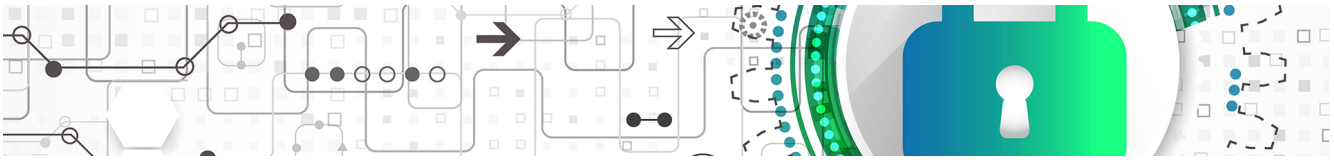

PANTHR: Ontario's Commendable Use of De-Identified Personal Health Information

April 14, 2020

By Jennifer McKenzie and Amanda Branch

This article provides an update on how the Ontario government is addressing the tension between the protection of individual privacy and the disclosure of information related to the COVID-19 pandemic. We previously touched upon the response from various Canadian regulators relating to the collection, use and disclosure of personal information in times of a pandemic.

"Personal information" is generally defined in federal and provincial law as information about an identifiable individual. Properly aggregated and anonymized data does not contain personal information and is therefore not subject to privacy legislation. The issue is how to de-identify an individual from his/her information so that information is truly anonymized and privacy is protected while also ensuring important data is available to those who need it.

Release of aggregated information by the Ontario government

In terms of information provided to the general public, the Ontario government releases a summary of cases, updated daily, which includes aggregated province wide data on confirmed cases of COVID-19 and includes details like age range (i.e. 19 and under; 20 – 39), gender, number of patients hospitalized, number of patients in the ICU and number of patients in ICU on a ventilator, and number of resolved cases. The province also releases a daily epidemiological summary (see an example here) which includes, among other things, a map which indicates a range of confirmed cases by geography (e.g., >500 confirmed cases in the Toronto region). It also includes a table of confirmed cases by public health unit (e.g., 592 confirmed cases reported by Ottawa Public Health).

It is clear that researchers need far more data than that which is provided to the general public. In recognition of this fact, according to an April 12, 2020 press release, the Ontario government, in consultation with The Office of the Information and Privacy Commissioner of Ontario ("IPC"), is developing a new health data platform called the Pandemic Threat Response (PANTHR) which is a new platform that "will hold secure health data that will allow researchers to better support health system planning and responsiveness, including the immediate need to analyze the current COVID-19 outbreak". PANTHR will gather de-identified data from publicly funded administrative health service records, such as physicians' claims to the Ontario Health Insurance Plan, drug claims submitted to the Ontario Drug Benefit Program and discharge summaries of hospital stays and emergency department visits.

This information was previously contained in silos. It is the hope that integrating the data will help researchers with modelling such as discovering risk factors, predicting when and where outbreaks may occur and identifying where to allocate equipment and other resources.

Christine Elliott, Ontario's Deputy Premier and Minister of Health, said: "Better access to integrated data will improve modelling and research to determine how COVID-19 is evolving, ensuring frontline staff are as prepared as possible in these unprecedented times". She added: "While access to data is important, we are taking all measures to ensure patient privacy is always respected and Ontarians are aware of how anonymized information may be shared."

Guidance from the Office of the Information and Privacy Commissioner of Ontario

Long recognizing the importance of de-identification, IPC has published a number of resources on the topic. In June 2016, IPC released the De-identification Guidelines for Structured Data (the "Guidelines"). The Guidelines define "de-identification" as "the general term for the process of removing personal information from a record or data set" and notes that de-identification is important because it protects the privacy of individuals.

De-identification doesn't completely eliminate the risk of re-identification; however, when done correctly, it does significantly reduce the risk. The Guidelines offer direction on taking a "risk-based approach" to de-identification, which involves calculating an acceptable level of re-identification risk for a given data release. This involves considering "prosecutor risk" (i.e., whether an "adversary" can know if a target individual is included in the data set) and "journalist risk" (i.e., if an adversary does not or cannot know if the target individual is in the data set). The Guidelines offer the following example of prosecutor risk:

> … if a teenager's parents know that their child has participated in a survey and the results are to be released in de-identified form, the risk of the parents attempting to re-identify their child's responses would qualify as prosecutor risk.

The Guidelines states that the amount of de-identification that needs to be applied to a data set is determined by how likely it is that an adversary will attempt to re-identify one or more individuals in the data set. Further, the context of any release must be considered and, specifically, public release of data requires the most stringent de-identification measures.

This position was emphasized in Order PO-3644, an appeal involving an access request made to the Ministry of Health and Long-Term Care. In this instance, the appellant, a media outlet, sought access to two de-identified records prepared by the Canadian Institute of Health Information ("CIHI") and submitted to the Ministry of Health and Long-Term Care. The Ministry denied access to the records stating the information was available to the public through CIHI. It relied on the exemption found in section 22(a) of FIPPA which states:

> A head may refuse to disclose a record where,
>
> (a) the record or the information contained in the record has been published or is currently available to the public

The appellant was unable to access the information through CIHI and appealed the Ministry's decision on the grounds that CIHI's disclosure of data to only approved requestors resulted in the records not being publicly available and therefore the section 22(a) exemption was not applicable.

The adjudicator noted section 22(a) is intended to provide institutions with the option to refer a requester to a publicly available source of information; however, in order to rely on this exemption, the records must be available to the public generally through a regularized system of access (e.g., unreported court decisions, property sale data, police accident reconstruction records, etc.). The Ministry also took the position that it would have to remove numerous data fields in order to ensure the records are properly de-identified and therefore, as a result, would be less detailed than what the appellant could obtain directly from CIHI.

The adjudicator found the Ministry's position to be consistent with the Guidelines and reiterated that public release of data requires the most stringent de-identification measures. The adjudicator noted that the fact the Ministry receives de-identified patient level data from CIHI for a limited purpose does not mean it can release the same information publicly without taking additional de-identification steps.

Conclusion

De-identification can be a complex and technically challenging process and should be considered by a team that includes legal counsel and technical experts. The PANTHR initiative is a commendable endeavor to help ensure researchers have access to de-identified information.

help you.