

Increasing Access to Health Administrative Data with ICES Data & Analytic Services

Lisa Ishiguro, Refik Saskin, Marian J. Vermeulen, Erika Yates, Nadia Gunraj and J. Charles Victor

Abstract

The Institute for Clinical Evaluative Sciences (ICES) is one of only a few organizations in Ontario permitted to access, link and analyze health administrative data for the purpose of generating evidence to inform decisions in policy and practice. Although ICES is a leading research institute, its access to the data has historically been restricted to scientists with an ICES affiliation. This requirement, designed to meet ICES' data privacy and security obligations, created barriers with respect to the widespread use of Ontario's data assets. In 2014, as part of the government's commitment to the Strategy for Patient-Oriented Research, ICES launched the Data & Analytic Services platform, which is aimed at increasing access to data and analytic services to investigators external to ICES. In making the data widely available to the broader research community, this initiative engages investigators involved in front-line care, stimulates new avenues of research and fosters collaboration that was previously challenging or unfeasible.

The Issue

Universal healthcare coverage and access to a wide variety of administrative data in Ontario allow for large population-based studies in many areas including diagnoses and treatment, outcomes and quality of care. Established in 1992, the Institute for Clinical Evaluative Sciences (ICES) houses a vast and secure array of linkable, encoded health-related data on more than 14 million Ontarians, including health services data, healthcare provider data, registries and population-based health surveys. Previously, research with health administrative data has been restricted to institutions recognized as a prescribed entity by the Office of the Information and Privacy Commissioner (IPC) of Ontario. This designation allows such institutions to access, link and analyze personal health information while adhering to strict policies and procedures in accordance with the Personal Health Information Protection Act (Information and Privacy Commissioner of Ontario 2016).

ICES is one of the few organizations that maintain prescribed entity status with the IPC and is subject to a rigorous triennial review. Therefore, analysis of health administrative data had previously been restricted to those with an affiliation at ICES; individuals and organizations external to ICES received access to only aggregate-level data.

In 2008, ICES launched the cancer data release program, cd-link. This collaboration between the Ontario Institute for Cancer Research, Cancer Care Ontario and ICES made the health administrative data housed at ICES more widely available and stimulated new research into the field of cancer health services research (Earle 2014). The cd-link program allows large sets of linked, encoded cancer registry data to be provided directly to investigators at Ontario academic institutions on an encrypted DVD delivered by a fully tracked courier. To date, over 42 data sets have been released to cd-link investigators, which has led to 34 publications for investigators who would not otherwise have had the opportunity to access linked data from ICES. Cd-link is a novel initiative for broadening access to linked health administrative data by investigators focused on cancer health services research. As per the agreement, cd-link investigators must either return to ICES all data to be destroyed upon completion of the term, or justify and enter into a new agreement.

The Response

In 2014, as part of Ontario's commitment to the Strategy for Patient-Oriented Research (SPOR) (Canadian Institutes of Health Research 2015) – a collaboration between government ministries and funding agencies – ICES built on cd-link and launched the Data & Analytic Services (DAS) platform. With a primary objective of increasing access to health administrative data for investigators outside ICES, DAS provides access to encoded, risk-reduced data sets created from ICES' data holdings; analytic support; and complete data analysis and report-writing services. DAS is also able to import researcher-collected data for linkage with ICES data holdings.

These services are available to publicly funded researchers, healthcare providers, healthcare administrators, policy makers and students. For Ontario-based researchers who benefit from support from the Ontario SPOR Support Unit, these services are currently provided on a cost-recovery basis. A modest increase is applied to the cost of these services for researchers outside of Ontario.

Interested investigators are required to complete a short request form and provide high-level information about the research question. The DAS agreement is executed between the principal investigator and ICES, outlining the scope, timelines and cost. Projects may also include a request for the importation of external data for linkage to ICES data. Once the external data have been transferred securely, a designated data covenantor uses an encrypted, unique identifier to link the imported data to ICES data holdings. All other personal health identifiers are removed from the data set prior to analysis.

