, G. Berlakovich, J. Fritz, B. Feurstein, W. Popp, V. Karam, P. Muiesan, J. O’Grady, N. Jamieson, S. J. Wigmore, J. Pirenne, S. A. Malek-Hosseini, E. Hidalgo, Y. Tokat, A. Paul, J. Pratschke, M. Bartels, P. Trunecka, U. Settmacher, M. Pinzani, C. Duvoux, P. N. Newsome, S. Schneeberger, European Liver (ELITA), and Intestine Transplant Association. Outcomes of liver transplantation for non-alcoholic steatohepatitis: A european liver transplant registry study. J Hepatol, 71(2):313–322, 2019. ISSN 1600-0641. doi:10.1016/j.jhep.2019.04.011. URL https://www.ncbi.nlm.nih.gov/pubmed/31071367.

L. B. Vanwagner, M. Bhave, H. S. Te, J. Feinglass, L. Alvarez, and M. E. Rinella. Patients transplanted for non-alcoholic steatohepatitis are at increased risk for postoperative cardiovascular events. Hepatology, 56(5):1741–50, 2012. ISSN 1527-3350. doi:10.1002/hep.25855. URL https://www.ncbi.nlm.nih.gov/pubmed/22611040.

M. R. Charlton, J. M. Burns, R. A. Pedersen, K. D. Watt, J. K. Heimbach, and R. A. Dierkhising. Frequency and outcomes of liver transplantation for non-alcoholic steatohepatitis in the United States. Gastroenterology, 141(4): 1249–53, 2011. ISSN 0016-5085. doi:10.1053/j.gastro.2011.06.061.

J. B. Henson, J. M. Wilder, M. R. Kappus, A. S. Barbas, C. A. Moylan, D. Niedzwiecki, A. J. Muir, C. L. Berg, and Y. A. Patel. Transplant outcomes in older patients with non-alcoholic steatohepatitis compared to alcohol-related liver disease and hepatitis c. Transplantation, 104(6):e164–e173, 2020. ISSN 1534-6080. doi:10.1097/TP.0000000000003219. URL https://www.ncbi.nlm.nih.gov/pubmed/32150036.

G. Punchhi, Yingji Sun, Sirisha Rambhatla, and Mamatha Bhat. Deep learning to predict trajectories & identify features associated with death & transplant in waitlisted nash patients. In Amer. Assoc. for the Stu. of Liver Dis. (AASLD), 2022a.

G. Punchhi, Yingji Sun, Sirisha Rambhatla, and Mamatha Bhat. Predicting future trajectories of the waitlisted nash patients using deep learning. In International Liver Transplantation Society (ILTS) Annual Congress, 2022b.

G. Punchhi, Yingji Sun, Sirisha Rambhatla, and Mamatha Bhat. Deep learning to predict trajectories and identify features associated with death and transplant in waitlisted nash patients. In CDTRP Annual Scientific Meeting, 2022c.

G. Punchhi, Y. Sun, S. Asrani, S. Rambhatla, and M. Bhat. Deepnash: A competing risk neural network model to forecast nash patient trajectories on the liver transplant waitlist. Journal Under Review, 2023.

Annette Spooner, Emily Chen, Arcot Sowmya, Perminder Sachdev, Nicole A. Kochan, Julian Trollor, and Henry Brodaty. A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction. Scientific Reports, 10(1):20410, 2020. ISSN 2045-2322. URL https://doi.org/10.1038/ s41598-020-77220-w.

Vardhmaan Jain, Agam Bansal, Nathan Radakovich, Vikram Sharma, Muhammad Zarrar Khan, Kevin Harris, Salam Bachour, Cerise Kleb, Jacek Cywinski, Maged Argalious, Cristiano Quintini, K. V. Narayanan Menon, Ravi Nair, Michael Tong, Samir Kapadia, and Maan Fares. Machine learning models to predict major adverse cardiovascular events after orthotopic liver transplantation: A cohort study. Journal of Cardiothoracic and Vascular Anesthesia, 35(7):2063–2069, 2021. ISSN 1053-0770. doi:https://doi.org/10.1053/j.jvca.2021.02.006. URL https://www. sciencedirect.com/science/article/pii/S1053077021001014.

Shunji Nagai, Anubhav Reddy Nallabasannagari, Dilip Moonka, Madhu Reddiboina, Sirisha Yeddula, Toshihiro Kitajima, Iman Francis, and Marwan Abouljoud. Use of neural network models to predict liver transplantation waitlist mortality. Liver Transplantation, n/a(n/a), 2022. ISSN 1527-6465. doi:https://doi.org/10.1002/lt.26442. URL https://doi.org/10.1002/lt.26442.

N. van Geloven, D. Giardiello, E. F. Bonneville, L. Teece, C. L. Ramspek, M. van Smeden, K. I. E. Snell, B. van Calster, M. Pohar-Perme, R. D. Riley, H. Putter, and E. Steyerberg. Validation of prediction models in the presence of competing risks: a guide through modern methods. Bmj, 377:e069249, 2022. ISSN 0959-8138. doi:10.1136/bmj-2021-069249.

Hemant Ishwaran, Udaya B. Kogalur, Eugene H. Blackstone, and Michael S. Lauer. Random survival forests. The Annals of Applied Statistics, 2(3):841 – 860, 2008. doi:10.1214/08-AOAS169. URL https://doi.org/10.1214/ 08-AOAS169.

J. L. Katzman, Uri Shaham, Alexander Cloninger, Jonathan Bates, Tingting Jiang, and Yuval Kluger. Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network. BMC medical research methodology, 18(1), 2018.

Changhee Lee, William Zame, Jinsung Yoon, and Mihaela van der Schaar. Deephit: A deep learning approach to survival analysis with competing risks. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1), 2018. URL https://ojs.aaai.org/index.php/AAAI/article/view/11842.

S. Rambhatla*, S. Huang*, L. Trinh, M. Zhang, M. Dong, V. Unadkat, H. A. Yenikomshian, J. Gillenwater, and Y. Liu. Dl4burn: Burn surgical candidacy using multimodal deep learning. American Medical Informatics Association (AMIA) Annual Symposium, 2021.

Begoña Acha, Carmen Serrano, Irene Fondón, and Tomás Gómez-Cía. Burn depth analysis using multidimensional scaling applied to psychophysical experiment data. IEEE Transactions on Medical Imaging, 32(6):1111–1120, 2013.

Marco D Cirillo, Robin Mirdell, Folke Sjöberg, and Tuan D Pham. Tensor decomposition for colour image segmentation of burn wounds. Scientific reports, 9(1):1–13, 2019.

Joohi Chauhan and Puneet Goyal. Bpbsam: Body part-specific burn severity assessment model. Burns, 2020.

Mario Aurelio Martínez-Jiménez, Jose Luis Ramirez-GarciaLuna, Eleazar Samuel Kolosovas-Machuca, Justin Drager, and Francisco Javier González. Development and validation of an algorithm to predict the treatment modality of burn wounds using thermographic scans: Prospective cohort study. PloS one, 13(11):e0206477, 2018.

Weizhi Li, Weirong Mo, Xu Zhang, John J Squiers, Yang Lu, Eric W Sellke, Wensheng Fan, J Michael DiMaio, and Jeffrey E Thatcher. Outlier detection and removal improves accuracy of machine learning approach to multispectral burn diagnostic imaging. Journal of biomedical optics, 20(12):121305, 2015.

Christopher Pham, Zachary Collier, and Justin Gillenwater. Changing the way we think about burn size estimation. Journal of Burn Care & Research, 40(1):1–11, 2019.

Carmen Serrano, Rafael Boloix-Tortosa, Tomás Gómez-Cía, and Begoña Acha. Features identification for automatic burn classification. Burns, 41(8):1883–1890, 2015.

DP Yadav, Ashish Sharma, Madhusudan Singh, and Ayush Goyal. Feature extraction based machine learning for human burn diagnosis from burn images. IEEE journal of translational engineering in health and medicine, 7:1–7, 2019.

Yuan Wang, Zuo Ke, Zhiyou He, Xiang Chen, Yu Zhang, Peizhen Xie, Tao Li, Jiao Zhou, Fangfang Li, Canqun Yang, et al. Real-time burn depth assessment using artificial networks: a large-scale, multicentre study. Burns, 46(8): 1829–1838, 2020.

Jordan A Browning and Renford Cindass. Burn debridement, grafting, and reconstruction. 2019.

Timothy HF Daugherty. Surgical excision of burn wounds. Burn Care: Rescue, Resuscitation, and Resurfacing, An Issue of Clinics in Plastic Surgery, E-Book, 44(3):619, 2017.

Lara Devgan, Satyanarayan Bhat, S Aylward, and Robert J Spence. Modalities for the assessment of burn wound depth. Journal of burns and wounds, 5, 2006.

Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, and Andrew Y Ng. Multimodal deep learning. In ICML, 2011.

Dezhao Song, Edward Kim, Xiaolei Huang, Joseph Patruno, Héctor Muñoz-Avila, Jeff Heflin, L Rodney Long, and Sameer Antani. Multimodal entity coreference for cervical dysplasia diagnosis. IEEE transactions on medical imaging, 34(1):229–245, 2014.

Tao Xu, Han Zhang, Xiaolei Huang, Shaoting Zhang, and Dimitris N Metaxas. Multimodal deep learning for cervical dysplasia diagnosis. In International conference on medical image computing and computer-assisted intervention, pages 115–123. Springer, 2016.

Garam Lee, Kwangsik Nho, Byungkon Kang, Kyung-Ah Sohn, and Dokyoon Kim. Predicting Alzheimer’s disease progression using multi-modal deep learning approach. Scientific reports, 9(1):1–12, 2019.

Fan Zhang, Zhenzhen Li, Boyan Zhang, Haishun Du, Binjie Wang, and Xinhong Zhang. Multi-modal deep learning model for auxiliary diagnosis of Alzheimer’s disease. Neurocomputing, 361:185–195, 2019.

Nils Gessert, Maximilian Nielsen, Mohsin Shaikh, René Werner, and Alexander Schlaefer. Skin lesion classification using ensembles of multi-resolution efficientnets with meta data. MethodsX, 7:100864, 2020. ISSN 2215-0161. doi:https://doi.org/10.1016/j.mex.2020.100864. URL https://www.sciencedirect.com/science/article/ pii/S2215016120300832.

Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR’05), volume 1, pages 886–893. IEEE, 2005.

Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In 2011 International conference on computer vision, pages 2564–2571. Ieee, 2011.

A. J. Hung, S. Rambhatla, D. I. Sanford, N. Pachauri, E. Vanstrum, J. H. Nguyen, and Y. Liu. Road to automating robotic suturing skills assessment: Battling mislabeling of the ground truth. Surgery, 2021a.

A. J. Hung, S. Rambhatla, D. I. Sanford, N. Pachauri, J. H. Nguyen, and Y. Liu. Automating suturing skills assessment with a limited surgeon dataset: Meta learning. American Urology Association, Journal of Urology, 1 2021b.
A. B. Chen, T. Haque, S. Roberts, S. Rambhatla, G. Cacciamani, P. Dasgupta, and A. J. Hung. Artificial intelligence applications in urology: Reporting standards to achieve fluency for urologists. Urology Clinics North America, 2 2022.

Tim Lane. A short history of robotic surgery. The Annals of The Royal College of Surgeons of England, 100(6_sup): 5–7, 2018.

Andrew J Hung, Paul J Oh, Jian Chen, Saum Ghodoussipour, Christianne Lane, Anthony Jarc, and Inderbir S Gill. Experts vs super-experts: differences in automated performance metrics and clinical outcomes for robot-assisted radical prostatectomy. BJU international, 123(5):861–868, 2019a.

Andrew J Hung, Jian Chen, and Inderbir S Gill. Automated performance metrics and machine learning algorithms to measure surgeon performance and anticipate clinical outcomes in robotic surgery. JAMA surgery, 153(8):770–771, 2018.

Andrew J Hung, Jian Chen, Saum Ghodoussipour, Paul J Oh, Zequn Liu, Jessica Nguyen, Sanjay Purushotham, Inderbir S Gill, and Yan Liu. A deep-learning model using automated performance metrics and clinical features to predict urinary continence recovery after robot-assisted radical prostatectomy. BJU international, 124(3):487–495, 2019b.

Alvin C Goh, David W Goldfarb, James C Sander, Brian J Miles, and Brian J Dunkin. Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. The Journal of urology, 187 (1):247–252, 2012.

JA Martin, Glenn Regehr, Richard Reznick, Helen Macrae, John Murnaghan, Carol Hutchison, and M Brown. Objective structured assessment of technical skill (osats) for surgical residents. British journal of surgery, 84(2):273–278, 1997.

Syed Johar Raza, Erinn Field, Christopher Jay, Daniel Eun, Michael Fumo, Jim C Hu, David Lee, Zayn Mehboob, John Nyquist, James O Peabody, et al. Surgical competency for urethrovesical anastomosis during robot-assisted radical prostatectomy: development and validation of the robotic anastomosis competency evaluation. Urology, 85(1):27–32, 2015.

John D Birkmeyer, Jonathan F Finks, Amanda O'reilly, Mary Oerline, Arthur M Carlin, Andre R Nunn, Justin Dimick, Mousumi Banerjee, and Nancy JO Birkmeyer. Surgical skill and complication rates after bariatric surgery. New England Journal of Medicine, 369(15):1434–1442, 2013.

Abhishek Trehan, Ashton Barnett-Vanes, Matthew J Carty, Peter McCulloch, and Mahiben Maruthappu. The impact of feedback of intraoperative technical performance in surgery: a systematic review. BMJ open, 5(6):e006759, 2015.

Thomas S Lendvay, Lee White, and Timothy Kowalewski. Crowdsourcing to assess surgical skill. JAMA surgery, 150 (11):1086–1087, 2015.

Loc Trinh, Samuel Mingo, Erik B Vanstrum, Daniel I Sanford, Runzhuo Ma, Jessica H Nguyen, Yan Liu, Andrew J Hung, et al. Survival analysis using surgeon skill metrics and patient factors to predict urinary continence recovery after robot-assisted radical prostatectomy. European Urology Focus, 8(2):623–630, 2022.

Andrew B Chen, Siqi Liang, Jessica H Nguyen, Yan Liu, and Andrew J Hung. Machine learning analyses of automated performance metrics during granular sub-stitch phases predict surgeon experience. Surgery, 169(5):1245–1249, 2021a.

Andrew J Hung, Sirisha Rambhatla, Daniel I Sanford, Nilay Pachauri, Erik Vanstrum, Jessica H Nguyen, and Yan Liu. Road to automating robotic suturing skills assessment: battling mislabeling of the ground truth. Surgery, 171(4): 915–919, 2022.

Karishma Sharma, Pinar Donmez, Enming Luo, Yan Liu, and I Zeki Yalniz. Noiserank: Unsupervised label noise reduction with dependence models. In European Conference on Computer Vision, pages 737–753. Springer, 2020.

A. Murugan, S. Rambhatla, and A. Wong. Whose health matters in healthcare models? unmasking data bias for data-aware modeling. Manuscript Under Review, 2023.

James Zou and Londa Schiebinger. Design ai so that it's fair. Nature, 559(7714):324–326, 2018.

Irene Chen, Fredrik D Johansson, and David Sontag. Why is my classifier discriminatory? Advances in neural information processing systems, 31, 2018.
Irene Y Chen, Peter Szolovits, and Marzyeh Ghassemi. Can ai help reduce disparities in general medical and mental health care? AMA journal of ethics, 21(2):167–179, 2019.

Alvin Rajkomar, Jeffrey Dean, and Isaac Kohane. Machine learning in medicine. New England Journal of Medicine, 380(14):1347–1358, 2019.

Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464):447–453, 2019.

Alvin Rajkomar, Michaela Hardt, Michael D Howell, Greg Corrado, and Marshall H Chin. Ensuring fairness in machine learning to advance health equity. Annals of internal medicine, 169(12):866–872, 2018.

Irene Y Chen, Emma Pierson, Sherri Rose, Shalmali Joshi, Kadija Ferryman, and Marzyeh Ghassemi. Ethical machine learning in healthcare. Annual review of biomedical data science, 4:123–144, 2021b.

Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. circulation, 101(23):e215–e220, 2000.

AE Johnson, Tom J Pollard, Lu Shen, L-w H Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, L Anthony Celi, and Roger G Mark. Mimic-iii, a freely accessible critical care database sci. Data, 3 (160035):10–1038, 2016.

Nathalie Japkowicz. The class imbalance problem: Significance and strategies. In Proc. of the Int'l Conf. on artificial intelligence, volume 56, pages 111–117, 2000.

Alistair EW Johnson, Lucas Bulgarelli, Lu Shen, Alvin Gayles, Ayad Shammout, Steven Horng, Tom J Pollard, Benjamin Moody, Brian Gow, Li-wei H Lehman, et al. Mimic-iv, a freely accessible electronic health record dataset. Scientific data, 10(1):1, 2023.

Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. Datasheets for datasets. Communications of the ACM, 64(12):86–92, 2021.

USMLE. Step 1 sample test questions, Jun 2022. URL https://www.usmle.org/prepare-your-exam/ step-1-materials/step-1-sample-test-questions.

OpenAI, 2023. URL https://platform.openai.com/.

F. Barnard, M. V. Sittert, and S. Rambhatla. Self-diagnosis and large language models: A new front for medical misinformation. Under Review, 2023. URL https://arxiv.org/abs/2307.04910.

Viriya Taecharungroj. “What Can ChatGPT Do?” Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. Big Data and Cognitive Computing, 7(1):35, March 2023. ISSN 2504-2289. doi:10.3390/bdcc7010035. URL https://www.mdpi.com/2504-2289/7/1/35. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

Mohammed Salah, Hussam Alhalbusi, Maria Mohd Ismail, and Fadi Abdelfattah. Chatting with ChatGPT: Decoding the Mind of Chatbot Users and Unveiling the Intricate Connections between User Perception, Trust and Stereotype Perception on Self-Esteem and Psychological Well-being. preprint, In Review, March 2023. URL https://www. researchsquare.com/article/rs-2610655/v2.

Clare Goyder, Ann McPherson, and Paul Glasziou. Self diagnosis. BMJ, 339:b4418, November 2009. ISSN 0959-8138, 1468-5833. doi:10.1136/bmj.b4418. URL https://www.bmj.com/content/339/bmj.b4418. Publisher: British Medical Journal Publishing Group Section: Practice.

Ryen W. White and Eric Horvitz. Experiences with Web Search on Medical Concerns and Self Diagnosis. AMIA Annual Symposium Proceedings, 2009:696–700, 2009. ISSN 1942-597X. URL https://www.ncbi.nlm.nih. gov/pmc/articles/PMC2815378/.

Bridget M. Kuehn. Clinician Shortage Exacerbates Pandemic-Fueled “Mental Health Crisis”. JAMA, 327(22):2179– 2181, June 2022. ISSN 0098-7484. doi:10.1001/jama.2022.8661. URL https://doi.org/10.1001/jama.2022. 8661.

Mathieu Boniol, Teena Kunjumen, Tapas Sadasivan Nair, Amani Siyam, James Campbell, and Khassoum Diallo. The global health workforce stock and distribution in 2020 and 2030: a threat to equity and ‘universal’ health coverage? BMJ Global Health, 7(6):e009316, June 2022. ISSN 2059-7908. doi:10.1136/bmjgh-2022-009316. URL https://gh.bmj.com/content/7/6/e009316. Publisher: BMJ Specialist Journals Section: Original research.

Jean-Pierre Michel and Fiona Ecarnot. The shortage of skilled workers in Europe: its impact on geriatric medicine. European Geriatric Medicine, 11(3):345–347, June 2020. ISSN 1878-7657. doi:10.1007/s41999-020-00323-0. URL https://doi.org/10.1007/s41999-020-00323-0.

Sue Turale and Apiradee Nantsupawat. Clinician mental health, nursing shortages and the COVID-19 pandemic: Crises within crises. International Nursing Review, 68(1):12–14, 2021. ISSN 1466-7657. doi:10.1111/inr.12674. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/inr.12674. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/inr.12674.

Ari Horesh. Using ChatGPT to study medicine: Learn the basics, Mar 2023. URL https://futuredoctor.ai/ chatgpt/.

Hyunsu Lee. The rise of ChatGPT: Exploring its potential in medical education. Anatomical Sciences Education, March 2023. ISSN 1935-9780. doi:10.1002/ase.2270.

Sam Sedaghat. Early applications of ChatGPT in medical practice, education and research. Clinical Medicine (London, England), pages clinmed.2023–0078, April 2023. ISSN 1473-4893. doi:10.7861/clinmed.2023-0078.

Tiffany H. Kung, Morgan Cheatham, Arielle Medenilla, Czarina Sillos, Lorie De Leon, Camille Elepaño, Maria Madriaga, Rimel Aggabao, Giezel Diaz-Candido, James Maningo, and Victor Tseng. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digital Health, 2(2): e0000198, February 2023. ISSN 2767-3170. doi:10.1371/journal.pdig.0000198. URL https://journals.plos. org/digitalhealth/article?id=10.1371/journal.pdig.0000198. Publisher: Public Library of Science.

Linta Iftikhar et al. Docgpt: Impact of chatgpt-3 on health services as a virtual doctor. EC Paediatrics, 12(1):45–55, 2023.

Dan Primack. Here come the robot doctors, Jan 2023. URL https://www.axios.com/2023/01/18/ chatgpt-ai-health-care-doctors.

Jason Holmes, Zhengliang Liu, Lian Zhang, Yuzhen Ding, Terence T. Sio, Lisa A. McGee, Jonathan B. Ashman, Xiang Li, Tianming Liu, Jiajian Shen, and Wei Liu. Evaluating Large Language Models on a Highly-specialized Topic, Radiation Oncology Physics, April 2023. URL http://arxiv.org/abs/2304.01938. arXiv:2304.01938 [physics].