According to the U.S. Health Insurance Portability and Accountability Act Privacy Rule (National Institutes of Health 2005), data can be de-identified by (1) the “safe harbour” method, which involves the removal of 18 variables including name, geographic subdivisions smaller than a state, all elements of dates (except for year) that are directly related to an individual, and telephone numbers; or (2) methods of statistical de-identification (US Department of Health and Human Services 2012). DAS uses both methods to de-identify and reduce the risk of re-identification of data.

The research-ready data sets are further evaluated for risk of re-identification using the Privacy Analytics Risk Assessment Tool (<www.privacyanalytics.ca>). This software uses algorithms to modify the data if necessary to maintain the risk of re-identification at an acceptable level. Its use in various scenarios has been demonstrated (El Emam et al. 2009a, 2009b, 2011, 2012). Examples of modification of variables include grouping age, scrambling or suppressing geographic variables or omitting a person’s sex. The data sets released must contain at least three persons who share the same characteristics.

Researchers who choose to perform their own analyses access a secure virtual desktop infrastructure via an encrypted Internet connection. The ICES Data & Analytic Virtual Environment (IDAVE) is fully compliant with requirements set out by the IPC. Multifactor authentication with a strong password is required to access the environment, and any external peripherals connected to the user’s local computer are disabled. Access is further restricted to directories and resources the users are authorized to access. External peripherals (printers, USB devices, etc.) connected to the user’s local computer are disabled and cannot be used within the environment. Furthermore, no data or documentation can be copied or transferred from the IDAVE without undergoing re-identification risk assessment as per ICES policies. Alternatively, projects seeking analytic services are led

by ICES scientists with clinical and subject matter expertise and are carried out by analysts and epidemiologists who are highly skilled in the use of ICES data and statistical methods.

There is a growing interest in comparing health services across provinces within Canada.

The Results

Since the launch of the DAS platform in March 2014, more than 200 requests from across Canada have been adjudicated, of which 187 have been deemed feasible and eligible for services. Requests have been submitted because of word of mouth, conference attendance and online presence (<www.ices.on.ca/DAS>). The majority of the investigators are Ontario-based with strong representation from Toronto (62%) and comparable representation from Ottawa (10%), Hamilton (9%) and Northern Ontario (9%). Research topics have varied widely and include assessments of healthcare provision by sector, chronic disease prevalence and treatment and statistical methods. Topics for population-based cohort studies have included assessing cardiovascular events associated with a specific drug therapy, exploring healthcare utilization and costs associated with smoking and examining complications following a prostate biopsy. Specific objectives have included assessments of existing practices in Ontario hospitals, such as the effectiveness of surgical safety checklists and the impact of specific quality-based procedures on resource use. Some statistical methods include econometric analyses to understand how surgical referral decisions are made and modelling a risk-adjusted parameter to estimate the performance of a current practice. Of projects deemed feasible and eligible, 136 have requested data access only and 24 have requested analytic services; 34 have requested the importation and linkage of external data to ICES data holdings. Typically, requests for access to the data are delivered within two to four weeks of contract execution. Requests for analytic services have varied greatly depending on the scope and complexity of the request. The response has been positive, and demand for data and analytic services remains high. Requests from repeat investigators are becoming increasingly complex including combinations of services, importation of investigator-collected data from across Canada and requests for access to restricted data sets.

Challenges

There is a growing interest in comparing health services across provinces within Canada. The launch of DAS allows investigators outside Ontario to request access to data or analyses to compare aspects of Ontario healthcare to their home province. One primary challenge in such projects includes the execution of data sharing agreements across jurisdictions, with laws of varying stringency governing the use and disclosure of data.

Understanding each of these laws to execute an agreement necessitates considerable consultation time, and thus increases the cost to investigators; this could be circumvented by having a national umbrella agreement in place that is governed by federal law. While healthcare policy makers would likely agree that facilitating cross-jurisdictional research would be beneficial to improving the quality of care nationwide, sharing personal health information, albeit de-identified, may also give rise to critics who would argue that such sensitive information should not be shared.