Aidan Gilson, Conrad Safranek, Thomas Huang, Vimig Socrates, Ling Chi, R. Andrew Taylor, and David Chartash. How Does ChatGPT Perform on the Medical Licensing Exams? The Implications of Large Language Models for Medical Education and Knowledge Assessment, December 2022. URL https://www.medrxiv.org/content/ 10.1101/2022.12.23.22283901v1. Pages: 2022.12.23.22283901.

Malik Sallam, Nesreen Salim, Muna Barakat, and Alaa Al-Tammemi. ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. Narra J, 3

(1):e103–e103, March 2023. ISSN 2807-2618. doi:10.52225/narra.v3i1.103. URL https://narraj.org/main/ article/view/103. Number: 1.

Malik Sallam. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare, 11(6):887, January 2023. ISSN 2227-9032. doi:10.3390/healthcare11060887. URL https://www.mdpi.com/2227-9032/11/6/887. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

Luigi De Angelis, Francesco Baglivo, Guglielmo Arzilli, Gaetano Pierpaolo Privitera, Paolo Ferragina, Alberto Eugenio Tozzi, and Caterina Rizzo. ChatGPT and the Rise of Large Language Models: The New AI-Driven Infodemic Threat in Public Health, February 2023. URL https://papers.ssrn.com/abstract=4352931.

Vivek Goel, Ewan Affleck, David Castle, Stafford Dean, Jonathan Dewar, Isabel Fortier, Michael Harvey, Steven Hoffman, Bartha Maria Knoppers, Alies Maybee, Muhammad Mamdani, Kim McGrail, Jeff Nes-bitt, Cory Neudorf, Glynda Rees, Janet Smylie, Gail Tomblin Murphy, Graham Tipples, and Michael Wolfson. The pan-Canadian health data strategy, the expert advisory report 1: Charting a path toward am-bition, 2022. URL https://www.canada.ca/en/public-health/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries.html

Trust in Canadian Government: What is it, (how) has it changed, can it be measured, and why does it matter for Population Health?

Dr. Samantha B. Meyer
Associate Professor
School of Public Health Sciences
University of Waterloo

Abstract

Researchers and those in health policy have made tremendous advancements in utilizing health data for the purpose of system improvement and promoting the health of Canadians. However, these technologies are heavily dependent on the public sharing their health data. We are asking the public to trust in government and related social institutions (e.g., universities, research institutes, public health) to ensure 1) the privacy of their data, and 2) that data is used in a manner that serves the interest of citizens. The present paper introduces the concept of institutional trust and speaks to the critical role of trust in society broadly and as it relates to health behaviours, including data sharing. Literature focused on understanding and explaining declining trust are explored, including the proliferation of misinformation on social media, as well as broader social issues in Canada – e.g., levels of inequality in the population. The importance of trust for population health is also discussed. The paper concludes with a discussion of the importance of measuring citizens' trust over time in an ever-changing political climate and how this might help to inform prospective actions aiming to demonstrate the trustworthiness of Canadian social institutions and subsequently support data sharing. Lastly, approaches to monitoring trust over time are discussed, with specific mention of how research might focus on identifying populations of Canadians that would benefit from interventions to (re)build trust.

Introduction.

Within this e-book we discuss the utility of collecting Canadian's health data – via mobile or other devices, and via hospital administration – to facilitate health system improvement. Technologies reliant on health data (e.g., artificial intelligence) have potential to change the way our health data are collected, shared, and used to improve the health of Canadians. However, as contributors in this e-book discuss, cybersecurity and privacy research to protect health data remain an area of ongoing development. The present paper considers these tremendous advancements and the potential for system improvement through the lens of the consumer – the individuals whom we are asking to trust the institutions generating these technologies and managing the security of personal health data. Our 'consumers' are the patients within our health system and the citizens for whom we are generating such incredible technologies. Whether or not, and how these technologies are applied, is heavily dependent on the public's trust in the science, government, and related social institutions (e.g., universities, research institutes, public health) that will collect, store, and use these data.

Trust is the acceptance of vulnerability (risk) and is based on positive expectations of the person or institution we are trusting. Trust is central to any decision an individual makes if there is something they consider to be at risk – in the case of sharing personal medical information, their privacy and security. Related to data sharing, we are asking people to trust that their medical information is being collected and stored securely (the system is competent), and that their data are collected for the purpose of creating a better health system (that those collecting the data have their best interests in mind). Without the trust of citizens and consequently their sharing of data, we will not have the information required to modernize our health system, generate evidence-informed decision making, or the data to inform AI technologies. Indeed, data shows that higher trust in technology is associated with increased technology use (1) and the failure to earn the trust of patients and citizens can have significant implications for the implementation of technologies in public systems. For example, a failure to earn the trust of patients, citizens, and healthcare providers is what led to the failure of Care.data – a big data initiative set to link and provide access to health and social care information from different settings, with the goal of facilitating the planning and provision of healthcare in the UK (2). A lack of public trust regarding how data might be used, anonymized, and stored led to the failure of what might have been an innovative integrated data repository to monitor clinical outcomes for quality improvement.

The aim of this paper is to discuss trust in government as it relates to health behaviour while noting implications for data sharing. The first section introduces the concept of trust and why trust in social institutions is so critical for population health. Data are then presented regarding the extent of trust in the Canadian government, followed by a discussion of how access to mis- and dis-information (3) challenge the trustworthiness of Canadian social institutions. The paper concludes with a discussion of how data, collected through valid measures of trust, might assist us in systematically documenting Canadian's trust over time, and help to generate evidence-informed interventions to (re)build trust through communication and/or public policy.

What is trust and why does it matter?

Trust, in both institutions and individuals, is critical for society to function (4). Given the complexity of our world and the decisions we make daily, we are often faced with a decision of whether to trust in a person or institutions to overcome the gaps in our knowledge. To trust is to reduce the complexity of our decisions, and to reduce uncertainty. For example, we trust a physician to provide a diagnosis because they have the training and expertise provided in medical school, to which we are not privy. We trust institutions – for example, public health – to ensure that safe and effective vaccines are administered to the population. We trust them because we, as citizens, do not have the expertise or ability to develop, test, procure, and regulate vaccines.

What distinguishes trust from other related concepts (e.g., hope, faith, dependence, confidence) is that trust is a decision, and one that carries an element of risk and uncertainty. As consumers, we cannot be certain that our doctor will give us the right advice or that a vaccine might not carry adverse effects. As such, if trust is broken, we have something to lose. The decision to trust is thus based on our assessment of risk which may relate to our past experiences with individuals or institutions upon which we are called to trust. A useful example here is the use of TikTok – while organizations and governments have made the decision to ban this application on corporate devices for privacy and security concerns (they do not trust the app or rather, the individuals/organisations behind the app). For many users, these security risks are either not considered to be a risk or are something of which they are unaware. It is the awareness of risk that forces us to trust, or not.

Acknowledging that trust is an active choice is critical for understanding why trust should be a central concern for governments globally. It is unlikely that citizens, at a population level, have ever, (at least to the extent that they are today), been required to consider *if* they trust their government. The unprecedented nature of the pandemic and related exposure to government representatives (e.g., the Prime Minister at the time held daily press conferences) brought to centre stage the role of government in setting the norms and rules for society. The risks we take in trusting government officials suddenly became very salient, as the decisions they made so greatly impacted our living and working conditions. Further, we had little knowledge with which to counter decisions made early in the pandemic and thus, our trust reduced our uncertainty as government led a chosen way forward out of the pandemic. In looking back, the public became more aware of the fallibility of government and officials in making some of these decisions, and the impact of these decisions on the lives of Canadians led to further questioning of government care and competence, and hence, trust.

Conceptually, scholars identify trust at two levels – institutional and interpersonal (5) – that are often inextricably linked; that is, trust in individuals (e.g., Trudeau), who serve as the face of institutions (e.g., Federal government), to some extent impacts trust in the organization they represent. The focus here is that of institutional trust in government, which may be evaluated by the public based on their perception of whether the government is benevolent, competent, reliable, and responsible toward citizens (7,8) – that is, their perception that the government cares and is competent. Indeed, research has demonstrated that public trust in political institutions, and their representatives, is predicated on public perception that the actions and intentions of these institutions are in the public's best interest and deliver results that meet their expectations (9–11). Trust differs from having a positive or negative opinion of an institution (6). For example, I may disagree with a policy put in place by government while still having trust in the institution.

Trust in government has been associated with increased social cohesion (solidarity among community members) and interpersonal trust between citizens (7, 8), and with public acceptance of government policy (9) and government recommended health behaviours (10-13). Critical in a world with increased access to harmful mis- and dis-information, trust in government is also negatively associated with believing misinformation and conspiracy theories (14). This is a critical association and demonstrates that if citizens do not trust the government, they may look elsewhere for information to guide their decisions. Declining trust thus threatens the legitimacy of the social institutions put in place to structure the rules and norms shaping behaviour at a population level.

Trust in government within Canada.

In 2023, it was reported that 37% of citizens trusted the national government (15). While this was a jump up from 22% in 2022, trust in Canadian parliament scored second to immigration services, in terms of the least trusted social institution in Canada (15). Trust in the current Liberal party has also improved since historic lows in 2022, jumping from 43% to 54% (16); however, this suggests that just over half of the population trusts the current political party in power. Trust in the Prime Minister currently sits at 32%, falling 1% since 2022 (16). These statistics may relate to citizens' concerns regarding the government's response to the pandemic (17), and/or relate to perceptions of inadequate government attention to pressing social issues in Canada, as discussed below. What some have called 'a trust crisis' (see https://trustgov.net/) is not isolated to Canada, however, as there is widespread concern that public trust in democratic institutions globally is under threat. As such, research that might inform strategies to build trust are of central focus.

That trust in government, and trust and cohesion between citizens (solidarity among community members), is at threat might be best illustrated by the January 2022 so-called 'Freedom Convoy' in Ottawa ON. This convoy – born out of a vehement distrust in government – led to rallies, widely considered unlawful, that posed a significant threat to public peace and safety. This demonstration ultimately led to less trust between citizens, with increasing division amongst the populace, and further, a loss of trust in government by the citizens of Ottawa whose lives were impacted by the unlawful behaviours of demonstrators (18). Residents of Ottawa impacted by the demonstration have since called upon officials, as means for officials to earn their trust, to admit that the official's inaction in protecting the public, caused harm (19).

The factors underpinning the trust of Canadians are complex and vary across the population. As noted earlier, as trust in government and social institutions is called into question, the public starts to look to alternative forms of information, that are often harmful and counter the actions of officials to protect the public. Herein the focus is one such source of harmful mis- and dis-information – social media.

Misinformation, alternative experts, and social media.

The circulation of misinformation since the start of the pandemic on social media, has led to popular conspiracy theories that are dangerous from a health perspective in that they are associated with reduced compliance with government health guidelines, posing a danger to population health (20). For example, COVID-19 misinformation adversely affects governments' abilities to mitigate the disease and is associated with accidental deaths and self-harm due to false beliefs about the virus, prevention measures, vaccines, and cures (21).

Social media has drastically changed the way that individuals come to access and consume news, as these sites provide instant news updates and thus rapid distribution of information to consumers (20). While there are benefits to social media use in terms of consumer engagement with news and engagement with different sources of news and information seeking, the negative implications of its use are also widely documented, and central to this paper. These implications include the increased access to dis- and misinformation that threaten the legitimacy and trust of science and social institutions promoting health and social behaviours (which may include data sharing) to support population health. Trust in social media, rather than evidence-informed recommendations guiding health recommendations, is an issue of concern. The notion of what is 'legitimate' information has changed within the context of social media, as users increasingly see the peers within their social media networks as an authoritative and legitimate source of information (22). The harms arise when this alternative form of 'expertise' contrasts with information that is generated by institutions with specialized expertise – for example, public health agencies. Indeed, scientific misinformation has been "actively propagated as a means to destabilise trust in governments" (22) (p. e277).

To combat the dangers of misinformation, various entities within the social media industry have worked (to varying degrees), to generate mechanisms for combating misinformation. However, the absence of adequate control mechanisms make it difficult to regulate the nature of information circulated on social media (23). Innovative approaches, such as having users downrank distrusted sources of information on social media have also been proposed (24). Still, policy makers continue to face growing challenges with the proliferation of misinformation on social media and the potential association with harmful health behaviours and subsequently, health outcomes (25).

There are also calls from authoritative organizations to try and combat misinformation by engaging with social media (26), though in doing so we need to be careful not to further polarize users (e.g., by pushing individuals to defend their existing beliefs) (27). Strategies for combating misinformation have also been generated to combat the acceptance of misinformation – e.g., by pre-bunking or inoculation communication strategies to counter disinformation, which involves forewarning people regarding possible exposure to information that will challenge their existing beliefs or behaviors (28). There are also calls for greater media literacy interventions to support individuals in identifying false information online (21), and particularly among the younger generation who, given their level of engagement online, may be at greater risk of exposure to information leading to the rejection of government health guidelines (20). However, the question remains – why do people trust information on social media over authoritative information guided by scientific institutions and (previously) trusted social institutions? It is important to consider that 'expertise' is socially constructed (29) – that is, while researchers might consider science as the authority in guiding public health recommendations, it is dangerous to assume that this occurs at a population level. Further, scientific terminology, such as ‘evidence’, or ‘research’, may signal that a source should be trusted, despite being used to manipulate and propagate misinformation by content creators (30).

What remains an important question is whether a lack of trust in government organizations is associated with greater trust in ‘alternative’ forms of expertise found online. And more importantly, we might look to understand why our institutions are not deserving of trust (31). These questions remain areas for empirical research in Canada, as discussed below.

Future research/practice agendas.

The demonstrated importance of trust – or rather, the consequences of a lack of trust – points to the continued need to better understand, measure, and respond to public concern. While it is well documented that a lack of trust in government threatens population health, less is known about *why* individuals are looking for alternative sources of information to trust, *how* this varies across the population, and *how* we might develop *strategies* to (re)build trust tailored to different subpopulations for whom the reasons why trust might be at threat, or indeed might have been lost, differ. To do this, we need to systematically monitor and understand the government (in)actions that lead to a change in attitudes and beliefs about government and importantly, how it affects public behaviour.

The CANTRUST index (https://getproof.com/trust/cantrust/) provides a tool for regular and rapid evaluation of the public’s trust in government. The current survey question “to what extent are you willing to trust the following institutions to operate competently and effectively and to do the right thing” allows us to see how levels of trust change over time and how they vary across population groups. The single item question that incorporates two dimensions of trust – care and competence – limits our understanding of the ‘why’ behind the data and thus, what can be done to improve trust. More comprehensive multidimensional measures developed for the purpose of evaluating and monitoring trust in government in Canada – at a local, provincial, and federal level – might complement CANTRUST data by providing a more nuanced understanding of public acceptance of government messaging and inform behavioural change measures, communication strategies, and policies in response. Using tools, for example, that contain survey items that capture the multidimensionality of trust in government, might better direct interventions for improving trust. The ‘Citizen Trust in Government Organizations’ scale (32) and a ‘Trust in Public Health Authorities scale’ (33), for example, evaluate trust based on responses related to government’s communication, decision making, transparency, honesty and delivering on promises. They permit greater observation of areas where efforts might be focused on (re)building trust.

There also remains a gap in measures of trust at a population level that are validated for specific sub populations of Canadians that might be used to inform the development of strategies for (re)building trust across communities and populations. As expected, there is less trust in government by population groups that have not been advantaged by social institutions to the same extent as more privileged populations in Canada. For example, the COVID-19 pandemic exacerbated and made visible existing inequities across subpopulations of Canadians – with higher rates of infection occurring in, for example, racially and ethnically-minoritized residents. Within Ontario, minoritized residents reported three times higher rates of COVID-19 infection, four times higher rates of hospitalization and intensive care unit admissions, and a two times higher rate of death, as compared to neighbourhoods with higher proportions of white residents in 2020 (34). Universal preventative interventions (i.e., social distancing or self-isolation) to mitigate COVID-19 spread, were designed with the assumption that Canadians had the will and capacity to accept recommendations, and that they trusted them, which has been demonstrated as critical for the acceptance of COVID-19 countermeasures (35). In reality, public health communications were limited in reaching populations already facing health inequities and for some, resulted in lost trust as a result of continued underservice in their already marginalized communities (see (36)). Importantly, underservice by government during the pandemic has been documented to create opportunities for misinformation to spread through trusted channels, such as peer-to-peer discussion via social media (36). It is therefore critical to identify evidence-informed approaches to redressing the social injustices and/or concerns that lead individuals to look to alternative sources of information that may be harmful. Development of measures for specific populations will require more engagement with communities of focus to ensure survey items reflect what matters in terms of the assessment of trust. For example, researchers might consider generating survey items to complement existing measures of trust that reflect the present day context – that is, the contemporary issues underpinning trust (37). This might allow us to better identify and understanding systemic explanations (e.g., discrimination) for lower levels of trust among equity-deserving populations that we might redress through tailored strategies to earn the trust of the public. Qualitative research investigating the nature of trust in populations of focus will be required to generate/refine existing items to ensure their relevance prior to validation studies.

Conclusion.

The Canadian government can only function in the interest of its citizens if there is public trust. Relevant to this e-book, any advancements in health technologies necessitating access to personal health data will require public trust. The public will need to trust that their data are being used in a manner that serves their interests and that of their fellow citizens (2). In addition to competence, government and related agencies involved in collecting these data need to be transparent in the risk of sharing such data and competence in minimizing these risks. While they may seem unrelated, public perception of government broadly – for example, regarding efforts toward Reconciliation with Indigenous Peoples for historical injustices**,** high levels of inequality in the population, general dissatisfaction with governmental approaches to managing economic hardship, climate change and the COVID-

19 pandemic – will translate to a lack of trust in other areas of government. A lack of trust in government related to (in)action challenges the legitimacy of an institution. As such, it will be critical that the government works to earn the trust of the public if we are to expect individuals to adopt technologies that are considered as posing a risk.

While trust takes a long time to develop, it is easily broken. Researchers have generated discussion regarding approaches to rebuilding trust, including those specific to promoting data share. For example, The Pan-Canadian Health Strategy states: “We need to [create] transparent, pro-active two-way communication with the public to help earn and sustain their trust, and demonstrate the value of health data sharing” (38). How recommended actions are enacted is complicated and will require dedicated investment in policy and communication, genuine commitment to redressing the reasons why trust was lost (or indeed never formed), and consistent monitoring to ensure our actions are generating trust. We need to follow through and demonstrate accountability which will involve ongoing evaluation of our efforts.

The health policy and research community are paying greater attention than ever before to the importance of trust. It is a critical time to focus our efforts on trust – that is, how it can be earned and maintained. Importantly, in populations where trust in government remains, it should not be taken for granted. It will be important to ensure that Canadians continue to feel that their governmental agencies are competent and that the interests of citizens inform their actions.

Acknowledgements.

I would like to acknowledge the investigators and trainees with whom I have discussed matters of trust as they relate to population health, informing much of the work I present in this paper. Investigators: P. Brown (University of Amsterdam, NL), M. Calnan (University of Kent, UK), P.R. Ward (Torrens University, AUS), J. Little (University of Waterloo, CAN), G.S. Betini (University of Waterloo, CAN), C.M. Perlman (University of Waterloo, CAN). University of Waterloo Trainees: E. Filice, H. Herati, N. Ike, M. Nascimento, L. Koshy, K. Burns.

References.