Lessons Learned

Based on the increasing number of requests, it is evident that the larger research community has a substantial interest in accessing the health services data of Ontarians for the purposes of research in population health, healthcare, program evaluation and quality improvement. By providing the research data in an accessible environment while ensuring its privacy and security, DAS has created a platform that is convenient and ideal for collaboration. As a result, it has been necessary to maintain a flexible pool of analysts and epidemiologists to account for variations in workload and increase staffing as necessary.

Access to restricted data sets was identified early on as a barrier for investigators. To address this, ICES is currently engaged in discussion with existing data partners to amend data sharing agreements to allow for external researchers to access these data as a part of DAS. Looking to the future, prospective data partners and stakeholders are increasingly interested in sharing their data with ICES for the purposes of linkage, further enhancing the rich data resources and knowledge sharing both internally at ICES and externally through DAS. **HQ**

Acknowledgements

The authors thank Dr. Craig Earle, director of the health services research program for Cancer Care Ontario and the Ontario Institute for Cancer Research, and a senior scientist at ICES, for establishing the cd-link program and for his work in broadening access to ICES data.

ICES Data & Analytic Services acknowledges its funders and partners: Canada's Strategy for Patient-Oriented Research (SPOR), the Ontario SPOR Support Unit, Canadian Institutes of Health Research, and the Government of Ontario.

References

Canadian Institutes of Health Research. 2015. "SPOR SUPPORT Units." Retrieved January 26, 2016. <<http://www.cihr-irsc.gc.ca/e/45859.html>>.

Earle, C. 2014. "Ontario Cancer Data Linkage Project: "cd-link".
Healthcare Quarterly 17(1): 11–13.

El Emam, K., D. Paton, F. Dankar and G. Koru. 2011. "De-identifying a Public Use Microdata File from the Canadian National Discharge Abstract Database." *BMC Medical Informatics and Decision Making* 11: 53.

El Emam, K., F.K. Dankar, R. Issa, E. Jonker, D. Amyot, E. Cogo et al. 2009a. "A Globally Optimal k-Anonymity Method for the De-identification of Health Data." *Journal of the American Medical Informatics Association* 16(5): 670–82.

El Emam, K., F.K. Dankar, R. Vaillancourt, T. Roffey and M. Lysyk. 2009b. "Evaluating the Risk of Re-identification of Patients from Hospital Prescription Records." *Canadian Journal of Hospital Pharmacy* 62(4): 307–19.

El Emam, K., L. Arbuckle, G. Koru, B. Eze, L. Gaudette, E. Neri et al. 2012. "De-identification Methods for Open Health Data: The Case of the Heritage Health Prize Claims Dataset." *Journal of Medical Internet Research* 14(1): e33.

Information and Privacy Commissioner of Ontario. 2016. *Personal Health Information Protection Act, 2004*. Retrieved January 26, 2016. <<https://www.ipc.on.ca/english/phipa/>>.

US Department of Health and Human Services. 2012. "Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule." Retrieved January 26, 2016. <http://www.hhs.gov/sites/default/files/oct/privacy/hipaa/understanding/coveredentities/De-identification/hhs_deid_guidance.pdf>.

About the Authors

Lisa Ishiguro, MSc, is a research project manager at ICES for Data & Analytic Services. The authors may be contacted at das@ices.on.ca.

Refik Saskin, MSc, is a staff scientist at ICES for Data & Analytic Services and an adjunct professor at the Institute of Health Policy, Management and Evaluation at the University of Toronto.

Marian J. Vermeulen, BScN, MHSc, is a staff scientist at ICES for the Applied Health Research Question portfolio.

Erika Yates, MSc, is a senior research project manager at ICES for the Applied Health Research Question portfolio.

Nadia Gunraj, MSc, is a senior analyst at ICES for Data & Analytic Services.

J. Charles Victor, MSc, PStat, is senior director, research and data, at ICES and a lecturer at the Institute of Health Policy, Management and Evaluation at the University of Toronto.

Breakfast with
the Chiefs

Longwoods.com/events