1. Aboueid S, Meyer SB, Wallace J, Chaurasia A. Latent classes associated with the intention to use a symptom checker for self-triage. PLOS ONE. 2021;16(11):e0259547.
2. van Staa T-P, Goldacre B, Buchan I, Smeeth L. Big health data: the need to earn public trust. BMJ. 2016;354:i3636.
3. Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, et al. The science of fake news. Science. 2018;359(6380):1094-6. Mis-information refers to information that is false or misleading and dis-information as information that is false but also purposely spread to deceive people
4. Ward PR. Improving Access to, Use of, and Outcomes from Public Health Programs: The Importance of Building and Maintaining Trust with Patients/Clients. Front Public Health. 2017;5:22.
5. Meyer SB, Ward PR, Coveney J, Rogers W. Trust in the health system: an analysis and extension of the social theories of Giddens and Luhmann. Health Sociology Review. 2008;17(2):177-86.
6. Devos T, Spini D, Schwartz SH. Conflicts among human values and trust in institutions. The British journal of social psychology. 2002;41(Pt 4):481-94.
7. Rothstein B, Stolle D, editors. The handbook of social capital. New York: Oxford University Press; 2008.
8. Spadaro G, Gangl K, Van Prooijen J-W, Van Lange PAM, Mosso CO. Enhancing feelings of security: How institutional trust promotes interpersonal trust. PLOS ONE. 2020;15(9):e0237934.
9. Algan Y, Cahuc P. Chapter 2 - Trust, Growth, and Well-Being: New Evidence and Policy Implications. In: Aghion P, Durlauf SN, editors. Handbook of Economic Growth. 2: Elsevier; 2014. p. 49-120.
10. Marlow LA, Waller J, Wardle J. Trust and experience as predictors of HPV vaccine acceptance. Human vaccines. 2007;3(5):171-5.
11. van der Weerd W, Timmermans DR, Beaujean DJ, Oudhoff J, van Steenbergen JE. Monitoring the level of government trust, risk perception and intention of the general public to adopt protective measures during the influenza A (H1N1) pandemic in The Netherlands. BMC Public Health. 2011;11:575.
12. Wang Q, Yang L, Jin H, Lin L. Vaccination against COVID-19: A systematic review and meta-analysis of acceptability and its predictors. Prev Med. 2021;150:106694.
13. Shanka MS, Menebo MM. When and How Trust in Government Leads to Compliance with COVID-19 Precautionary Measures. Journal of Business Research. 2022;139:1275-83.
14. Majid U, Wasim A, Truong J, Bakshi S. Public trust in governments, health care providers, and the media during pandemics: A systematic review. Journal of Trust Research. 2021;11(2):119-41.
15. Proof Strategies. Special Report: The crisis of trust in government and democracy. CanTrust Index; 2023.
16. Proof Strategies. 2023 Results Report.
17. Levitt EE, Gohari MR, Syan SK, Belisario K, Gillard J, DeJesus J, et al. Public health guideline compliance and perceived government effectiveness during the COVID-19 pandemic in Canada: Findings from a longitudinal cohort study. The Lancet Regional Health - Americas. 2022;9:100185.
18. Ottawa People's Commission. Part 1: What we heard. https://www.opc-cpo.ca/wp-content/uploads/2023/02/OPC-Report-Part-I-What-We-Heard.pdf; 2023.
19. Ottawa People's Commission. Part II - After the occupation: Change. https://www.opc-cpo.ca/wp-content/uploads/2023/04/Recommendations-from-Part-II-After-The-Occupation-Change_EN.pdf; 2023.
20. Xiao X, Borah P, Su Y. The dangers of blind trust: Examining the interplay among social media news use, misinformation identification, and news trust on conspiracy beliefs. Public Understanding of Science. 2021;30(8):977-92.
21. Melki J, Tamim H, Hadid D, Makki M, El Amine J, Hitti E. Mitigating infodemics: The relationship between news exposure and trust and belief in COVID-19 fake news and social media spreading. PLOS ONE. 2021;16(6):e0252830.
22. Limaye RJ, Sauer M, Ali J, Bernstein J, Wahl B, Barnhill A, et al. Building trust while influencing online COVID-19 content in the social media world. The Lancet Digital Health. 2020;2(6):e277-e8.
23. Dabbous A, Aoun Barakat K, de Quero Navarro B. Fake news detection and social media trust: a cross-cultural perspective. Behaviour & Information Technology. 2022;41(14):2953-72.
24. Epstein Z, Pennycook G, Rand D. Will the Crowd Game the Algorithm? Using Layperson Judgments to Combat Misinformation on Social Media by Downranking Distrusted Sources. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; Honolulu, HI, USA: Association for Computing Machinery; 2020. p. 1–11.
25. Chou W-YS, Oh A, Klein WMP. Addressing Health-Related Misinformation on Social Media. JAMA. 2018;320(23):2417-8.
26. Witteman HO, Zikmund-Fisher BJ. The defining characteristics of Web 2.0 and their potential influence in the online vacciantion debate. Vaccine. 2012;30:3734-40.
27. Meyer SB, Violette R, Aggarwal R, Simeoni M, MacDougall H, Waite N. Vaccine hesitancy and Web 2.0: Exploring how attitudes and beliefs about influenza vaccination are exchanged in online threaded user comments. Vaccine. 2019;37:1769-74.
28. Vivion M, Anassour Laouan Sidi E, Betsch C, Dionne M, Dubé E, Driedger SM, et al. Prebunking messaging to inoculate against COVID-19 vaccine misinformation: an effective strategy for public health. Journal of Communication in Healthcare. 2022;15(3):232-42.
29. Attwell K, Leask J, Meyer SB, Rokkas P, Ward PR. Vaccine rejecting parents' engagement with expert systems that inform vaccination programs. Journal of Bioethical Inquiry. 2017;14(1):65-76.

30. Meyer SB, Lunnay B, Warin M, Foley K, Olver IN, Wilson C, et al. Examining social class as it relates to heuristics women use to determine the trustworthiness of information regarding the link between alcohol and breast cancer risk. PLoS One. 2022;17(9):e0270936.
31. Benjamin R. Informed Refusal: Toward a Justice based Bioethics. Science, Technology, & Human Values. 2016;41(6):967-90.
32. Grimmelikhuijsen S, Knies E. Validating a scale for citizen trust in government organizations. International Review of Administrative Sciences. 2017;83(3):583-601.
33. Holroyd TA, Limaye RJ, Gerber JE, Rimal RN, Musci RJ, Brewer J, et al. Development of a Scale to Measure Trust in Public Health Authorities: Prevalence of Trust and Association with Vaccination. J Health Commun. 2021;26(4):272-80.
34. Public Health Ontario. COVID-19 in Ontario – A focus on diversity: January 15, 2020 to May 14, 2020. Queen’s Printer for Ontario. 2020;https://files.ontario.ca/moh-covid-19-report-en-2020-05-15.pdf.
35. Han Q, Zheng B, Cristea M, Agostini M, Bélanger JJ, Gützkow B, et al. Trust in government regarding COVID-19 and its associations with preventive health behaviour and prosocial behaviour during the pandemic: a cross-sectional and longitudinal study. Psychol Med. 2021:1-11.
36. Bhalla M, Boutros H, Meyer SB. Aunties, WhatsApp, and "haldi da doodh": South Asian communities' perspectives on improving COVID-19 public health communication in Ontario, Canada. Canadian journal of public health = Revue canadienne de sante publique. 2022;113(Suppl 1):46-53.
37. Aboueid SE, Herati H, Nascimento MHG, Ward PR, Brown P, Calnan M, et al. How do you measure trust in social institutions and health professionals? A systematic review of the literature (2012-2021). Sociology Compass. In Press.
38. Expert Advisory Group Pan-Canadian Health Data Strategy. Pan-Canadian Health Data Strategy: Building Canada’s Health Data Foundation. https://www.canada.ca/content/dam/phac-aspc/documents/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries/expert-advisory-group-report-02-building-canada-health-data-foundation/expert-advisory-group-report-02-building-canada-health-data-foundation.pdf: Public Health Agency of Canada; 2021.

Addressing the Controversy on Mobility Data to Analyze Daily COVID-19 Cases: Implications for Population Data Literacy

Anindya Sen
Professor of Economics
Associate Director, Cybersecurity & Privacy Institute
Member, Waterloo Artificial Intelligence Institute
University of Waterloo

Abstract

This study explores the benefits of exploiting publicly available Google mobility data in studying the efficacy of COVID-19 policies and predicting future trends of disease spread. In part, this research is motivated by the controversy that occurred when the Public Health Agency of Canada (PHAC) was accused by some media outlets of engaging in mass surveillance of Canadians through cell-phone data, thereby compromising their privacy. This was not true. An investigation by the Office of the Privacy Commissioner confirmed that the data obtained by PHAC was aggregated and below a threshold of concern, with respect to preserving privacy. Concerns raised by the media can be attributed to poor data literacy in terms of understanding not only when data can be privacy protected through de-identification and aggregation, but also with respect to how the analysis of data can lead to significant public benefits. In this case, population mobility data can yield significant insights on policy efficacy and disease spread, which I demonstrate through analyzing available mobility data obtained from Google for the Ontario region. Results from relatively simple statistical models reveal correlations between more stringent policy restrictions on mobility and lower daily cases. The findings suggest that retail store focused lockdowns curbed the spread of infection. Given the simplicity of models and availability of free coding software that can be employed to conduct such analysis, universities and colleges should consider having publicly available courses that will contribute to expanded population data literacy. Further, governments at all levels should implement minimum levels of statistical training for relevant job positions.

Introduction and Problem Statement

COVID-19 revealed significant deficiencies in data collection and sharing protocols by governments across the world. Despite the occurrence of SARS in 2003 and H1N1 influenza in 2009, Canada in 2020 still did not have uniform data collection and reporting protocols between providers (physicians, hospitals) and across provinces. There were also several warnings from public health professionals, scholars, and the Auditor General of Canada, on Canada's dysfunctional system of public health data sharing (Attaran and Houston, 2020). It was left to a Ph.D. student (Isha Berry) along with other epidemiologists from U of T to construct a Github that tracked daily cases, deaths, and hospitalizations across provinces.[1] In response to these and other long-standing concerns on the collection and sharing of health data across the country, in 2020 the Public Health Agency of Canada announced the Pan Canadian Health Data Strategy (PCHDS), which established an Expert Advisory Group tasked with providing guidance on modernizing national health data collection, sharing and interoperability, streamlining and updating the approach to privacy and access to health data, and constructing standards for accountability and health data governance.[2]

Many lives lost to COVID-19 could have potentially been saved with more precise data and better availability. In its introduction, the First Report of the PCHDS (June 2021) noted that responses by various levels of public health to the pandemic were severely limited as a result of "significant gaps" in Canada's health data ecosystem. In 2022, the Federal Government communicated to

---

[1] The data can be accessed at https://github.com/ccodwg/Covid19Canada.

[2] https://www.canada.ca/en/public-health/programs/pan-canadian-health-data-strategy.html.

provincial and territorial governments that future funding on healthcare would be tied to their agreement of sharing health data and contributing to a national health data system.[3]

On February 7th 2023, the Prime Minister released a letter announcing that the federal government would increase health funding to provinces and territories by $196.1 billion over 10 years, including $46.2 billion in new funding.[4] As noted by Duong (2023), although the federal government's offer fell short of the original demands made by provinces, most provincial and territorial governments quickly accepted it, along with conditions involving data collection and sharing, and measures towards accountability. Specifically, the provinces and territories will receive funds under the condition of committing to improving how health system data are "collected, shared, used and reported" to promote transparency and manage public health emergencies, and also developing plans detailing steps implemented to measure progress (Duong, 2023).

With a view to building awareness of the public health benefits that can accrue from relatively straightforward statistical modelling, this paper will discuss some of my recent work on the effects of different policies on daily COVID-19 cases and appropriate forecasting models. An objective of this research is to demonstrate what can be accomplished with real-time data reporting and sharing, and relatively simple statistical models. However, while the widespread availability of relatively inexpensive data collection, sharing, and coding software makes real time collection, reporting, and Artificial Intelligence (AI) based analytics feasible, the gains from better data sharing and analysis become redundant with low data literacy. Poor data and AI literacy can impede the adoption of better policies as well as resulting in the loss of insights that have the potential to save lives. I define data literacy as basic knowledge of: (1) differences in types of data; (2) how insights can be generated from data using a variety of statistical methods; and (3) how individual privacy can be reasonably protected through de-identification and aggregation. AI literacy refers to knowledge of AI methods that can be used for statistical analysis and data visualization.

In this respect, a good example is the inaccurate reporting that appeared in many Canadian newspapers and outlets relating primarily to the use of population mobility data – derived from de-identified data aggregated across individual cellphone subscribers - by the Public Health Agency of Canada (PHAC). Sensationalistic headlines (such as '[Canada's public health agency admits it tracked 33 million mobile devices during lockdown](https://nationalpost.com/news/canada/canadas-public-health-agency-admits-it-tracked-33-million-mobile-devices-during-lockdown)') gave the impression that the federal government was violating the individual privacy of Canadians by attempting to monitor movements of individuals and using cellphone-based mobility data, which they did not have permission to access.[5] Nothing could be further from the truth. PHAC was interested in the ability of mobility data to yield insights on the determinants of COVID-19 spread and persistence. Understanding patterns in disease movements undoubtably has significant public health benefits in terms of efficient policy responses, such as guiding resources to areas with high levels of COVID-19 incidence. While much of the media coverage was inaccurate, a positive spillover was an inquiry into the matter and a subsequent report by the House of Commons Standing Committee on Access to Information, Privacy, and Ethics (Hereafter, The Report (2022). The Report has made several important recommendations relating to modernizing Canada's privacy laws, and an examination of the processes relating to the use of privately collected data by the federal government and improving population data literacy.

Improved data literacy is also one of the key recommendations contained in the Third Report of the Expert Advisory Group (May 2022), focusing on how to move forward with respect to national data collection and sharing. A summary of the recommendations are: (1) the need to ensure respect for, and alignment with, First Nations, Inuit, and Métis Nation data principles and strategies, which includes (but is not restricted to) control over their own data to develop evidence-informed approaches to health for their populations across Canada, and when data about First Nations, Inuit, and Métis Nation populations are used by other jurisdictions, that that use is collaborative, transparent, and ultimately guided by Indigenous communities; (2) establishing an independent Health Information Stewardship Council to advise, design, and drive investments in health data in coordination with First Nations, Inuit, and Métis Nation representative bodies and that can recommend clear data standards and ensure interoperability of data across

---

[3] See for example, 'Federal Health Minister seeks national health data system in exchange for increased funding to provinces', by Andrea Woo and Justine Hunter, published November 7th, 2022, and available at https://www.theglobeandmail.com/canada/article-federal-health-minister-seeks-national-health-data-system-in-exchange/. Also see https://www.pm.gc.ca/en/news/news-releases/2023/02/07/working-together-improve-health-care-canadians

[4] For further details please see 'Working together to improve health care for Canadians', available at https://www.pm.gc.ca/en/news/news-releases/2023/02/07/working-together-improve-health-care-canadians and Duong (2023). A key point in the Prime Minister's letter was the need for access to a patient's own electronic health information that is shared between the health professionals they consult.

[5] Please see 'Canada's public health agency admits it tracked 33 million mobile devices during lockdown.
The Public Health Agency of Canada accessed data such as cell-tower location to monitor people's activity during lockdown, it said' by Swikar Oli, published Dec 24, 2021, and available at https://nationalpost.com/news/canada/canadas-public-health-agency-admits-it-tracked-33-million-mobile-devices-during-lockdown. Also refer to 'Canada's privacy watchdog probing health officials' use of cellphone location data', by Alex Boutilier , Global News, posted January 11, 2022, and available at https://globalnews.ca/news/8503895/watchdog-probing-officials-cell-location-data/.
And 'Canadians' trips to liquor stores, pharmacies tracked via phones during COVID-19 pandemic', by Justin Tang PUBLISHED MAY 5, 2022, and available at https://www.theglobeandmail.com/canada/article-canadians-trips-to-liquor-stores-pharmacies-tracked-via-phones-during/.

jurisdictions, specifically through common standards for data linkage and architecture; (3) establishing a harmonized pan-Canadian health data policy framework for data stewardship that mandates appropriate data exchange, re-use, and collaboration across critical health data supply chains, balanced with security, confidentiality, and respect for privacy; (4) coordinating meaningful and ongoing engagement with the public and different stakeholders to understand their health data needs and expectations, which will assist in building public trust; and (5) consistent with the need to develop high levels of public trust and engagement, improving population level health data literacy, and demonstrating public health benefits of data collection and analysis, along with defining professional standards for critical health data roles.

Analyzing all these recommendations is beyond the scope of this paper; therefore, I shall focus on the importance of data literacy. Given how low data literacy can undermine the efficacy of policies with strong public health benefits, I shall attempt to explain the combined findings from some of my recent papers that evaluate the impacts of COVID-19 policies from a relatively non-technical perspective. My hope is that this study can be understood by a large segment of the population to build awareness of the type of statistical research that can lead to credible findings and the reasonable expectations that individuals should have on how their data can be safely used for public good. While I cannot completely abstain from discussing results based on specific advanced statistical models, I will refrain from employing unnecessary notation and advanced concepts. I shall also discuss different ways in which population level data literacy can be increased, which is critical not only from the perspective of addressing issues related to the efficacy of government communication but increasing trust in government as well.

PHAC and Population Mobility Data

Given the associated controversy, it is important to provide some background on the use of population mobility data by the Public Health Agency of Canada. As explained by Wolfson (2021), public mobility measures can be constructed from matching cell-phone movements to the nearest cell tower on a provider's network, and then calculating the percentage of phones near a single tower. This information can be used to produce aggregated data on population mobility consisting of visits to certain destinations (such as to retail outlets, pharmacies, and grocery stores) for specific geographic areas. Such data can then be used to evaluate the efficacy of lockdown measures that are intended to discourage movements outside individual residences and reduce the spread of COVID-19.

Individual data privacy can be accomplished through different methods. A typical first step is 'anonymization' which involves stripping away any specific personal identifiers (such as name, address, date of birth) that can be used to directly identify an individual. However, rigorous anonymization methods must go beyond removing direct identifiers. For example, imagine a dataset which does not include the name or address of individuals. On the other hand, it includes the number of visits to a specific physician's office. Proper anonymization techniques must ensure that the person's identity cannot be uncovered by combining this information with other external data that might be publicly available. A second recommended strategy would be to aggregate data across individuals through appropriate summary statistics. While data aggregation does not guarantee that such information cannot be unravelled to access individual specific details, the level of aggregation employed must be reasonable enough to minimize such risks, which is usually a function of sample size. A trade-off to data aggregation is that the richness of analysis that can be conducted with individual data, may be lost.

In its investigation of the use of mobility data by PHAC, the Office of the Privacy Commissioner (2023) details how PHAC used and accessed different data sources.[6] Besides data from TELUS, PHAC also obtained de-identified mobility data from Bluedot (https://bluedot.global/), a private sector firm specialized in data and analysis relating to disease outbreaks and spread. These data were based on apps downloaded by individual cell-phone users, which used GPS to track movements. PHAC did not have direct access to either the TELUS or Bluedot data. Another Government of Canada organization that specialized in the analysis of telecommunications data - Communications Research Centre Canada ("CRC") – processed the TELUS Data and generated mobility reports to provide aggregated data and statistics to PHAC's scientists for analysis. Analogously, Bluedot constructed aggregated mobility data that were either uploaded directly to PHAC's cloud or included into written reports that are sent by email to PHAC. Authorized individuals from PHAC could also access the information through a "Mobility Dashboard" developed by Bluedot. Therefore, it is unsurprising that in its investigation of the use of mobility data by PHAC, the Office of the Privacy Commissioner concluded that the combination of de-identification measures and policies implemented by PHAC and its data providers, as well as the reliance on aggregated as opposed to individual level data, reduced the possibility of re-identification below its "serious possibility" threshold.[7] The Report (2022) also agreed with this assessment, specifically stating in its summary;

"Other witnesses agreed that based on the publicly available evidence on the PHAC case, it did not appear that the government had used anything other than properly deidentified data to assess mobility patterns."

---

[6] Please see https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-federal-institutions/2022-23/pa_20230529_phac/.

[7] Please see https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-federal-institutions/2022-23/pa_20230529_phac/.

The data aggregation strategy employed by TELUS was no different than that used by Google to release aggregate population mobility measures based on individual use of its Google Maps app, and that were released for public use and research.[8] I used these data myself, to evaluate the impacts of public mobility on daily COVID-19 cases and for forecasting models (Stevens et al., 2021). Further, TELUS takes great care to ensure the privacy of individuals through practices such as not sharing real time mobility patterns, which have a greater likelihood of de-identification when matched with other external datasets/information.[9] The data aggregation relied upon by both TELUS and Bluedot resulted in geographic specific averages of public movements that could be used to analyze the relationship between population mobility and disease incidence.[10]

From a privacy perspective, it is important to emphasize that PHAC did not have explicit permission from each individual subscriber to access their data. However, they did not have to, under the conditions of TELUS's 'Data for Good' initiative.[11] Through this initiative, TELUS aggregates individual level data that can be used to compute population mobility measures for a certain geographic area. The objective of this program is to stimulate study and research on important societal issues that need mobility data. Any individual subscriber who does not want their data to be used for such purposes can request to opt-out of the program. This opt-out approach is akin to an 'implied consent' strategy. However, this is not the only method through which data vital to public health research might be obtained. Further research is required in the area of comparing the costs and benefits of opting-out of public health relevant data collection relative to having 'explicit consent or permission', which is the most common strategy employed by app providers to collect data from individuals in return for certain services, which may or may not be free. The Report (2022) recommended that Canadians be given the option to opt out of the data collection and that instructions for the method for opting out must be easily understood, widely communicated, and remain publicly available. Finally, it is important to acknowledge that both the Investigation by the Office of the Privacy Commissioner and Report of the Standing Committee concurred that PHAC could have been more transparent in its use of mobility data.

Research Findings

This section summarizes findings from three of my recent published papers that use population mobility data generated by Google. The mobility data being employed in this research have been extracted from the location history associated with Google Maps app use. The information was made available for use by researchers and policymakers through Google's 'COVID-19 Community Mobility Reports' that can be downloaded from https://www.google.com/covid19/mobility.

The mobility data made available through the public website captures total visits to the following specific destinations commonly frequented by individuals: (1) grocery and pharmacy stores, including grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies; (2) parks, which consist of local parks, national parks, public beaches, marinas, dog parks, plazas, and public gardens; (3) transit stations, comprised of subway, bus, and train stations; (4) retail stores & recreation outlets consisting of places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters; and (5) workplaces. The Google mobility measure with respect to places of residence, represents duration of stay and not total visits to different residences. A benefit to using Google indicators is that they are specific to mobility at different destinations such as retail stores and groceries, and therefore, have the ability to yield insights on the type of mobility that is most correlated with increases in daily cases. Values for each day are aggregated across individuals who have enabled their location history. As explained on its website, these values are calculated relative to the baseline, which is the median for the corresponding day of the week, during the 5-week period Jan 3–Feb 6, 2020. Hence, each daily value is the percentage change in the social mobility category relative to its baseline, which shows how visits and length of stay at different destinations have changed since the onset of the pandemic.

As detailed on its website, Google takes care to ensure that the data are de-identified/anonymized and then the information is aggregated across users who have turned on the Location History setting of their individual Google accounts on their phones and

---

[8] The data can be accessed at https://www.google.com/covid19/mobility/.

[9] Please see and https://www.telus.com/en/about/privacy/data-for-good/commitments for more information.

[10] Please refer to the Report of the Standing committee (2023) for a detailed discussion on how TELUS and Bluedot ensured that their data were properly de-identified an aggregated for PHAC.

[11] Please see https://www.telus.com/en/about/news-and-events/media-releases/telus-data-for-good-program-to-provide-de-identified-network-mobility-data-and-insights and https://www.telus.com/en/about/privacy/data-for-good/commitments.

have agreed to share this information. Hence, in contrast to TELUS's data, the Google mobility data are generated through explicit consent. However, a limitation to acknowledge is that Google has not made public its precise methodology for calculating social mobility. Further, the Google data are based on a sample of users who own mobile devices and who chose to share their location history, and the resulting data therefore may not be representative of the population. Of course, all data obtained from a specific source suffer from the same shortcoming.

Sen (2021) empirically investigates the relationship between daily COVID-19 cases across the twelve largest Public Health Units (PHUs) from May 2020 - December 2020. The first approach is to employ a linear regression model (Cochrane-Orcutt correct for autocorrelation) to assess the effects of: one and two day lagged cases; and seven-day lags in retail and grocery mobility with respect to daily COVID-19 cases for each of the twelve PHUs. The second contribution is to evaluate the forecasting performance of these simple models through predictions on daily cases from November 16th – December 15th. I constructed forecasts with data updating on a weekly basis, and the corresponding Root Mean Square Errors (RMSEs), and the Mean Absolute Percentage Deviations (MAPDs) are reported for each Public Health Unit.[12] The results for the six largest PHUs are reproduced in the below table. The daily COVID-19 cases are from https://www.publichealthontario.ca/en/data-and-analysis/infectious-disease/covid-19-data-surveillance/covid-19-data-tool.

TABLE 1. Based on Table 3 in Sen (2021)

Large PHUs	**Hamilton**	**Durham**	**Ottawa**	**Peel**	**Toronto**	**York**
One Day Lagged Cases	0.670 (0.121)^{a}	0.58 (0.107)^{a}	0.595 (0.145)^{a}	0.641 (0.206)^{a}	0.604 (0.254)^{a}	0.606 (0.149)^{a}
Two Day Lagged Cases	0.282 (0.116)^{b}	0.397 (0.104)^{a}	0.331 (0.135)^{b}	0.371 (0.206)^{c}	0.392 (0.250)^{c}	0.416 (0.149)^{a}
Seven Day Lagged Retail Mobility	0.041 (0.052)	0.065 (0.043)	0.243 (0.126)^{c}	0.418 (0.197)^{b}	0.776 (0.339)^{b}	0.284 (0.129)^{b}
Seven Day Lagged Grocery Mobility	-0.0498 (0.061)	-0.064 (0.061)	-0.182 (0.157)	-0.656 (0.267)^{b}	-0.807 (0.484)^{b}	-0.480 (0.178)^{a}
Weekend Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R Square	0.8040	0.8423	0.8509	0.9803	0.9775	0.9441
Root Mean Square Error (RMSE)	20.97	24.63	9.39	41.092	61.66	36.58
Mean Absolute Percentage Deviation (MAPD)	21.92%	22.61%	20.33%	7.11%	6.24%	16.79%

Notes: The regressions in this table are based on data from May 1st - November 15th consisting of 229 observations for each Public Heath Unit. Regression estimates are obtained from Cochrane-Orcutt transformations. Standard errors are in parentheses below coefficient estimates. a, b, and c denote statistical significance at the 1%, 5%, and 10% levels. Root Mean Square Errors (RMSE) and Mean Absolute Percentage Deviations (MAPDs) are constructed with respect to data from November 16th – December 15th.

The above results suggest that increases in grocery mobility are significantly correlated with reductions in daily cases for the largest PHUs of Peel, Toronto, and York. This is to be expected given that grocery stores were among the few places that could be consistently accessed by the public during this time-period. The high Adjusted R Square values demonstrate that the simple linear specifications fit the data quite well. What is perhaps surprising is the extremely strong forecasting performance of the basic linear regression models for the largest PHUs of Peel and Toronto. On average, the daily forecasts are only 6%-7% off actual values.

Table 2 below contains linear regression or OLS estimates of equation (1) with the dependent variables in columns (1), (2), (3), and (4) being: total cases by specimen test collection date; close contact cases; community cases; and outbreak cases, respectively. Daily close contact, community based, and outbreak cases are also reported by date of specimen collection. The results in Table 2 constitute a sensitivity analysis to the findings contained in Table 1, as the data are for the entire province of Ontario. Further, I

---

[12] The MAPD is calculated by taking the percentage difference of each daily forecast from its actual value between November 16th-December 15th, converting it to an absolute value and then calculating the mean across this sample range.

employ different variations of the dependent variable. Close contact refers to confirmed COVID-19 cases resulting from close proximity with an infected person. A case is considered community transmission when the source of a positive test result cannot be confirmed because the person: did not travel; did not knowingly have contact with another infected person; and is not associated with an outbreak. An outbreak setting refers to cases in which COVID-19 is transmitted in a specific shared space or setting such as a workplace, long-term care home, or daycare.

The daily number of new confirmed cases (total, close contact, community, or outbreak) based on date of specimen collection is modeled as a function of one- and two-day lags of itself, one-day lag in the number of tests, seven-day lag in the Bank of Canada (BOC) Policy Stringency Index, and seven-day lagged values of population Google mobility at: retail and recreation venues; groceries and pharmacies; parks; workplaces; transit hubs; and residences. All Google population mobility variables are employed to minimize potential omitted variables bias. Right hand side variables are lagged to capture the time required for symptoms to manifest after transmission. Dummy variables representing each weekday and weekends are employed to account for day specific variation in mobility that remains invariant over time. The number of daily tests is an important control variable as differences in daily cases is also a function of the corresponding number of tests.

The effects of policies at the provincial level are measured through the Bank of Canada Policy Stringency Index created by Cheung et al. (2021). This index is based on the methodology of the Oxford COVID-19 Government Response Tracker (OxCGRT) developed by the University of Oxford's Blavatnik School of Government.[13] The index is comprehensive in capturing different policies aimed at restricting public mobility and include: school and university closures; workplace and office closures; public event cancellations and restrictions; restrictions on private gatherings; public transport closures; stay-at-home requirements; restrictions on intra-provincial travel (between cities or regions within a province); restrictions on international travel; restrictions on interprovincial travel (between provinces); enforcement mechanisms for individuals; enforcement mechanisms for firms; and public information campaigns.[14] The index ranges from 0 (no restrictions) to 1 (maximum restrictions).

---

[13] For more details on methodology please see Hale et al. (2020).

[14] Please refer to Cheung et al. (2021) for further details.

TABLE 2. OLS Estimates of the Effects of Non-Pharmaceutical Interventions (NPIs) on Daily COVID-19 Cases in Ontario (Time-Series Evidence)

	(1) Total Cases	(2) Close Contact Cases	(3) Community Cases	(4) Outbreak Cases	(5) Retail Mobility	(6) Groceries & Pharmacies Mobility	(7) Workplace Mobility	(8) Residential Mobility
Dependent Variable					0.251	0.161	0.115	0.209
- One Day Lag					(0.133)^{c}	(0.123)	(0.05)^{b}	(0.049)^{a}
Dependent Variable					0.362	0.247	-0.600	-0.437
- One Day Lag					(0.102)^{a}	(0.100)^{b}	(0.257)^{b}	(0.213)^{b}
COVID-19 Policy					-80.097	-87.234	-102.31	55.895
Stringency Index					(36.63)^{b}	(52.69)^{c}	(41.09)^{b}	(15.31)^{a}
COVID-19 Policy					21.443	41.523	-43.297	13.527
Stringency Index					(34.50)	(51.11)	(42.11)	(12.22)
One Day Lag								
Seven Day Lagged	-161.7	-120.97	-14.42	-157.84				
COVID-19 Policy	(131.6)	(49.68)^{b}	(40.81)	(89.19)^{c}				
Stringency Index								
One Day Lagged	0.001	0.0003	0.0004	0.000001				
Tests	(0.001)	(0.0004)	(0.0004)	(0.0005)				
One Day Lagged	0.603	0.52	0.694	0.428				
Cases	(0.085)^{a}	(0.076)^{a}	(0.062)^{a}	(0.090)^{a}				
Two Day Lagged	0.376	0.459	0.316	0.399				
Cases	(0.089)^{a}	(0.074)^{a}	(0.068)^{a}	(0.086)^{a}				
Seven Day Lagged	2.097	0.1450	0.867	-0.521				
Retail Mobility	(1.197)^{c}	(0.448)	(0.309)^{b}	(0.896)				
Seven Day Lagged	-0.413	0.202	-0.408	0.701				
Grocery Mobility	(0.831)	(0.332)	(0.197)^{b}	(0.778)				
Seven Day Lagged	0.454	-0.284	-0.0234	1.04				
Work Mobility	(1.032)	(0.363)	(0.2698)	(0.53)^{b}				
Seven Day Lagged	-1.155	-0.848	-0.945	-0.225				
Transit Mobility	(1.642)	(0.691)	(0.543)	(1.092)				
Seven Day Lagged	0.0339	0.0065	-0.026	0.066				
Park Mobility	(0.106)	(0.046)	(0.04)	(0.066)				
Seven Day Lagged	6.423	0.584	0.061	5.307				
Residential Mobility	(3.432)^{c}	(1.130)	(0.678)	(2.255)^{a}				
Day of Week Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R Square	0.9354	0.9287	0.9530	0.9064	0.8401	0.5039	0.7303	0.8195

Notes: The regressions in this table are based on aggregate Ontario data from April 2nd – September 30th, 2020. The dependent variables in columns (1), (2), (3), (4), (5), (6), (7), and (8) are: total cases; close contact cases; community cases; outbreak cases; Google retail mobility; Google groceries & pharmacies mobility; Google workplace mobility; and Google residential mobility. All regression estimates are obtained from OLS regression with robust standard errors that are also Newey-West corrected for first order autocorrelation. Standard errors are in parentheses below coefficient estimates. a, b, and c denote statistical significance at the 1%, 5%, and 10% levels.

The seven-day lagged BOC Index has the expected negative sign across all columns. However, it is only statistically significant with respect to close contact and outbreak cases, at the 5% and 10% levels, respectively. One-day lagged number of tests is statistically insignificant across all columns. In all columns, one- and two-day lagged cases are positively and significantly correlated with trends in daily cases at the 1% level. In column (1), when the dependent variable is total cases by specimen collection, seven-day lagged retail mobility is positively correlated with daily cases at the 10% level. Seven-day lagged residential mobility is also positive and statistically significant at the 10% level. The significance of retail mobility with respect to total cases is driven by community cases (column 3), while the significance of residential mobility is because of outbreak cases (column 4). With respect to these sources of infection, a 10-percentage point increase in retail mobility is associated with roughly a 21- and 9-case daily increase in total and community cases. Further, a 10-percentage point rise in residential mobility is associated with roughly a 5.3 daily case increase, respectively, in daily outbreak-based spread. Finally, all regressions have adjusted R-squares higher than 0.9.

These findings correspond with intuition. The positive coefficients of retail mobility with respect to community-based transmission support the implementation of retail focused lockdowns in order to curb the spread of infections. On the other hand, the statistical significance of the seven-day lagged Google residential mobility indicators with respect to outbreaks are consistent with actual events at long-term care facilities. These results also underscore the importance of estimating the effects of policy stringency on cases by source of infection, as while the coefficient estimate of the BOC Index is statistically significant with respect to total cases, it is negatively and significantly correlated with close contact and outbreak associated cases. Further, the magnitude of coefficient estimates of the BOC Index from columns 2 and 4, imply that effects of stricter policies can be substantial. Finally, it is also important to note the significance of certain Google mobility indicators even after controlling for policy stringency.

Columns (5)-(8) report results of basic OLS regressions with different dependent Google mobility indicators as dependent variables. The motivation is to explore the impacts of policy stringency on mobility. The results, thus far, establish a statistically significant correlation between stricter policies on mobility and certain types of COVID-19 mobility cases. These results are premised on the belief that stricter policies reduce mobility, which after a certain time lag, lead to lower daily case counts. Hence, it is important to investigate whether changes in the BOC Policy Index are correlated with Google mobility indicators, to alleviate concerns that results in columns (1)-(4) are the result of spurious correlations. Further, establishing whether stricter policies are associated with reduced social mobility is, of itself, an important finding.

The results in columns (5)-(8) are based on simple specifications with the BOC Index and day of week dummies as explanatory variables. Although the underlying model is simple, the adjusted R Square is above 0.7 with respect to retail and residential mobility. Coefficient estimates of the BOC Index are negative and statistically significant at the 1% level for retail, work, and groceries mobility, and positive for residential mobility. The largest coefficients are with respect to retail and recreation mobility followed by work mobility, which makes intuitive sense. The smaller coefficient estimate for groceries & pharmacies reflects the essential nature of these businesses, while the negative sign, I suspect, captures the fact that individuals did reduce their visits. The positive coefficient estimate of residential mobility reveals the greater time at home spent by the public, in response to stricter controls on mobility.

In tandem, the results in Table 2 imply that stricter policies on mobility do reduce population movements, which in turn, are correlated with lower daily case counts. In this respect, reducing population mobility at retail and recreational venues and workplaces is of critical importance. The statistical significance of the BOC Policy Index with respect to daily cases, (that are conditioned on the use of Google mobility indicators), probably captures the effects of other initiatives, such as the wearing of masks and public awareness initiatives on hand washing and sanitization, that impact the likelihood of case transmission, independent of mobility trends.

IV. Conclusion

The use of public mobility data by PHAC to understand the relationships between population movements and COVID-19 attracted considerable media coverage and concerns on whether the privacy of Canadians was being compromised by unwarranted government surveillance. In my opinion, this coverage was inaccurate and a result of poor data literacy. The objective of this paper is to discuss whether the data used in the analysis commissioned by PHAC was poorly de-identified, whether it could potentially reveal individual-specific information that Canadians might be unwilling to share with government agencies, and possible public health benefits from the use of such data for research purposes.

Based on obtainable publicly available information, I concur with the essence of the investigation conducted by the Office of the Privacy Commissioner (2023). Specifically, while PHAC did not direct access to the data, it ensured that the data were aggregated in a manner that individual privacy was well protected. Given the care that was taken in ensuring data security and individual privacy, from a public health perspective, the use of such data was well justified by the knowledge that the analysis on the data would reveal. While Google has not shared its methodology on how it created its own mobility measures, the data aggregation principles seem comparable to those employed by TELUS, in terms of only releasing aggregate data for large geographic areas, for which the risk to individual privacy was minimal. However, as noted in The Report (2022), it is also incumbent upon the federal government to have more wide-ranging discussions on its uses of these types of data and to have public consultations to ensure better awareness among Canadians. This would also help to minimize inaccurate media coverage.

I also discuss some of my own research findings from employing Google mobility data. The use of Google data also enables an evaluation of the types of mobility that are the most correlated with increases in daily COVID-19 cases in Ontario. I find evidence of a modest correlation between more stringent policy restrictions on mobility and lower daily cases. However, stricter policies are important, as they reduce public mobility, which ultimately leads to lower daily cases. Results based on aggregate Ontario data suggest that independent of changes in policy, increases in retail mobility can be linked with higher daily COVID-19 cases, and that therefore, retail store focused lockdowns can help curb the spread of infection. This effect is a result of the impacts of retail mobility on total and close contact cases. On the other hand, higher work and residential mobility is unsurprisingly associated with increases in outbreak cases. These findings have obvious public policy implications and are only possible through the use of aggregated public mobility variables.

What is needed from a policy perspective, are clear tests enshrined in the federal Privacy Act which give guidance on when the Government of Canada can employ privately collected data to conduct analysis aimed at public benefits. The guidance could also be stated clearly in guidelines issued by the Office of the Privacy Commissioner that address the hard-to-quantify aspects of social benefits and enhanced risks to individuals, which are always possible when data are collected, shared, and analyzed. Specific standards on data de-identification and aggregation must also be constructed. These points are consistent with recommendations from The Report (2022). Such standards would be very helpful in mitigating concerns on the use of such data for public policy analysis. Questioning the validity of aggregated data for scientific analysis in the interest of privacy, leads to the possibility of a strong ‘privacy chill’ that might have unintended consequences in deterring important research. In its second report, the Expert Advisory Group noted:

“The 'privacy chill' also hampered Canada's response to COVID-19. Scientists were unable to link crucial data about variants of concern to determine whether the patient with a positive COVID test had been vaccinated, where they were located, and other important details that would have benefited insights and not impacted the privacy of the individual. Scientists were not able to develop the detailed insight that could have improved Canada's response and potentially saved lives.”

Individual privacy is a basic human right. There must be strong and sufficient policy safeguards, which ensure that individual privacy is protected, and the likelihood of individuals being harmed from the use of their data is minimal. However, in its second report, the Expert Advisory Group also recommended that a pillar of health data governance be the prioritization of data use for the public good while ensuring its security. These two objectives are not mutually exclusive. Minimizing legitimate access to data by researchers results in significant lost knowledge and social benefits to society.

With respect to data literacy, governments in Canada at all levels have a large pool of employees who are trained in research methods and are able to engage in statistical research, communications, and public conversations. This may be the case for Statistics Canada but should be true for most government organizations. Telling the right stories with data is critical in building trust with governments. Further, government employees should be trained in asking questions on whether particular data should be collected, critically assessing both the algorithms being run and the potential spillover effects on society. These recommendations are consistent with observations from senior public servants. The Report (2022) noted that the Chief Public Health Officer of Canada, Dr. Theresa Tam, testified that public health in Canada is presently insufficiently capacitated to employ big data, and that more capacity, tools, and data platforms are needed to inform its decisions.

Population data literacy must be improved to lead to meaningful public data related conversations. A policy goal could be that all government employees with certain job titles (e.g., policy, program, research analyst, manager, director) take courses each year that teach and update such skills (according to job needs). Data literacy is required not only for governmental representatives but for elected officials as well. In my opinion, the conversation on improving population data literacy and skills is as important as deciding our best approaches to future pandemics. Policies that result in better data literacy will lead to an enhanced awareness on how facts evolve not from cherry-picked examples, but robust and well documented evidence that goes through debate and peer-review. Such policies would be a strong platform for increasing trust in government and ensuring that society is well-prepared and ready to work together should we ever face a future crisis like COVID-19.

References

Attaran, Amir and Adam Houston (2020). “Pandemic Data Sharing: How the Canadian Constitution Turned into a Suicide Pact” in Colleen M Flood, Vanessa MacDonnell, Jane Philpott, Sophie Theriault & Sridhar Venkapuram, eds, *Vulnerable: The Policy, Law and Ethics of COVID-19* (Ottawa: University of Ottawa Press)

Cheung, Calista, Lyons, Jerome. Madsen, Bethany, Miller, Sarah, Sheikh, Saarah. (20201). “The Bank of Canada COVID-19 Stringency Index: Measuring Policy Response across Provinces.” Staff Analytical Note 2021-1, available at https://www.bankofcanada.ca/2021/02/staff-analytical-note-2021-1/.

Duong, Diana (2023). “Provinces accept federal health funding deal.” *CMAJ* 195 (9) E344-E345.DOI: https://doi.org/10.1503/cmaj.1096041.

Expert Advisory Group Report 1:
Charting a Path toward Ambition PAN-CANADIAN
HEALTH DATA STRATEGY available at https://www.canada.ca/content/dam/phac-aspc/documents/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries/expert-advisory-group-report-01-charting-path-toward-ambition/pCHDS-first-EAG-report-en.pdf.

Expert Advisory Group Report 1: Charting a Path toward Ambition https://www.canada.ca/en/public-health/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries/expert-advisory-group-report-01-charting-path-toward-ambition.html. June 2021

Expert Advisory Group Report 2: Building Canada’s Health Data Foundation available at https://www.canada.ca/en/public-health/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries/expert-advisory-group-report-02-building-canada-health-data-foundation.html#annexd. November 2021

Expert Advisory Group Report 3: Toward a world-class health data system. Available at Expert Advisory Group Report 3: Toward a world-class health data system Final Report May 2022 https://www.canada.ca/en/public-health/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries/expert-advisory-group-report-03-toward-world-class-health-data-system.html.

Hale, T., Angrist, N., Cameron-Blake, E., Hallas, L., Kira, B., Majumdar, S., Petherick, A., Phillips, T., Tatlow. H., and S. Webster (2020). “Variation in Government Responses to COVID-19.” Blavatnik School of Government, University of Oxford, Working Paper No. 2020/032, available https://www.bsg.ox.ac.uk/research/publications/variation-government-responses-covid-19.

Office of the Privacy Commissioner of Canada (2023). “Investigation into the collection and use of de-identified mobility data in the course of the COVID-19 pandemic.” Available at https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-federal-institutions/2022-23/pa_20230529_phac/.

Report of the House of Commons Standing Committee on Access to Information, Privacy and Ethics (2022). “Collection and use of Mobility Data by the Government of Canada and Related Issues”. Available at https://www.ourcommons.ca/Content/Committee/441/ETHI/Reports/RP11736929/ethirp04/ethirp04-e.pdf.

Sen, Anindya (2021). “Using Google Data to Estimate the Effects of Regional Mobility on Daily COVID-19 Cases: Evidence from Ontario Public Health Units.” *Canadian Journal of Regional Science* , 44(3), pp. 167–175

Sen, Anindya Sen, John David Baker, Qihuang Zhang, Rishav Raj Agarwal, Jean-Paul Lam (2023). “Do more stringent policies reduce daily COVID-19 case counts? Evidence from Canadian provinces”, *Economic Analysis and Policy*, 78, pp. 225-242.

Stevens, Nathaniel T., Anindya Sen, Francis Kiwon, Plinio P. Morita, Stefan H. Steiner, Qihuang Zhang (2021). “Estimating the Effects of Non-Pharmaceutical Interventions and Population Mobility on Daily COVID-19 Cases: Evidence from Ontario.” *Canadian Public Policy* 48(1), pp. 144-161.

Wolfson, Michael (2022). “Privacy outrage over the use of cellphone data by public health officials is unwarranted.” The Globe and Mail, Published January 24th, available at https://www.theglobeandmail.com/opinion/article-privacy-outrage-over-the-use-of-cellphone-data-by-public-health/.